

Micro/Nano Technologies

Series Editors: Zheng You · Xiaohao Wang

SPRINGER  
REFERENCE

Qing-An Huang *Editor*

# Micro Electro Mechanical Systems

 Springer

---

# Micro/Nano Technologies

## **Series Editors**

Zheng You

Department of Precision Instrument

Tsinghua University

Beijing, China

Xiaohao Wang

Department of Precision Instrument

Tsinghua University

Beijing, China

The series consists of five volumes: *Micro/Nano Fabrication Technology*, *Micro Electro Mechanical Systems (MEMS)*, *Nanomaterial*, *Nanomedicine*, and *Applications of Micro-/Nanotechnologies* in IT. Experienced researchers and experts are invited to contribute in each of these areas. The series is published under Springer Major Reference Works, which allow continuous online update and publication. These features allow newcomers and other readers to keep in touch with the most up-to-date information in micro-/nanotechnologies. It presents an overview of the knowledge base, as well as selected topics, and provides comprehensive and authoritative information on the field for researchers, engineers, scientists, and graduate students who are involved in different aspects of micro-/nanotechnologies. This publication will provide inspiration for innovative research and application ideas for continued growth of the field. Micro-/nanotechnologies are techniques to fabricate and build systems with dimensions ranging from nanometers to microns. Advances in these technologies have created many interdisciplinary research opportunities and have been applied in various areas from nanomedicine to space systems. After several decades of development, a knowledge base of micro-/nanotechnologies covering all areas of research has now been created. It includes design and simulation, fabrication and manufacturing, packaging and assembly, device and system technologies, and application in different areas.

More information about this series at <http://www.springer.com/series/13903>

---

Qing-An Huang  
Editor

# Micro Electro Mechanical Systems

With 1122 Figures and 126 Tables

 Springer



*Editor*

Qing-An Huang

Key Laboratory of MEMS of the Ministry of Education

Southeast University

Nanjing, Jiangsu, China

ISBN 978-981-10-5944-5

ISBN 978-981-10-5945-2 (eBook)

ISBN 978-981-10-5946-9 (print and electronic bundle)

<https://doi.org/10.1007/978-981-10-5945-2>

Library of Congress Control Number: 2018936659

© Springer Nature Singapore Pte Ltd. 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

---

## Series Preface

Micro/nano technologies are those techniques to fabricate and build systems with dimensions ranging from nanometer to micron. The advancement of these technologies has been creating many interdisciplinary research opportunities and has been applied in various areas from nanomedicine to space systems. After several decades of development, a knowledge base of micro/nano technologies forms. It comes from every achievement of research and consists of design and simulation, fabrication and manufacturing, packaging and assembly, device and system technologies, and application in different areas.

Main structure of the knowledge base and selected topics are shown in this series of books. This work provides a comprehensive and authoritative knowledge of the field for researchers, engineers, scientists, and graduate students who are engaged in different aspects of micro/nano technologies. The series consists of five books, which are micro/nano fabrication, MEMS, nanomaterial, nanomedicine, and applications of micro/nano technologies in IT/IoT. In the books, the latest researches of micro-nano technology are summarized; the history and roadmap of related techniques, devices, and applications are discussed. The authors invited come from China, USA, Germany, Great Britain, Austria, Japan, etc. Their research activities in this area are remarkable and prospective. The readership is not merely specialists in this field but also a broad range of students and researchers from other related disciplines.

The Chinese Society of Micro/Nano Technology (CSMNT) has organized the Micro/Nano Technologies series in cooperation with Springer. This series would be the first systematic series of books in micro/nano technologies and provides a comprehensive and authoritative knowledge system. The series is published under Springer Major Reference, which allows continuous online update and publication. These features allow the newcomers and other content users to keep in touch with the most updated information in micro/nano technologies. It is expected that this publication will provide a new window to sparkle innovative research and application ideas for the continuous growth of the field.

Department of Precision Instrument  
Tsinghua University  
Beijing, China  
December 2017

Zheng You  
Xiaohao Wang

---

## Volume Preface

The acronym for micro electro mechanical systems (MEMS) was worldwide adopted in 1989. Since then, we have witnessed an increasing maturity of the MEMS industry and a rapid introduction of new products addressing applications ranging from smartphones, automobiles, aerospace, electronics instrumentation, industrial process control, appliances, biotechnology, healthcare, office equipment, and telecommunications to Internet-of-Things. MEMS devices and microsystems are now found everywhere.

In China, MEMS projects have been funded by the National Science Foundation of China (NSFC) since 1989, by the Major State Basic Research Program of China (973 Program) since 1998, and by the National High Technology Research and Development Program (863 Program) since 2000. The National Key Research and Development Program of China (the newly established national R&D program in 2016) continues to support the MEMS field. Hundreds of groups from Chinese universities/institutes have been involved in the MEMS research and development, and more than 200 MEMS companies have been set up in China until 2015. China has made great progress in the MEMS field. The aim of this volume is to introduce the research and development of MEMS in China, which has not yet been globally presented in a systematic way.

Because microfluidic chips are to be collected in Vol. 4: Nanomedicine and the processing and packaging of MEMS are to be collected in Vol. 1: Micro and Nano Fabrication Technology, this volume is divided into nine main parts consisting of 40 chapters.

Part 1 addresses the modeling of processing, structures, and devices in MEMS. Part 2 covers the characterization and test of materials and microstructures in MEMS. Part 3 describes physical sensors such as pressure sensors, flow sensors, wind sensors, accelerometers, gyroscopes, vacuum sensors, and electric current sensors. Part 4 comprehensively discusses chemical sensors including humidity, gas, trace energetic chemicals, and biochemical sensors. Part 5 presents micro-actuators for relays, inertial switches, microgrippers, and micropumps. Part 6 specifically describes RF MEMS devices including switches, microwave power sensors, reconfigurable attenuators, and acoustic devices for wireless communication. Part 7 focuses on uncooled infrared focal plane array. Part 8 is devoted to power MEMS with an emphasis on direct methanol fuel cells, piezoelectric vibration energy

harvesters, and triboelectric nanogenerators. Part 9 provides implantable medical devices in which particular attention is paid to microelectrode arrays.

This volume would not have been possible if the top academic authors had not taken some of their valuable time to do this work. And for this, I want to thank them most heartily. I would like to extend words of thanks and gratitude to the reviewers who selflessly gave time for carefully reviewing the chapters, to the Series Editors Zheng You and Xiaohao Wang for providing the vision in navigating the completion of the chapters, to the CSMNT secretary Jun Wang for her support in contacting the Series Editors and Springer Editors, and to the editorial team from Springer Nature, Li Shen, Lijuan Wang, Andrew Spencer, Stephen Yeung, Shanmuga Priya Kaliyamoorthy, and Pavithra Balakrishnan, for their patience in dealing with our unique and independent experts, and their persistence in contacting each of us in order to develop the logistics for the book publishing process.

It is my pleasure to serve as an editor for organizing the MEMS volume in Micro/Nano Technologies Series planned by the Chinese Society of Micro/Nano Technology (CSMNT). In turn, it is my hope that readers may enjoy the volume and gain further insights and full understanding of MEMS in China through this volume.

Key Laboratory of MEMS of the Ministry of  
Education, Southeast University  
Nanjing, Jiangsu, China  
December 2017

Qing-An Huang

---

# Contents

## Volume 1

<b>Part I Modeling of Processing, Structures, and Devices in MEMS</b> .....	<b>1</b>
<b>Modeling and Simulation of Silicon Anisotropic Etching</b> .....	<b>3</b>
Zai-Fa Zhou and Qing-An Huang	
<b>Modeling and Simulation of Silicon Dry Etching</b> .....	<b>27</b>
Yan Xing	
<b>Modeling and Simulation of SU-8 Thick Photoresist Lithography</b> .....	<b>67</b>
Zai-Fa Zhou and Qing-An Huang	
<b>Modeling of Electrostatically Actuated Microplates</b> .....	<b>99</b>
Libo Zhao, Zhuangde Jiang, Zhikang Li, and Yihe Zhao	
<b>Behavior Modeling and Simulation of an Inertial Sensor</b> .....	<b>155</b>
Honglong Chang, Jianbing Xie, and Ying Liu	
<b>Part II Characterization and Test of Materials and Microstructures in MEMS</b> .....	<b>195</b>
<b>Online Test Microstructures of the Mechanical Properties for Micromachined Multilayered Films</b> .....	<b>197</b>
Zai-Fa Zhou, Qing-An Huang, Xin-Ge Guo, and Yi-Fan Gu	
<b>Online Test Microstructures of the Thermophysical Properties of MEMS Conducting Films</b> .....	<b>237</b>
Hai-Yun Liu, Zai-Fa Zhou, and Qing-An Huang	
<b>Fracture Properties of MEMS/NEMS Thin Films</b> .....	<b>303</b>
Jinling Yang and Quan Yuan	

<b>Part III Physical Microsensors</b> .....	<b>323</b>
<b>High Temperature Silicon Pressure Sensors</b> .....	325
Zhuangde Jiang, Yulong Zhao, Libo Zhao, and Tingzhong Xu	
<b>A Micromachined Silicon Resonant Pressure Sensor</b> .....	387
Junbo Wang, Deyong Chen, Bo Xie, Jian Chen, Lin Zhu, and Yulan Lu	
<b>A Micromachined Vibratory Gyroscope</b> .....	421
Qiancheng Zhao, Chunhua He, Jian Cui, and Guizhen Yan	
<b>A Double Differential Torsional MEMS Accelerometer with Improved Temperature Robustness</b> .....	461
Dingbang Xiao, Xuezhong Wu, Qingsong Li, and Zhanqiang Hou	
<b>Micromachined Gas Inertial Sensors</b> .....	517
Rong Zhu	
<b>A Micromachined Thermal Wind Sensor</b> .....	539
Zhenxiang Yi, Ming Qin, and Qing-An Huang	
<b>Integrated Vacuum Microsensor Systems in CMOS Technology</b> .....	577
Jiaqi Wang and Zhenan Tang	
<b>Micro Thermal Flow Sensor</b> .....	595
Rong Zhu	
<b>Passive MEMS DC Electric Current Sensors</b> .....	625
Dong F. Wang, Huan Liu, Xuesong Shang, Weikang Xian, Yipeng Hou, Xu Yang, Toshihiro Itoh, and Ryutaro Maeda	
<b>Volume 2</b>	
<b>Part IV Chemical Microsensors</b> .....	<b>657</b>
<b>Microcalorimeters for Detection of Trace Energetic Chemicals</b> .....	659
Zheyao Wang and Wenzhou Ruan	
<b>Biochemical Sensors Based on Piezoresistive Microcantilevers</b> .....	689
Xiaomei Yu and Rui Zhao	
<b>Micro-Heater-Based Gas Sensors</b> .....	717
Tie Li, Lei Xu, and Yuelin Wang	
<b>Picogram-Order Mass Sensors via Cantilever-Based Micro-/ Nanostructures</b> .....	753
Dong F. Wang, Xu Du, Xiaodong Li, Di Zhou, Cao Xia, Guowen Zheng, Shenglai Wan, and Xin Wang	

---

<b>Micromachined Humidity Sensors</b> .....	787
Jianqiu Huang and Qing-An Huang	
<b>Piezoelectric Micro/Nano Mechanical Devices for Frequency Control and Chemical Sensing</b> .....	817
Wei Pang, Menglun Zhang, and Ji Liang	
<b>Microsensors and Systems for Water Quality Determination</b> .....	847
Shanhong Xia, Jianhua Tong, Chao Bian, Jizhou Sun, and Yang Li	
<b>Part V Microactuators</b> .....	<b>905</b>
<b>Electrostatic Comb-Driven Actuator for MEMS Relay/Switch</b> .....	907
Jiahao Zhao and Yongfeng Gao	
<b>MEMS Actuators Driven by Lorentz Force</b> .....	929
Jinling Yang and Jinying Zhang	
<b>A MEMS Inertial Switch Based on Nonsilicon Surface Micromachining Technology</b> .....	945
Zhuoqing Yang, Guifu Ding, Yan Wang, and Xiaolin Zhao	
<b>A Rotary Microgripper</b> .....	997
Honglong Chang and Yongcun Hao	
<b>Traveling-Wave Micropumps</b> .....	1017
Guohua Liu and Wei Zhang	
<b>Part VI RF MEMS</b> .....	<b>1037</b>
<b>RF MEMS Switch</b> .....	1039
Li-Feng Wang, Qing-An Huang, and Lei Han	
<b>A Microwave Power Sensor</b> .....	1077
Zhiqiang Zhang and Xiaoping Liao	
<b>A Micromachined Reconfigurable Attenuator</b> .....	1117
Zewen Liu and Xin Guo	
<b>High-Performance Acoustic Devices for Wireless Communication and Sensor Applications</b> .....	1149
Changjian Zhou, Xiangguang Tian, and Tian-Ling Ren	
<b>Part VII Optical MEMS</b> .....	<b>1195</b>
<b>Silicon-Based Optical Sensor: Uncooled Infrared Focal Plane Array Based on Bi-Materials Cantilever Microstructures</b> .....	1197
Tianchun Ye, Dapeng Chen, Shali Shi, and Binbin Jiao	

---

<b>Part VIII Power MEMS</b> .....	<b>1265</b>
<b>MEMS Direct Methanol Fuel Cells</b> .....	1267
Yufeng Zhang, Weijian Yuan, Rui Xue, and Xiaowei Liu	
<b>MEMS Piezoelectric Vibration Energy Harvesters</b> .....	1297
Licheng Deng, Zhiyu Wen, and Xingqiang Zhao	
<b>Triboelectric Nanogenerators</b> .....	1335
Chi Zhang and Zhong Lin Wang	
<b>Part IX Implantable Medical Devices</b> .....	<b>1377</b>
<b>Microelectrode Array</b> .....	1379
Renxin Wang, Huaiqiang Yu, and Zhihong Li	
<b>Electrodes for Nerve Recording and Stimulation</b> .....	1413
Jing-Quan Liu, Hong-Chang Tian, Xiao-Yang Kang, and Ming-Hao Wang	
<b>Electrode Array for Neural Interfaces</b> .....	1437
Weihua Pei and Hongda Chen	
<b>Index</b> .....	1467



---

## About Series Editors



**Zheng You** is currently a professor in the Department of Precision Instrument in Tsinghua University, Beijing, China. He is also serving as the director of Micro-Nano Technology Research Center at Tsinghua University, the director of State Key Laboratory of Precision Measurement Technology and Instruments, and vice president of National Institute for Nano-Technology and Engineering. He has been elected as an academician of Chinese Academy of Engineering since 2013 and president of Chinese Society of Micro-Nano Technology (CSMNT).

Prof. Zheng You's main research interests fall in micro-nano technology, MEMS, satellite technology, and measurement and instruments.

Prof. You has published 322 papers and 32 research reports, holds 12 Chinese invention patents, among which over 114 papers are retrieved and embodied by the SCI or EI journals. He developed the first micro satellite in China and rewarded with numerous awards including the Second-class National Scientific and Technical Advance Awards, the First-class Scientific and Technical Awards of China Instrument and Control Society, the First-class and the Second-class Scientific and Technical Advance Awards of the Education Ministry, the First-class and Second-class Scientific and Technical Advance Award of the Beijing Municipality, and Second-class Science Technology and Industry Awards for National Defense.



**Xiaohao Wang** is currently a professor in the Department of Precision Instrument, Tsinghua University. He also serves as associate dean of Education and director of the Division of Advanced Manufacturing, Graduate School at Shenzhen, Tsinghua University, and executive secretary general of the Chinese Society of Micro-Nano Technology (CSMNT).

Prof. Xiaohao Wang received a bachelor degree and a Ph.D. both from Tsinghua University in 1994 and 1999, respectively. During 1998 to 2001, he joined the Department of Precision Instrument in Tsinghua University as an assistant professor and was promoted to associate professor in 2001. During 2007 to 2008, he was a visiting scholar at Technische Universitaet Berlin. In 2010, he became a professor of the Department of Precision Instrument in Tsinghua University.

Prof. Xiaohao Wang's research interests cover MEMS-based sensors, actuators, ionizing sources and portable mass spectrometer, ion mobility spectrometer, and high-field asymmetric waveform ion mobility spectrometer. He has published over 200 technical papers and applied tens of patents in the area of MEMS and the portable instrumentation.

---

## About the Editor



**Qing-An Huang** received the B.S. degree from the Hefei University of Technology, Hefei, China, in 1983; the M.S. degree from Xidian University, Xi'an, China, in 1987; and the Ph.D. degree from Southeast University, Nanjing, China, in 1991, all in electronics engineering. His Ph.D. research was focused on micromachined GaAs piezoelectric sensors.

He joined the Department of Electronic Engineering, Southeast University, after graduation, where he became a Full Professor in 1996, and was appointed as the Chang-Jiang Scholar by the Ministry of Education in 2004. He was a visiting scholar at the Hong Kong University of Science and Technology from 1997 to 1998, and a senior visiting scholar (1 month) at the Berkeley Sensor and Actuator Center, University of California at Berkeley in 2017. He has been the Founding Director of the Key Laboratory of MEMS (the Ministry of Education) with the Southeast University since 2001. He has authored a book entitled *Silicon Micromachining Technology* (Science Press, 1996), authored or coauthored 4 international book chapters, over 200 peer-reviewed international journals/conference papers, and holds over 100 Chinese patents.

Dr. Huang was a TPC member of TRANSDUCERS from 2009 to 2017 and IEEE SENSORS Conference from 2002 to 2015. He was the conference Co-chair of the *SPIE Microfabrication and Micromachining Process Technology and Devices* (Proceedings of SPIE, Vol. 4601, 2001), the TPC Co-chair of the *7th IEEE International Conference on Nano/Micro Engineered and Molecular Systems* (IEEE NEMS, Kyoto, Japan, 2012), and the TPC Co-chair of the *6th Asia-Pacific Conference of Transducers and Micro/Nano*

*Technologies* (APCOT, Nanjing, China, 2012). He served as the Steering Committee Chair of APCOT from 2012 to 2014 and the Founding Chairman of IEEE ED/SSC Nanjing Chapter from 2008 to 2014. He has been serving as the Editor-in-Chief of the *Chinese Journal of Sensors and Actuators* since 2005. He is an Editorial Board Member of the *Journal of Micro-mechanics and Microengineering*.

Dr. Huang served as the Head of Expert Group of Micro/Nano Fabrication Branch under the National High Technology Research and Development Program of China during 2011 to 2015, the member of Expert Group of Microelectromechanical Systems Branch under the National High Technology Research and Development Program during 2001 to 2005, and the member of Evaluation Expert Group of the Information Department under the National Natural Science Foundation of China during 2010 to 2011.

Dr. Huang is a Fellow of the IEEE. He was a recipient of the National Outstanding Youth Science Foundation Award of China in 2003.

---

## Contributors

**Chao Bian** Institute of Electronics, Chinese Academy of Sciences, Beijing, China

**Honglong Chang** MOE Key Laboratory of Micro/Nano Systems for Aerospace, Northwestern Polytechnical University, Xi'an, China

Northwestern Polytechnic University, Xi'an, China

**Dapeng Chen** Institute of Microelectronics of Chinese Academy of Sciences (IMECAS), Beijing, China

**Deyong Chen** State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China

**Hongda Chen** State Key Laboratory of Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

**Jian Chen** State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China

**Jian Cui** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Beijing, China

**Licheng Deng** College of Electronic and Optical Engineering and College of Microelectronics, Nanjing University of Posts and Telecommunications, Nanjing, China

Defense Key Disciplines Lab of Novel Micro-Nano Devices and System Technology, Chongqing University, Chongqing, China

**Guifu Ding** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University (SJTU), Shanghai, China

**Xu Du** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, China

**Yongfeng Gao** Tsinghua University, Beijing, China

**Yi-Fan Gu** Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, China

**Xin Guo** Tsinghua University, Beijing, China

**Xin-Ge Guo** Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, China

**Lei Han** Southeast University, Nanjing, Jiangsu, China

**Yongcun Hao** Northwestern Polytechnic University, Xi'an, China

**Chunhua He** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Beijing, China

**Yipeng Hou** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, Jilin, China

**Zhanqiang Hou** National University of Defense Technology, Changsha, China

**Jianqiu Huang** Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, Jiangsu, China

**Qing-An Huang** Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, Jiangsu, China

**Toshihiro Itoh** Department of Human and Engineered Environmental Studies, University of Tokyo, Chiba, Japan

**Zhuangde Jiang** State Key Laboratory for Manufacturing Systems Engineering, Collaborative Innovation Center of Suzhou Nano Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China

International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, Xi'an Jiaotong University, Xi'an, Shaanxi, China

**Binbin Jiao** Institute of Microelectronics of Chinese Academy of Sciences (IMECAS), Beijing, China

**Xiao-Yang Kang** Shanghai Jiao Tong University, Shanghai, China

**Qingsong Li** National University of Defense Technology, Changsha, China

**Tie Li** Science and Technology on Micro-system Laboratory, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China

**Xiaodong Li** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, China

**Yang Li** Institute of Electronics, Chinese Academy of Sciences, Beijing, China

**Zhihong Li** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Beijing, China

**Zhikang Li** State Key Laboratory for Manufacturing Systems Engineering, Collaborative Innovation Center of Suzhou Nano Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China

International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, Xi'an Jiaotong University, Xi'an, Shaanxi, China

**Ji Liang** State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin, China

**Xiaoping Liao** Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, China

**Guohua Liu** Key Laboratory of Photo-electronic Thin Film Devices and Technology of Tianjin, College of Electronic Information and Optical Engineering, Nankai University, Tianjin, China

**Hai-Yun Liu** Department of Electronics and Information Engineering, Hohai University, Nanjing, Jiangsu, China

**Huan Liu** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, Jilin, China

**Jing-Quan Liu** Shanghai Jiao Tong University, Shanghai, China

**Xiaowei Liu** MEMS Center, Harbin Institute of Technology, Harbin, China

**Ying Liu** School of Arts and Sciences, Shaanxi University of Science and Technology, Xi'an, China

**Zewen Liu** Tsinghua University, Beijing, China

**Yulan Lu** State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China

**Ryutaro Maeda** Ubiquitous MEMS and Micro Engineering Research Center, AIST Tsukuba, Ibaraki, Japan

**Wei Pang** State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin, China

**Weihua Pei** State Key Laboratory of Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

CAS Center for Excellence in Brain Science and Intelligence Technology, Beijing, China

University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

**Ming Qin** Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, China

**Tian-Ling Ren** Institute of Microelectronics, Tsinghua University, Beijing, China

**Wenzhou Ruan** Tsinghua University, Beijing, China

**Xuesong Shang** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, Jilin, China

**Shali Shi** Institute of Microelectronics of Chinese Academy of Sciences (IMECAS), Beijing, China

**Jizhou Sun** Institute of Electronics, Chinese Academy of Sciences, Beijing, China

**Zhenan Tang** School of Microelectronics, Dalian University of Technology, Dalian, China

**Hong-Chang Tian** Shanghai Jiao Tong University, Shanghai, China

**Xiangguang Tian** Institute of Microelectronics, Tsinghua University, Beijing, China

**Jianhua Tong** Institute of Electronics, Chinese Academy of Sciences, Beijing, China

**Shenglai Wan** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, China

**Dong F. Wang** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, Jilin, China  
Research Center for Ubiquitous MEMS and Micro Engineering, AIST, Tsukuba, Japan

**Jiaqi Wang** Institute of Microelectronics, Dalian University of Technology, Dalian, China

**Junbo Wang** State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing, China  
University of Chinese Academy of Sciences, Beijing, China

**Li-Feng Wang** Southeast University, Nanjing, Jiangsu, China

**Ming-Hao Wang** Shanghai Jiao Tong University, Shanghai, China

**Renxin Wang** Science and Technology on Electronic Test and Measurement Laboratory, North University of China, Taiyuan, China

**Xin Wang** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, China



**Yan Wang** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University (SJTU), Shanghai, China

**Yuelin Wang** Science and Technology on Micro-system Laboratory, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China

**Zheyao Wang** Tsinghua University, Beijing, China

**Zhong Lin Wang** Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China

School of Material Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA

**Zhiyu Wen** Defense Key Disciplines Lab of Novel Micro-Nano Devices and System Technology, Chongqing University, Chongqing, China

**Xuezhong Wu** National University of Defense Technology, Changsha, China

**Cao Xia** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, China

**Shanhong Xia** Institute of Electronics, Chinese Academy of Sciences, Beijing, China

**Weikang Xian** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, Jilin, China

**Dingbang Xiao** National University of Defense Technology, Changsha, China

**Bo Xie** State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China

**Jianbing Xie** MOE Key Laboratory of Micro/Nano Systems for Aerospace, Northwestern Polytechnical University, Xi'an, China

**Yan Xing** Southeast University, Nanjing, China

**Lei Xu** School of Information Science and Technology, University of Science and Technology of China, Hefei, China

**Tingzhong Xu** State Key Laboratory for Manufacturing Systems Engineering, International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, Collaborative Innovation Center of Suzhou Nano Science and Technology, School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China

**Rui Xue** MEMS Center, Harbin Institute of Technology, Harbin, China

**Guizhen Yan** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Beijing, China

**Jinling Yang** Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

State Key Laboratory of Transducer Technology, Shanghai, China

School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing, China

**Xu Yang** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, Jilin, China

**Zhuoqing Yang** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University (SJTU), Shanghai, China

**Tianchun Ye** Institute of Microelectronics of Chinese Academy of Sciences (IMECAS), Beijing, China

**Zhenxiang Yi** Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, China

**Huaiqiang Yu** Research Institute, China Electronics Technology Group Corporation, Chongqing, China

**Xiaomei Yu** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Haidian District, Beijing, China

**Quan Yuan** Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

State Key Laboratory of Transducer Technology, Shanghai, China

**Weijian Yuan** MEMS Center, Harbin Institute of Technology, Harbin, China

**Chi Zhang** Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China

**Jinying Zhang** Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

State Key Laboratory of Transducer Technology, Shanghai, China

School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing, China

**Menglun Zhang** State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin, China

**Wei Zhang** Key Laboratory of Photo-electronic Thin Film Devices and Technology of Tianjin, College of Electronic Information and Optical Engineering, Nankai University, Tianjin, China

**Yufeng Zhang** MEMS Center, Harbin Institute of Technology, Harbin, China

**Zhiqiang Zhang** Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, China

**Jiahao Zhao** Tsinghua University, Beijing, China

**Libo Zhao** State Key Laboratory for Manufacturing Systems Engineering, Collaborative Innovation Center of Suzhou Nano Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China

International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, Xi'an Jiaotong University, Xi'an, Shaanxi, China

**Qiancheng Zhao** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Beijing, China

**Rui Zhao** Science and Technology on Electronic Test and Measurement Laboratory, School of Instrument and Electronics, North University of China, Taiyuan, Shanxi, China

**Xiaolin Zhao** National Key Laboratory of Science and Technology on Micro/Nano Fabrication, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University (SJTU), Shanghai, China

**Xingqiang Zhao** School of Information and Control, Nanjing University of Information Science and Technology, Nanjing, China

**Yihe Zhao** State Key Laboratory for Manufacturing Systems Engineering, Collaborative Innovation Center of Suzhou Nano Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China

International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, Xi'an Jiaotong University, Xi'an, Shaanxi, China

**Yulong Zhao** State Key Laboratory for Manufacturing Systems Engineering, International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, Collaborative Innovation Center of Suzhou Nano Science and Technology, School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China

**Guowen Zheng** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, China

**Changjian Zhou** School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China

**Di Zhou** Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, China

**Zai-Fa Zhou** Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, Jiangsu, China

**Lin Zhu** State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing, China  
University of Chinese Academy of Sciences, Beijing, China

**Rong Zhu** State Key Laboratory of Precision Measurement Technology and Instrument, Department of Precision Instrument, Tsinghua University, Beijing, China

---

**Part I**

**Modeling of Processing, Structures, and  
Devices in MEMS**



# Modeling and Simulation of Silicon Anisotropic Etching

Zai-Fa Zhou and Qing-An Huang

## Contents

Overview .....	4
Microscopic Etch Rates of Surface Atoms .....	5
Microscopic Activation Energies for Surface Atoms .....	15
Simulations and Discussions .....	20
Conclusions .....	23
References .....	23

## Abstract

Silicon-based wet anisotropic etching is one of the most widely used techniques for the realization of various kinds of components for the fabrication of micro-electromechanical systems (MEMS)-based devices. This technique uses liquid-based etchants to selectively etch the bulk silicon to fabricate the microstructures. Various geometric models or atomic-level models have been presented for silicon anisotropic etching simulations to optimize the anisotropic etching processes and improve the efficiency of MEMS design. This chapter develops a novel model for silicon anisotropic etching in alkaline solutions based on the surface atom configurations on different crystal planes. The surface atoms are divided into some categories according to the atom configurations, and the microscopic etch rates of surface atoms are related to the macroscopic etch rates of different silicon crystal planes. The microscopic activation energies for some typical surface atoms are further calculated by the microscopic etch rates for surface atoms, to simplify the fitting procedures. The model has been extended to a simulation systems based on a dynamic cellular automaton method, and a series of

Z.-F. Zhou (✉) · Q.-A. Huang

Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, Jiangsu, China

e-mail: [zfzhou@seu.edu.cn](mailto:zfzhou@seu.edu.cn); [hqa@seu.edu.cn](mailto:hqa@seu.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,  
[https://doi.org/10.1007/978-981-10-5945-2\\_1](https://doi.org/10.1007/978-981-10-5945-2_1)

simulations have been performed using the simulation system for various etching conditions. The simulation results demonstrate to be in agreement with the experimental results. This indicates the effectiveness of the model, and this is useful for the research of anisotropic etching technology and the development of MEMS design.

---

**Keywords**

Silicon · Anisotropic etching · Simulation · Modeling · Activation energy

---

**Overview**

Wet anisotropic etching of silicon is one of the most popular etching methods for silicon bulk micromachining for the fabrication of various types of microstructures such as cantilever, diaphragm, and cavity. Anisotropic etching of silicon has many advantages over other techniques including its orientation-dependent etching, low cost, and flexibility of batch fabrication which makes it an inevitable choice for industrial applications. This type of etching is widely employed for fabricating various MEMS structures for some advantages, such as simple etching equipment, and its capability to fabricate some special microstructures (Bean 1978; Petersen 1982; Madou 1997; Liu 2006). Various geometric models (Buser and Rooij 1991; Sequin 1992; Danel and Delapierre 1992; Fruhauf et al. 1993; Schroder et al. 2001; Radjenovic et al. 2006) or atomic-level models (Than and Buttgenbach 1994; Hubbard and Antonsson 1997; van Veenendaal et al. 2000; Nishidate et al. 1998; Zhu and Liu 2000; Kakinaga et al. 2004; Camon and Moktadir 1995; Gosalvez et al. 2001, 2002; Zubeľ 2003; Zhou et al. 2007, 2009) have been presented for silicon anisotropic etching simulations to optimize the anisotropic etching processes and improve the efficiency of MEMS design. These models, however, have not yet completely discovered the reaction mechanisms of silicon anisotropic wet chemical etching. Many experiments and techniques have been implemented to measure the macroscopic anisotropic etch rates of various silicon crystal planes in different etching conditions for engineering applications (Seidel 1990; Seidel et al. 1990; Herr and Baltes 1991; Tabata et al. 1992; Allongue et al. 1993; Yun et al. 1998; Sato et al. 1998; Wind and Hines 2000; van Veenendaal et al. 2001; Zubeľ 2001; Zubeľ and Kramkowska 2004; Gosalvez et al. 2007; Hubbard and Antonsson 1994). These macroscopic etch rates cannot be directly adopted for the atomic-level simulation of silicon anisotropic etching processes, since there are some differences between the macroscopic etch rates for silicon crystal planes and the microscopic etch rates for surface atoms located on silicon crystal planes (Zhou et al. 2007). Attempts have been made to relate the microscopic etch rates of surface atoms with the macroscopic etch rates of different crystal planes (Zhou et al. 2007, 2009). Although these studies only concern ideal (hkl) surfaces with smoothness, and the etching surface roughness is completely neglected, these studies are useful for the investigation of the silicon anisotropic etching mechanisms and computer simulations of silicon anisotropic etching processes.

For the simulation of silicon anisotropic etching, the surface atoms can be divided into different categories according to the numbers of the first nearest atoms (FNAs),

the second nearest atoms (SNAs), and their locations with respect to the etching surface to the surface atom etching processes (Zhou et al. 2007). A series of equations are then derived to relate the macroscopic etch rates for different silicon crystal planes and the microscopic etch rates for surface atoms with different atom configurations (Zhou et al. 2009). Based on these equations, the microscopic etch rates for different surface atoms are fitted out for different etching conditions. After that, the microscopic activation energies for some typical surface atoms are further calculated based on these microscopic etch rates, to avoid too much laboring fitting procedures for future simulations. Although this study is only an initial attempt to deduce the microscopic activation energies for surface atoms from macroscopic etch rates of silicon crystal planes, it is useful to reduce the efforts to fulfill the etch rate database for silicon anisotropic etching simulation requirements and to stimulate the systematical measurements of silicon anisotropic macroscopic etch rates for simulation applications. The presented model has been extended to a simulation system based on a dynamic CA method (Zhou et al. 2009), and a series of simulations have been performed using the simulation system. Some simulation results are presented and compared with the experimental results.

---

## Microscopic Etch Rates of Surface Atoms

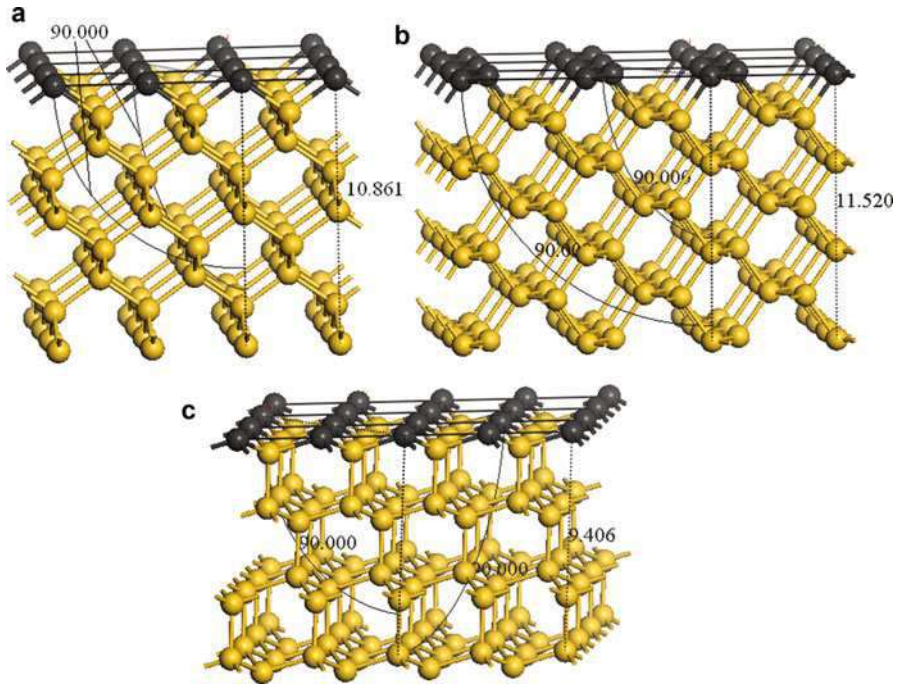
The correspondence between the macroscopic etch rates of silicon crystal planes and the microscopic etch rates of classified atoms with different configurations are determined by individually etching different orientation silicon wafers. The computer software Materials Studio 4.0 (Accelrys Inc.) is used for the analysis of atom configurations. The atom configurations for different crystal planes can be simply notated as (*as*, *bi*, *cs*, *di*), where “s” stands for surface, “i” stands for interior, and “*as*” and “*bi*” indicate that the numbers of the first nearest surface atoms is *a* and the numbers of the first nearest interior atoms is *b*, respectively. Similarly, “*cs*” and “*di*” indicate that the numbers of the second nearest surface atoms is *c* and the numbers of the second nearest interior atoms is *d*, respectively. The second nearest surface atoms include indirect and direct second nearest surface atoms. The definitions of indirect and direct second nearest surface atoms (neighbors) are originally introduced in Zubel (2003). The direct second nearest surface atoms are directly linked to a target atom through first nearest atom(s). While the indirect nearest neighbor atoms are not directly linked to the target atom through first nearest atom(s), in other words, the first nearest atom(s), linking the indirect nearest neighbor atoms and the target atom, has (have) been removed away during the etching process. Figure 1 shows the schematic diagram of FNAs and SNAs for target atoms located on (110) crystal planes. The microscopic etch rates for surface atoms with (*as*, *bi*, *cs*, *di*) configuration are noted as  $R_{(a, b, c, d)}$ , and the macroscopic etch rates for (*hkl*) crystal planes are noted as  $R_{(hkl)}$ . The continuous state variable (Zhu and Liu 2000) is adopted in this chapter; thus a surface atom has a state value ranging continuously from 0 (“removed”) to 1 (“unetched”).





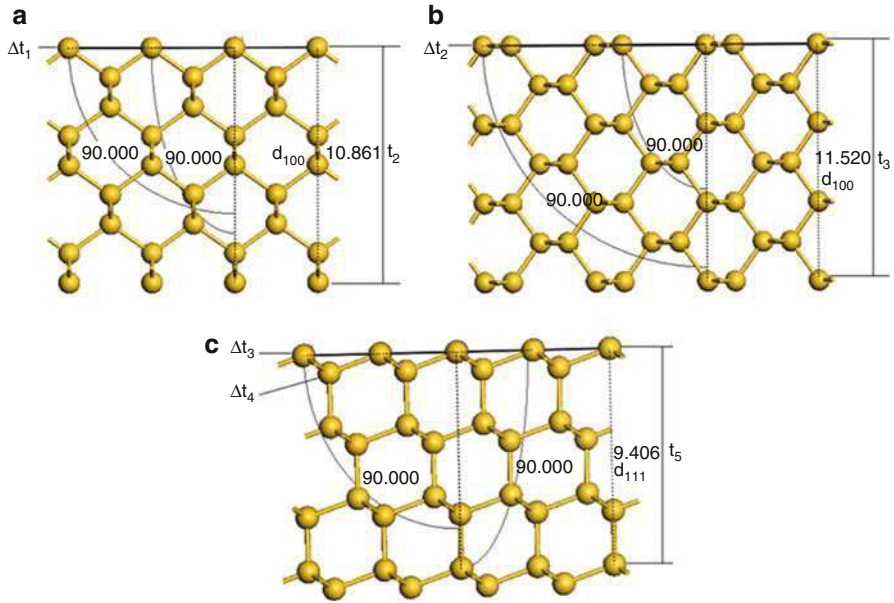
**Fig. 1** Surface atom configurations for (110) and (311) crystal planes

During the etching process of various silicon wafers without mask, only several types of surface atoms, even only one or two types of surface atoms, will be involved in etch rate determination, as shown in Figs. 2, 3, and 4. Thus the procedures to deduce the correspondence equations between the macroscopic etch rates of silicon crystal planes and the microscopic etch rates of classified atoms can be significantly simplified. Figures 2 and 3 show the three-dimensional views and the respective side views of atom configurations for (100), (110), and (111) wafers. Figure 4 shows the respective side views of atom configurations for (120) and (130) wafers. Considering that the crystal structures are dynamically repeated during the etching processes if the morphologies of the etching surfaces are neglected, namely, the surface atom configurations are reproduced during the etching processes, a series of equations can be deduced for the etching processes of some typical wafers, based on the bonding (atom linking) situations of these wafers. The basic principle to deduce these equations is that the necessary time to remove silicon atoms for several layers equals to the time to advance the crystal surface (or the time for certain etching depth), which can be directly determined by the macroscopic etch rates for different silicon. For example, since the state variable of the silicon atoms at the beginning of etching is 1 and the time to remove one atom located on (100) plane is  $\Delta t_1 = 1/R_{(0s, 2i, 4s, 4i)}$ , the necessary time to remove the top 8 layer atoms for (100) planes is  $t_1 = 8 \cdot (1/R_{(0s, 2i, 4s, 4i)})$ , as shown in Fig. 3a, while the time to advance the crystal surface is directly determined by the macroscopic etch rate for (100) crystal plane as  $t_2 = d_{100} \cdot \text{Re} / R_{100}$ , where  $\text{Re}$  is the simulation scale factor and  $d_{100} \cdot \text{Re}$  indicates the etching distance (depth) from the top of the (100) crystal wafer. So in other words,  $t_2$  is the time for  $d_{100} \cdot \text{Re}$  etching

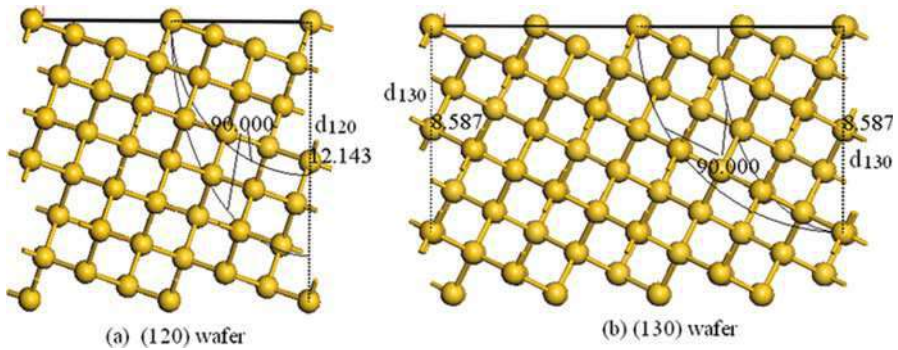


**Fig. 2** (a), (b), and (c) are the three-dimensional views of atom configurations for (100), (110), and (111) wafers

depth from the top of the (100) wafer. In the following part,  $Re \cdot d_{hkl}$  indicates the etching distance (depth) from the top of the  $(hkl)$  crystal wafer. Several repeated crystal structures, or in other words several layer atoms, are counted in this study for two reasons. One reason is to reduce the distance measurement error, and the other is to determine the etching depth directly by using Materials Studio 4.0 (Accelrys Inc.). The simulation scale factor means that each unit cell in the simulation is  $Re$  times bigger than that of a real unit cell in silicon crystal. The reason to incorporate the simulation scale factor into the model lies in that the exact simulation size is always more than several microns for engineering applications. The typical simulation array will be larger than  $10,000 \times 10,000 \times 10,000$  for general cases, if the real size of the silicon unit cell is adopted for simulations. Thus the silicon anisotropic etching simulations become impossible to be implemented by common personal computers. To solve this problem, the simulation scale factor is incorporated into the model; thus the simulation array can be readjusted according to the etching size. Note that, when parameter  $Re$  equals to 1, the etching processes is the exact etching processes of silicon crystal. So to some degree, we can say that the parameter  $Re$  is for efficient simulations of different etching sizes, especially for practical engineering applications. Parameter  $Re$  will not prevent us reveal the exact etching processes of silicon anisotropic etching techniques. Obviously, we have  $t_1 = t_2$  according to Fig. 3a, so the equation to relate



**Fig. 3** The respective side views of atom configurations for (100), (110), and (111) wafers



**Fig. 4** The respective side views of atom configurations for (120) and (130) wafers

the microscopic etch rates and the macroscopic etch rates for (100) wafer can be deduced as given by

$$\frac{1}{R_{(0s, 2i, 4s, 4i)}} \cdot 8 = \frac{d_{100} \cdot Re}{R_{100}} = \frac{10.861 \cdot Re}{R_{100}} \quad (1)$$

Similarly, according to Fig. 3b, the equation to relate the microscopic etch rates and the macroscopic etch rates for (110) wafer can be deduced as given by

$$\frac{1}{R_{(2s, 1i, 2s, 5i)}} \cdot 6 = \frac{d_{110} \cdot \text{Re}}{R_{110}} = \frac{11.520 \cdot \text{Re}}{R_{110}} \quad (2)$$

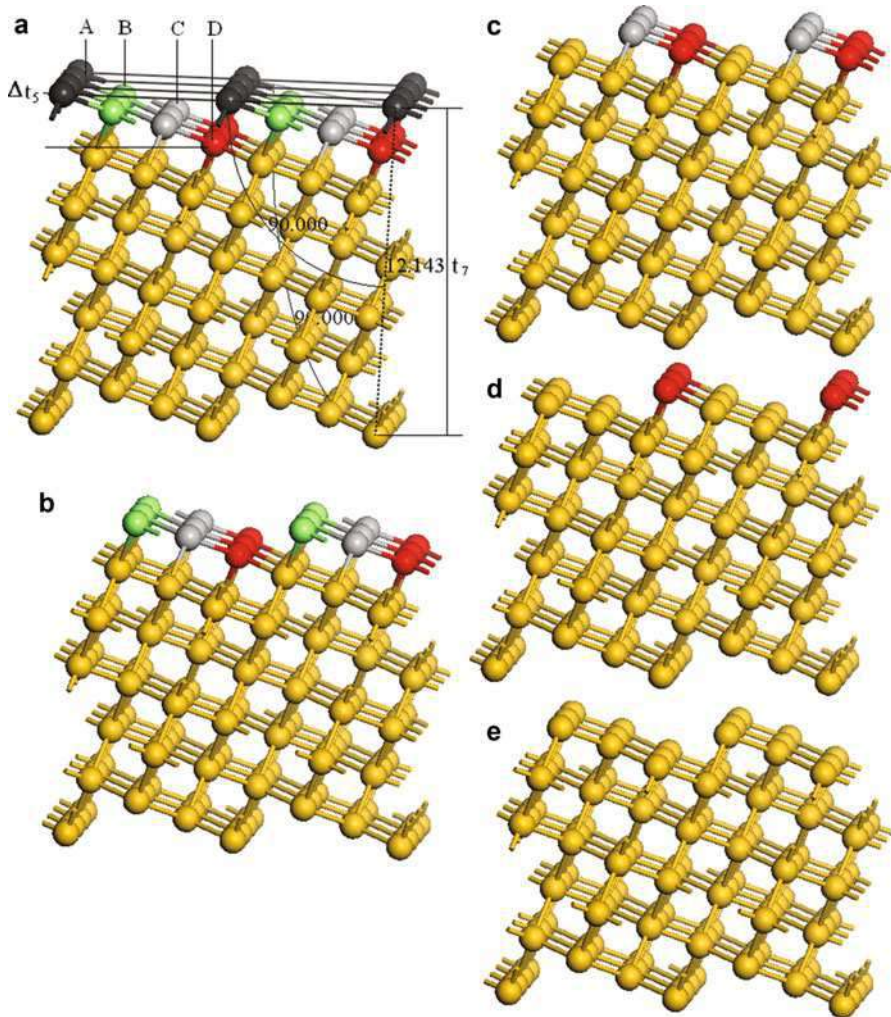
where  $\text{Re} \cdot d_{110}$  indicates the etching distance (depth) from the top of the (110) wafer.

The respective side view crystal structures for (100) and (110) planes are simply repeated, namely, only one crystal structure (atom configuration) is reproduced during the etching processes. The microscopic etch rates for surface cells on (100) and (110) planes can be directly related to the corresponding macroscopic etch rates. However, for other crystal planes such as (111) plane, more than one crystal structure will be reproduced during the etching processes, as shown in Fig. 3c. Similarly, the time to remove one atom with  $(0s, 3i, 6s, 3i)$  configuration is  $\Delta t_3 = 1/R_{(0s, 3i, 6s, 3i)}$ , and the time to remove one atom with  $(0s, 1i, 6s, 3i)$  configuration after the top layer atoms are removed is  $\Delta t_4 = 1/R_{(0s, 1i, 6s, 3i)}$ . Thus the total time to remove the top 6 layer atoms for (111) planes is  $t_4 = 3 \cdot (\Delta t_2 + \Delta t_3)$ , as shown in Fig. 3c, while the time to advance the crystal surface is directly determined by the macroscopic etch rate for (111) crystal plane as  $t_5 = d_{111} \cdot \text{Re} / R_{111}$ . According to  $t_4 = t_5$ , we get the equation to relate the microscopic etch rates and the macroscopic etch rates for (111) wafer

$$\left( \frac{1}{R_{(0s, 3i, 6s, 3i)}} + \frac{1}{R_{(0s, 1i, 6s, 3i)}} \right) \cdot 3 = \frac{R_{(0s, 1i, 6s, 3i)} + R_{(0s, 3i, 6s, 3i)}}{R_{(0s, 3i, 6s, 3i)} \cdot R_{(0s, 1i, 6s, 3i)}} \cdot 3 = \frac{d_{111} \cdot \text{Re}}{R_{111}} = \frac{9.406 \cdot \text{Re}}{R_{111}} \quad (3)$$

For the etching processes of other crystal planes such as (120), (130), (211), (411), (311), (331), and (221) wafers (planes), several atom configurations will be reproduced during the etching processes, as shown in Fig. 4. Namely, during the etching processes of a surface atom located the three low index planes, the atom configuration is kept unchanged, as shown in Figs. 2 and 3. However, during the etching processes of some other surface atoms, the atom configurations may be changed, and this means the microscopic etch rates of the surface atoms will change during their etching processes. Even so, the same principle to deduce Eqs. 1, 2, and 3 is adopted for the equations to relate the microscopic etching rates and the macroscopic etching rates for (120), (130), (211), (411), (311), (331), and (221) wafers. As an example, the equation to relate the microscopic etching rates and the macroscopic etching rates for (120) wafers is introduced in the following. During the etching processes of (120) crystal plane, three types of atoms with different atom configurations appear at the same time, and the atom configuration is changed with the etching processes of a surface atom, as shown in Fig. 5. The crystal structures are dynamically repeated during the etching processes, assuming that the necessary time to remove A-type atoms (the top layer atoms) is  $\Delta t_4$  for their current status, as shown in Fig. 5a, b, and then during the following  $\Delta t_4$  time, B-type atoms (the second top layer atoms) will be removed (Fig. 5c). In fact, the B-type atoms will hold the same atom configurations as A-type atoms after the first  $\Delta t_4$ , namely, after A-type atoms are removed away.





**Fig. 5** Schematic diagram for the etching processes of (120) wafer. (a) is the original structure and (b), (c), (d), and (e) are the etching results of for the etching time of  $\Delta t_5$ ,  $2\Delta t_5$ ,  $3\Delta t_5$ , and  $4\Delta t_5$ , respectively

Similarly, C-type atoms will be removed after B-type atoms are removed away, and D-type atoms will be removed after C-type atoms are removed away, as shown in Fig. 5d, e. We choose this as an etching cycle; thus the total etching cycle time is  $4\Delta t_4$ . Then the problem is to determine the exact value of  $\Delta t_4$ . From the whole etching processes of D-type atoms, we can see that the microscopic etch rate for D-type atoms is “0” during the first  $\Delta t_4$  time, since D-type atoms are interior atoms. During the second  $\Delta t_4$  time, D-type atoms become surface atoms and hold microscopic

etch rate,  $R_{(1s, 2i, 3s, 6i)}$ . During the following  $\Delta t_4$  time, D-type atoms are surface atoms and hold microscopic etch rate,  $R_{(2s, 1i, 0s, 6i)}$ . During the final  $\Delta t_4$  time, D-type atoms are still surface atoms and hold microscopic etch rate,  $R_{(1s, 1i, 3s, 3i)}$ , and D-type atoms are totally removed away after the final  $\Delta t_4$  time. Considering that the state of a surface atom is “1” at the beginning of the etching processes and it is reduced to “0” when it is totally removed away, we can have  $R_{(1s, 2i, 3s, 6i)}\Delta t_4 + R_{(2s, 1i, 0s, 6i)}\Delta t_4 + R_{(1s, 1i, 3s, 3i)}\Delta t_4 = 1$ . This method has also been presented by Gosalvez and his colleagues (Xing et al. 2007; Hubbard and Antonsson 1994). Thus we get  $4\Delta t_4 = 4/(R_{(1s, 2i, 3s, 6i)} + R_{(2s, 1i, 0s, 6i)} + R_{(1s, 1i, 3s, 3i)})$ . Five etching cycles are considered, as shown in Fig. 5, and then the time to advance the crystal surface is directly determined by the macroscopic etching rate for (120) crystal plane as  $t_7 = d_{120} \cdot \text{Re} / R_{120}$ . So we get the equation to relate the microscopic etching rates and the macroscopic etching rates for (120) wafer

$$\begin{aligned} t_6 &= 4\Delta t_4 \cdot 5 = \frac{20}{R_{(1s, 2i, 3s, 6i)} + R_{(2s, 1i, 0s, 6i)} + R_{(1s, 1i, 3s, 3i)}} \\ &= \frac{d_{120} \cdot \text{Re}}{R_{120}} = \frac{12.143 \cdot \text{Re}}{R_{120}} \end{aligned} \quad (4)$$

Similar to (120) planes, the equations to relate the microscopic etch rates and the macroscopic etch rates for (130), (211), (411), (311), (331), and (221) wafers can be deduced as given by

$$\frac{10}{R_{(1s, 2i, 2s, 6i)} + R_{(1s, 1i, 2s, 4i)}} = \frac{d_{130} \cdot \text{Re}}{R_{130}} = \frac{8.587 \cdot \text{Re}}{R_{130}} \quad (5)$$

$$\begin{aligned} 6. \quad &\frac{R_{(0s, 3i, 5s, 5i)} + R_{(0s, 2i, 5s, 2i)} + R_{(1s, 1i, 2s, 5i)} - R_{(1s, 2i, 4s, 3i)}}{(R_{(0s, 3i, 5s, 5i)} + R_{(0s, 3i, 4s, 5i)} + R_{(0s, 2i, 5s, 2i)}) \cdot R_{(1s, 1i, 2s, 5i)}} = \frac{d_{211} \cdot \text{Re}}{R_{211}} \\ &= \frac{6.651 \cdot \text{Re}}{R_{211}} \end{aligned} \quad (6)$$

$$\begin{aligned} &\frac{18(R_{(1s, 1i, 2s, 5i)} + R_{(0s, 3i, 5s, 5i)} + R_{(0s, 2i, 5s, 2i)} - R_{(1s, 2i, 2s, 5i)})}{2R_{(1s, 1i, 2s, 5i)}(R_{(0s, 3i, 5s, 5i)} + R_{(1s, 2i, 3s, 5i)} + R_{(0s, 2i, 3s, 4i)} + R_{(0s, 2i, 5s, 2i)})} \\ &\quad - R_{(1s, 1i, 2s, 5i)}(R_{(1s, 2i, 3s, 5i)} + R_{(1s, 2i, 2s, 5i)} + R_{(0s, 2i, 3s, 4i)}) \\ &= \frac{d_{411} \cdot \text{Re}}{R_{411}} = \frac{11.520 \cdot \text{Re}}{R_{411}} \end{aligned} \quad (7)$$

$$\begin{aligned} 11. \quad &\frac{R_{(0s, 3i, 5s, 5i)} + R_{(0s, 2i, 5s, 2i)} + R_{(1s, 1i, 2s, 5i)} - R_{(1s, 2i, 2s, 5i)}}{(R_{(0s, 3i, 5s, 5i)} + R_{(0s, 2i, 5s, 2i)}) \cdot R_{(1s, 1i, 2s, 5i)}} = \frac{d_{311} \cdot \text{Re}}{R_{311}} \\ &= \frac{18.012 \cdot \text{Re}}{R_{311}} \end{aligned} \quad (8)$$

$$\begin{aligned}
 19. & \frac{R_{(0s, 3i, 4s, 5i)} + R_{(2s, 1i, 4s, 3i)} + R_{(2s, 1i, 4s, 5i)} + R_{(0s, 1i, 4s, 3i)} - 2R_{(2s, 1i, 2s, 5i)}}{(R_{(0s, 3i, 4s, 5i)} + R_{(2s, 1i, 4s, 3i)}) \cdot (R_{(2s, 1i, 4s, 5i)} + R_{(0s, 1i, 4s, 3i)}) - R_{(2s, 1i, 2s, 5i)}^2} \\
 & = \frac{d_{331} \cdot \text{Re}}{R_{331}} = \frac{23.672 \cdot \text{Re}}{R_{331}}
 \end{aligned} \tag{9}$$

$$\begin{aligned}
 18. & \frac{R_{(0s, 3i, 6s, 3i)} + R_{(2s, 1i, 4s, 5i)} + R_{(0s, 1i, 4s, 3i)} - R_{(2s, 1i, 2s, 5i)}}{(R_{(0s, 3i, 4s, 5i)} + R_{(0s, 3i, 6s, 3i)} + R_{(2s, 1i, 4s, 3i)}) \cdot (R_{(2s, 1i, 4s, 5i)} + R_{(0s, 1i, 4s, 3i)})} \\
 & \quad - R_{(2s, 1i, 2s, 5i)}(R_{(0s, 3i, 4s, 5i)} + R_{(2s, 1i, 4s, 3i)}) \\
 & = \frac{d_{221} \cdot \text{Re}}{R_{221}} = \frac{16.292 \cdot \text{Re}}{R_{221}}
 \end{aligned} \tag{10}$$

Equations 5, 6, 7, 8, 9, and 10 above are for the etching processes of (130), (211), (411), (311), (331), and (221) wafers, respectively. Since the respective side view crystal structures for (100) and (110) planes are simply repeated, the microscopic etch rates for surface atoms on (100) and (110) planes can be directly calculated using the corresponding macroscopic etch rates. However, determinations of the microscopic etch rates for other surface atoms are a little complex, since there are too many variables (microscopic etch rates for many surface atoms with different atom configurations) in the corresponding equations. Thus, laboring fitting procedures are necessary to determine these microscopic etch rates. Similar to the approaches in Xing et al. (2007), Eqs. 3, 4, 5, 6, 7, 8, 9, and 10 are simplified by equaling microscopic etch rates of some atoms with similar atom configurations, for example,  $R_{(0s, 3i, 6s, 3i)} = R_{(0s, 3i, 4s, 5i)} = R_{(0s, 3i, 5s, 5i)}$ , etc. Based on the measured macroscopic etch rates for (100), (110), (111), (120), (130), (211), (411), (311), (331), and (221) wafers, the microscopic etch rates for different surface atoms can then be obtained by fitting procedures (including some trials). The fitting results for surface atoms located on some selected typical crystal planes are summarized in Table 1. Note that, the factor of  $Sc = \text{Re} / 10,000$  is used for scaling of [ $\text{\AA}$ ] to [ $\mu\text{m}$ ] in this chapter, since most researchers in MEMS and silicon anisotropic etching area are more familiar with the unit ( $\mu\text{m}/\text{min}$ ) for silicon etching rates. Here the etch rates in 30 wt% and 40 wt% KOH solutions are taken from experiments by Sato et al. (1998). We should point out that more equations like Eqs. 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 can also be deduced using the same method for other orientation wafers, such as (320), (530), etc. This is helpful to increase the accuracy of the fitting procedures, but this will make it more difficult to fulfill the macroscopic etch rate database for different orientation wafers (planes). In our previous study (Zhou et al. 2009), macroscopic etch rates for nine crystal planes were adopted to fit out the microscopic etch rates for some typical surface atoms. However, we choose macroscopic etch rates for ten crystal planes in this study as introduced above. We should point out that this is not related to technical method innovation. The reason lies in that one or two more macroscopic

**Table 1** Microscopic etch rates ( $R_x = R_x' \cdot Re / 10,000$ ) for some typical surface atoms for the silicon anisotropic etching simulations in 30 wt% and 40 wt% KOH solution

Atom type ( $\alpha$ )	(0 s, 2i, 4 s, 4i)	(2 s, 1i, 2 s, 5i)	(0 s, 3i, 6 s, 3i)	(1 s, 1i, 3 s, 3i)	(1 s, 1i, 2 s, 4i)	(1 s, 2i, 4 s, 3i)	(0 s, 2i, 5 s, 2i)	(2 s, 1i, 4 s, 3i)	(1 s, 1i, 2 s, 5i)	(1 s, 2i, 2 s, 6i)	(2 s, 1i, 0 s, 6i)
$R_\alpha'$ ( $\mu\text{m}/\text{min}$ ) (30 wt%, 70 °C)	0.5871	0.7578	0.00159	1.2014	1.1742	1.0542	1.6412	0.8847	1.3210	0.5323	0.8324
$R_\alpha'$ ( $\mu\text{m}/\text{min}$ ) (30 wt%, 80 °C)	1.0373	1.3592	0.0031	2.1663	2.1220	1.9118	2.8930	1.5729	2.4004	0.9528	1.4928
$R_\alpha'$ ( $\mu\text{m}/\text{min}$ ) (40 wt%, 70 °C)	0.4413	0.6740	0.00287	0.9842	0.9045	0.7456	1.2014	0.6534	0.9823	0.4041	0.7022
$R_\alpha'$ ( $\mu\text{m}/\text{min}$ ) (40 wt%, 80 °C)	0.7632	1.1249	0.00561	1.7514	1.6158	1.3370	2.0937	1.1410	1.7649	0.7130	1.2472

Note that the factor of  $\mathcal{S}_C = Re / 10,000$  is used for scaling of [ $\text{\AA}$ ] to [ $\mu\text{m}$ ], since most researchers in MEMS and silicon anisotropic etching area are more familiar with the unit ( $\mu\text{m}/\text{min}$ ) for silicon etching rates



etch rates for some other crystal planes, besides the nine basic macroscopic etch rates, will be useful to readjust some fitted values to improve the accuracy of fitting procedures, according to our experience. So the macroscopic etch rate for (221) plane is further included in this study, compared with our previous research (Zhou et al. 2009).

We should point out that Gosálvez and his colleagues have presented a more complete study of the equations relating the macroscopic etch rates of different silicon crystal planes and the microscopic etch rates of surface silicon by incorporating step flow aspects in their recent papers. Based on their approaches, different corner undercutting phenomena and etching processes have been analyzed and predicted (Gosálvez et al. 2011, 2016; Pal et al. 2014). The presented model has been extended to a simulation system based on a CCA method. Compared with Eqs. 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10, similar equations have been presented except the scaling factor and some atom classifications finally adopted. Unlike the presented method, Gosálvez's study is more systematical and more reasonable, by deriving their equations in a more general framework based on a classification of the different surface orientations into surface families. By considering the silicon anisotropic etching process as a dynamic repeated process, our method is to deduce the equations based on the geometrical structures of crystal silicon, and this method can be replaced by Gosálvez's method. However, in order to keep the integrity of this chapter and our study and to help the readers to efficiently appreciate the approaches to obtain these equations, we still describe the derivation of Eqs. 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 in this chapter.

With the fitted microscopic etch rates for different surface atoms using Eqs. 1–10, the atomic-level silicon anisotropic etching simulations can be implemented for the corresponding etching conditions. However, some problems deserve discussions. From the abovementioned fitting procedures and our previous study (Zhou et al. 2009), we notice that the macroscopic etch rates for different etching conditions should be measured for the corresponding fitting procedures. Thus too many efforts are expected to fulfill the etch rate database for silicon anisotropic etching simulations under different etching conditions. Furthermore, the fitting procedures should be carried out for every new etching condition. Theoretically speaking, these fitting procedures can be implemented by the computer program for silicon anisotropic etching simulation, with the input macroscopic etch rates of the typical crystal planes. However, since the fitting procedures include complex trials, it is difficult to accurately and efficiently fit out all microscopic etch rates only by computer program. In other words, it is not practical to accurately solve Eqs. 1–10 only by computer program. There are many variables in Eqs. 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 even after the simplification, so mostly it is not easy to obtain the relative proper fitted values for these variables. This is the main reason why our previous simulation system SEAES (Zhou et al. 2007) is not as practical as we expected before. To solve these problems and extend the model to a simulation system more accurately and efficiently, some useful attempts will be presented in the following part.

## Microscopic Activation Energies for Surface Atoms

The macroscopic etch rates of silicon crystal planes in alkaline solutions are acknowledged to obey the Arrhenius law. According to Eqs. 1 and 2, the microscopic etch rates for surface atoms located on (100) and (110) planes also obey the Arrhenius law. Following this, we have reasons to assume that the microscopic etch rates ( $R_\alpha$ ) for other surface atoms located on the selected crystal planes approximately obey the Arrhenius law. Thus we have the microscopic etch rates for different atom types ( $\alpha$ ) as given by

$$R_\alpha = R_{\alpha 0} e^{-\Delta E_\alpha / k_B T} \quad (11)$$

where  $R_{\alpha 0}$  is the pre-exponential factor,  $\Delta E_\alpha$  is the microscopic activation energy for  $\alpha$ -type atom,  $k_B$  is Boltzmann constant, and  $T$  is the etching temperature.

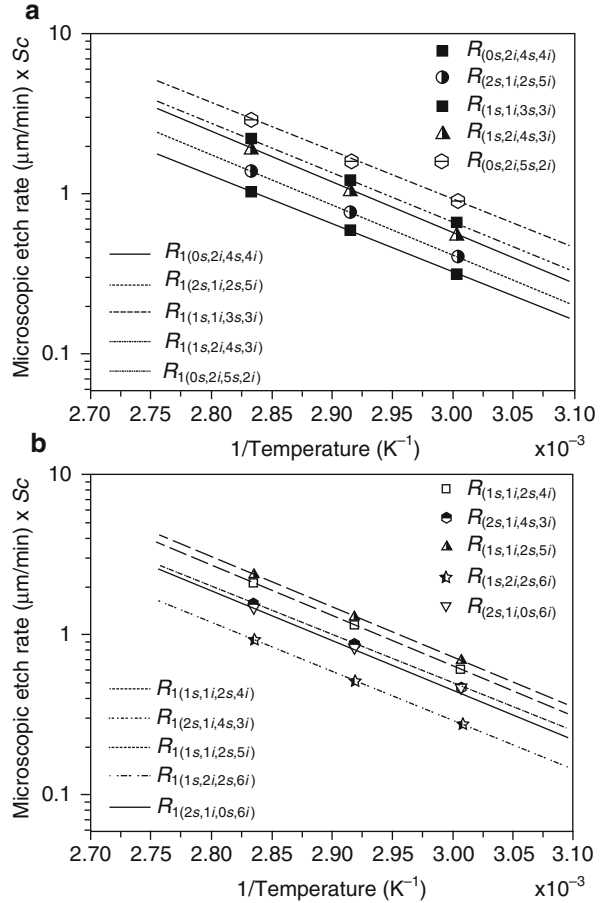
The microscopic etch rates of some typical surface atoms located on (100), (110), (111), (120), (130), (211), (411), (311), (331) and (221) planes for different etching conditions (e.g., 60 °C, 70 °C, and 80 °C 30 wt% KOH solutions; 60 °C, 70 °C, and 80 °C 40 wt% KOH solutions; and 60 °C, 70 °C, and 80 °C 50 wt% KOH solutions) are fitted out using the methods introduced in the previous section. Then the corresponding microscopic activation energies ( $\Delta E_\alpha$ ) and the corresponding pre-exponential factors for these surface atoms can be determined using Eq. 11, according to the fitted microscopic etch rates for these surface atoms under different etching temperatures. In this chapter, the microscopic activation energies for 11 types surface atoms are calculated out using the Arrhenius law, with the obtained microscopic etch rates for these surface atoms. With these microscopic activation energies for these typical surface atoms, some microscopic etch rates can be directly calculated for different etching temperatures. Thus variables (microscopic etch rates for many surface atoms with different atom configurations) in Eqs. 3–10 are reduced, and they can directly calculated out using Eqs. 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. So the macroscopic etch rates for different crystal planes under various etchant concentrations with two or three different temperatures are enough for the implementation of the fitting procedures. Furthermore, considering that the macroscopic etch rates of different crystal planes can be obtained using interpolation method (Wind et al. 2002), for similar etchant concentrations, the fitting procedures are significantly simplified, compared with the previous model (Zhou et al. 2007). Thus the experimental efforts in fulfillment of the macroscopic etch rate database for silicon anisotropic etching simulations are reduced, and this also makes it practical to accurately and efficiently obtain all microscopic etch rates by fitting procedures implemented by computer programs in the future.

Table 2 shows the fitted microscopic activation energies for some typical surface atoms appearing in the etching processes of (100), (110), (111), (120), (130), (211), (411), (311), (331), and (221) planes, by using the fitted microscopic etch rates for 60 °C, 70 °C, and 80 °C 30 wt% KOH solutions and 60 °C, 70 °C, and 80 °C 40 wt% KOH solutions. As an example, Fig. 6 shows the Arrhenius dependence of the microscopic etch rates for ten types of atoms after fitting the pre-exponential factors

**Table 2** Activation energies ( $\Delta E_a$ ) and pre-exponential factors ( $R_{\alpha 0} = R_{\alpha 0}' \cdot \text{Re}^{-1/10}$ , 000) for some typical surface atoms for the silicon anisotropic etching simulations in 30 wt% and 40 wt% KOH solution

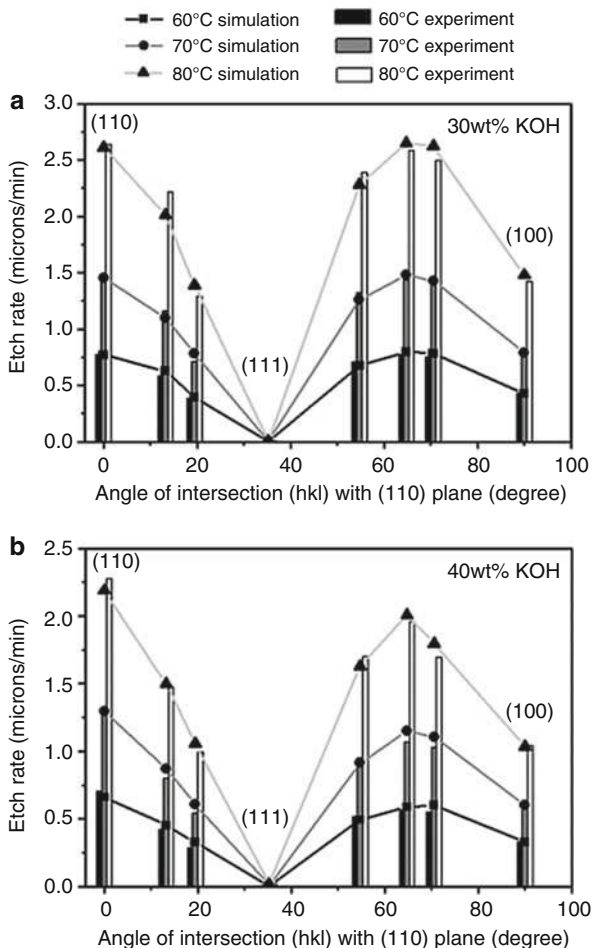
Atom type ( $\alpha$ )	(0 s, 2i, 4 s, 4i)	(2 s, 1i, 2 s, 5i)	(0 s, 3i, 6 s, 3i)	(1 s, 1i, 3 s, 3i)	(1 s, 1i, 2 s, 4i)	(1 s, 2i, 4 s, 3i)	(0 s, 2i, 5 s, 2i)	(2 s, 1i, 4 s, 3i)	(1 s, 1i, 2 s, 5i)	(1 s, 2i, 2 s, 6i)	(2 s, 1i, 4 s, 3i)
$R_{\alpha 0}$ ( $\mu\text{m}/\text{min}$ ) (30 wt%)	4.003 $\times 10^8$	8.871 $\times 10^8$	4.203 $\times 10^7$	1.324 $\times 10^9$	1.385 $\times 10^9$	1.423 $\times 10^9$	8.039 $\times 10^8$	5.874 $\times 10^8$	1.908 $\times 10^9$	4.478 $\times 10^8$	7.492 $\times 10^9$
$\Delta E_a$ (eV) (30 wt%)	0.602	0.618	0.710	0.616	0.618	0.622	0.592	0.601	0.624	0.608	0.610
$R_{\alpha 0}$ ( $\mu\text{m}/\text{min}$ ) (40 wt%)	1.430 $\times 10^8$	4.065 $\times 10^8$	5.411 $\times 10^7$	6.759 $\times 10^8$	7.111 $\times 10^8$	6.711 $\times 10^8$	3.923 $\times 10^8$	2.283 $\times 10^8$	9.459 $\times 10^8$	2.047 $\times 10^8$	4.507 $\times 10^8$
$\Delta E_a$ (eV) (40 wt%)	0.580	0.600	0.700	0.602	0.606	0.610	0.580	0.582	0.612	0.593	0.600

**Fig. 6** The Arrhenius dependence of the microscopic etch rates for ten types of atoms after fitting the pre-exponential factors and microscopic activation energies in 30 wt% solutions.  $R_{(as, bi, cs, di)}$  indicates the atomistic rate values of (as, bi, cs, di) atoms that were calibrated at two or three temperatures using Eqs. 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10, and  $R_{1(as, bi, cs, di)}$  indicates the corresponding atomistic etch rates after fitting the pre-exponential factors and microscopic activation energies



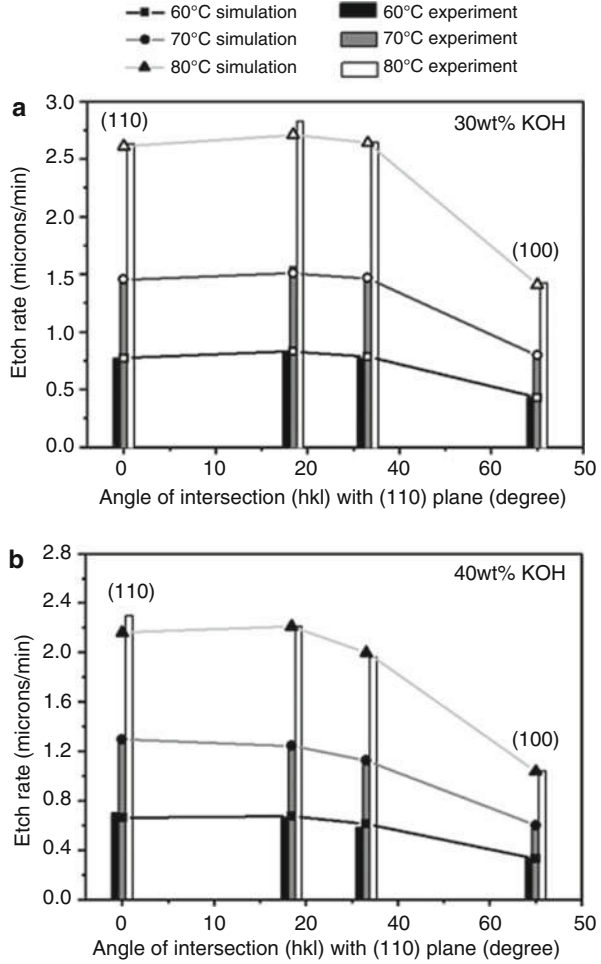
and microscopic activation energies in 30 wt% KOH solutions. These fitted microscopic etch rates are in good agreement with the corresponding microscopic etch rates that were calibrated at two or three temperatures using Eqs. 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. Note that the macroscopic etch rates for the implementations of the fitting procedures are taken from several published materials by different authors (Seidel 1990; Seidel et al. 1990; Herr and Baltes 1991; Sato et al. 1998; Wind and Hines 2000; Zubel and Kramkowska 2004), in terms of macroscopic activation energies and macroscopic etch rates for (100), (110), (111), (120), (130), (211), (411), (311), (331), and (221) planes. In fact, these macroscopic etch rates are obtained by combining these materials. Following this, some discussions from two aspects are valuable. On the one hand, since different authors using different measurement methods and etching equipment, their measured macroscopic etch rates and macroscopic activation energies for different crystal planes inevitably have some errors, leading to some inaccuracy of the fitting results of the microscopic activation energies and also the anisotropic etching simulations. However, the anisotropy of

**Fig. 7** Etch rates from simulation and experiments for some typical planes from the [110] crystal zone in KOH solution. The angle intersections ( $hkl$ ) with (110) planes for (110), (331), (221), (111), (211), (311), (411), and (100) are  $0^\circ$ ,  $13.26^\circ$ ,  $19.47^\circ$ ,  $35.26^\circ$ ,  $54.74^\circ$ ,  $64.76^\circ$ ,  $70.53^\circ$ , and  $90^\circ$ , respectively



macroscopic etch rate for different crystal planes from our simulations using the fitted microscopic activation energies is good, and the simulation results are in good agreement with the experimental results, if we acknowledge that some inaccuracy are inevitable and acceptable. Figures 7 and 8 show the comparisons of the macroscopic etch rates from simulations and available experiments for some typical planes from [110] crystal zone and [100] crystal zone, respectively. The angle intersection ( $hkl$ ) with (110) planes for different planes from [110] crystal zone and [100] crystal zone is taken from Zubel's work (Zhou et al. 2007; Zubel and Kramkowska 2004). The well match between simulation results and the experimental results reveal the effectiveness of our model. A series of simulations and experiments have been performed to further investigate the performance of our model, and detailed comparisons will be presented in the following section. On the other hand, this indicates that the previous study of macroscopic etch rates for different crystal planes under

**Fig. 8** Etch rates from simulation and experiments for some typical planes from the [100] crystal zone in KOH solution. The angle intersections (hkl) with (110) planes for (110), (120), (130), and (100) are 0°, 18.43°, 26.57°, and 45°, respectively



various etching conditions seems not to be systematical, at least not systematical enough, for accurate silicon anisotropic etching simulation applications. Even lots of efforts have been poured into the measurements of microscopic etch rates for different crystal planes under various etching conditions during the past 20 years (Seidel 1990; Seidel et al. 1990; Herr and Baltes 1991; Tabata et al. 1992; Allongue et al. 1993; Yun et al. 1998; Sato et al. 1998; Wind and Hines 2000; van Veenendaal et al. 2001; Zubel 2001; Zubel and Kramkowska 2004; Gosalvez et al. 2007), the microscopic etch rates measured (or recorded) by different researchers generally cannot systematically fulfill the macroscopic etch rate database for silicon anisotropic etching process simulations, especially atomic-level simulations. For most of the cases, these measured etch rates cannot be combined with each other for accurate simulation applications, since researchers have no standards to measure (record) macroscopic etch rates for how many silicon crystal planes or which set of silicon

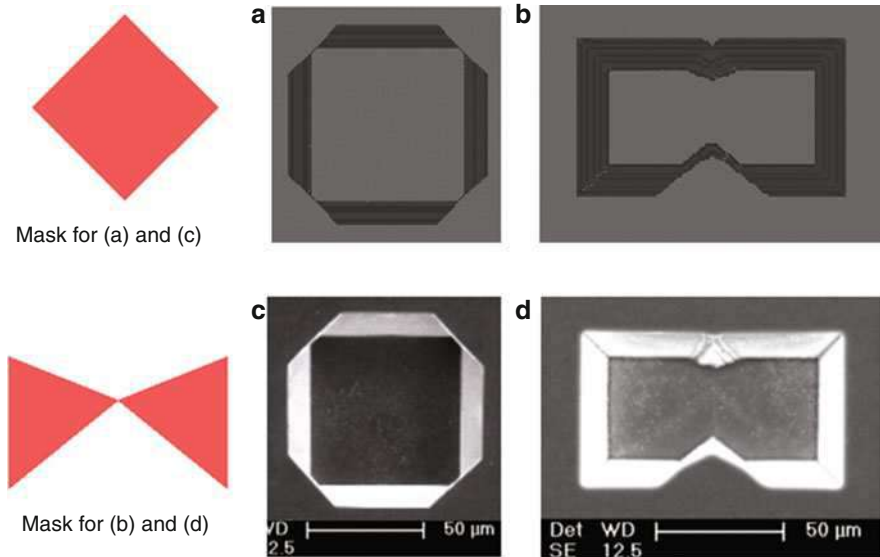
crystal planes. One of the major reasons lies in that some previous silicon anisotropic etching models (simulators), for example, geometric models (Buser and Rooij 1991; Sequin 1992; Danel and Delapierre 1992; Fruhauf et al. 1993; Schroder et al. 2001; Radjenovic et al. 2006), require the etch rates of many crystal planes for relatively accurate simulations and need to measure the corresponding set of etch rates for every etching conditions. Theoretically speaking, the macroscopic etch rate database can be fulfilled for every etching condition, but this is not practical, since too many experimental efforts are required. On the contrary, some other models need much less macroscopic etch rates (Zhu and Liu 2000; Kakinaga et al. 2004; Camon and Moktadir 1995; Gosalvez et al. 2001), since only some etching simulations with limited accuracy can be implemented by these models.

Our method simplifies the fitting procedures by describing the microscopic etch rates of surface atoms located in various crystal planes using the Arrhenius law. Only the macroscopic etch rates for ten crystal planes under various etchant concentrations with two or three different temperatures are necessary for the implementation of macroscopic etch rate database. In this way, the fulfillment of the macroscopic etch rate database for silicon anisotropic etching simulations becomes very practical, and the efforts for the implementation of the macroscopic etch database are significantly reduced. This also makes the measurement of macroscopic etch rates of different crystal planes becomes a systematical task, and this will stimulate researchers to systematically measure the macroscopic etch rates for silicon anisotropic etching simulation applications. Furthermore, our method makes it practical to accomplish the fitting procedures only by computer program. This will be useful to extend the model to an efficient simulator for silicon anisotropic etching processes. At last, our method further proves that the microscopic etch rate for some surface atoms obey the Arrhenius law, which has also been considered and studied by Hines' group and Gosalvez's group (Pal et al. 2014; Gosalvez et al. 2016; Wind et al. 2002).

---

## Simulations and Discussions

The atomic-level simulations of silicon anisotropic etching can be carried out by using cellular automaton (CA) approaches (Radjenovic et al. 2006; Than and Buttgenbach 1994; Hubbard and Antonsson 1997; van Veenendaal et al. 2000; Nishidate et al. 1998; Zhu and Liu 2000; Kakinaga et al. 2004; Zhou et al. 2009) or Monte Carlo (MC) approaches (Camon and Moktadir 1995; Gosalvez et al. 2001, 2002). As for CA approaches, we recently presented a dynamic algorithm matched with our silicon anisotropic etching model to reduce the memory usage and speed up the etching simulation (Zhou et al. 2009). The abovementioned model, combining our efficient dynamic algorithm, has been extended to a silicon anisotropic etching simulation system, based on the system structure of SEAES (Zhou et al. 2009). For the etching of the pattern on different orientation wafers, the data structure should be adjusted properly according to the atom bonding situations of these wafers. In this chapter, a series of simulations have been performed for anisotropic etching processes of different etching conditions and mask shapes on (100) silicon wafers in



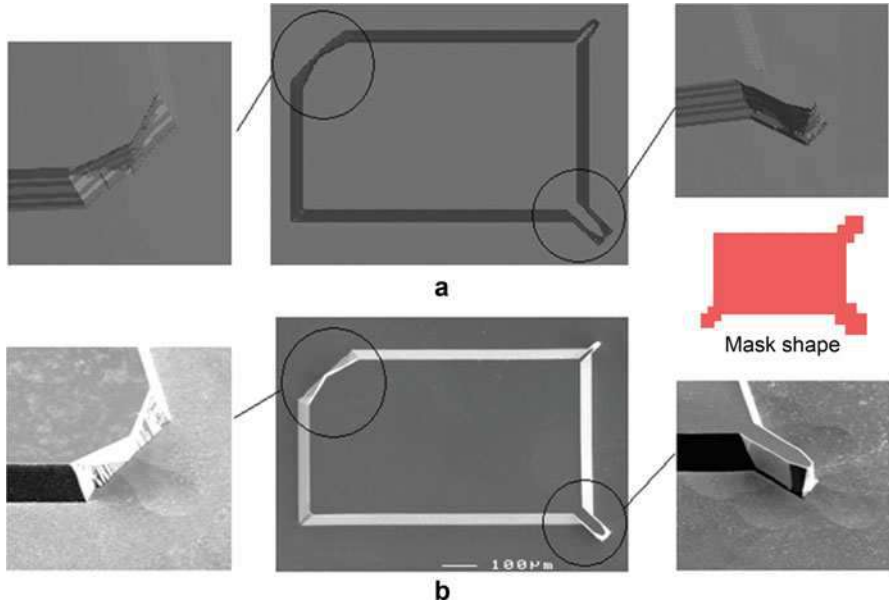
**Fig. 9** (a) and (c) are simulation and experimental results for 24 min etching using a square opening with the edges aligned to  $\langle 100 \rangle$  direction in  $70^\circ\text{C}$  30 wt% KOH solution. (b) and (d) are simulation and experimental results for 24 min etching using a special designed etching opening in  $70^\circ\text{C}$  30 wt% KOH solution

KOH solutions to investigate the performance of the model and the simulation system.

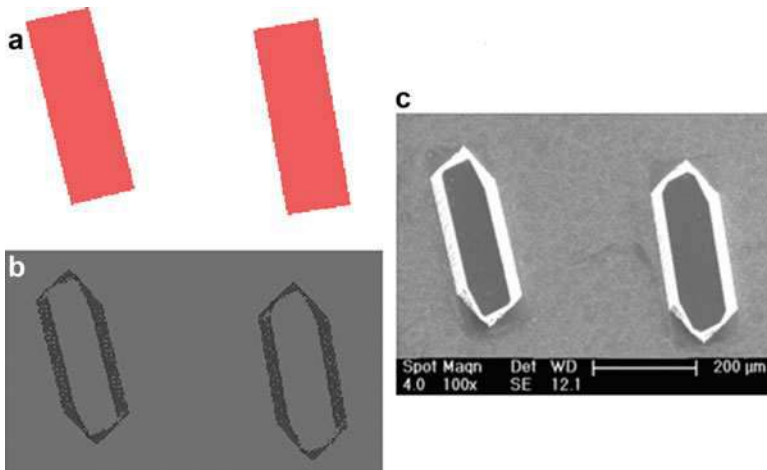
At first, a square etching opening aligned to  $\langle 100 \rangle$  direction and a special designed etching opening are adopted for the etching simulation on (100) wafer, as shown in Fig. 9. The temperature of KOH solution is  $70^\circ\text{C}$ , and the concentration of KOH solution is 30 wt%. The simulation result corresponding to etching time of 24 min shows good agreement with the experimental results, as shown in Fig. 9. Following this, a special mask shape to with three convex corner compensation structures of different scales, as shown in Fig. 10, is adopted, and the etching temperature is changed to  $80^\circ\text{C}$ . The simulation result for 40 min etching is also in good agreement with the corresponding experimental result, as shown in Fig. 10 (a) and (b). After that, the anisotropic etching processes in 40 wt% KOH solutions are also simulated. Figure 11 shows the simulation and experimental results corresponding to 78 min etching, when two rectangular shapes cross shapes with different misalignment with respect to  $\langle 110 \rangle$  direction are adopted. At last, a cross mask is adopted and the etching temperature is changed to  $80^\circ\text{C}$ . The simulation result for 128 min etching is shown in Fig. 12a, and the simulation result matches the experimental result as given by Fig. 12b. The agreements of simulation and experimental profiles indicate the effectiveness of the model and the silicon anisotropic etching simulation system.

Our model is only validated above by performing simulation in limited etching conditions, since only limited macroscopic etch rates are available currently as

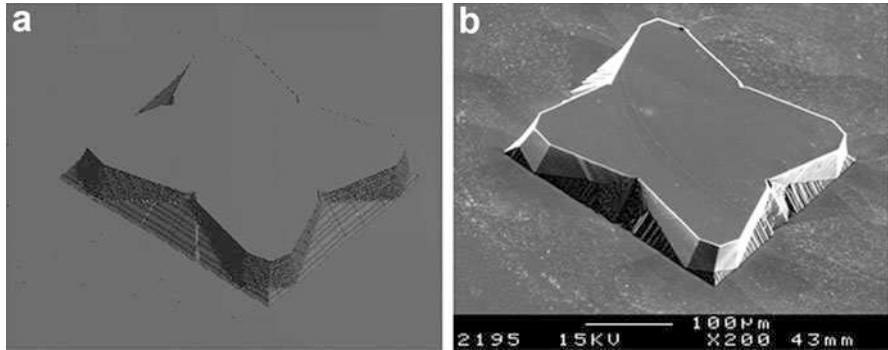




**Fig. 10** Simulation and experimental results using a mask with compensation structures in 80 °C 30 wt% KOH solution (a) simulation results for 40 min etching and (b) experimental results for 40 min etching



**Fig. 11** Simulation and experimental results for two rectangular shapes with different misalignment with respect to  $\langle 110 \rangle$  direction in 70 °C 40 wt% KOH solution (a) mask shape; (b) and (c) are simulation and experimental results for 78 min etching



**Fig. 12** (a) is the simulation experimental results for 128 min etching using a cross mask in 80 °C 40 wt% KOH solution, and (b) is the experimental results corresponding to (a)

introduced in the previous section. However, the obtained results and conclusions are probably general. With more systematical study being carried out on the measurement of silicon anisotropic etch rates in KOH, KOH/IPA, and TMAH etching solutions, the model will become more practical in the future. The model is useful to further extend the applications of silicon wet chemical etching techniques and to improve the design efficiency of some MEMS devices. Furthermore, the procedures in this chapter to simulate the silicon anisotropic etching processes will also benefit other similar researches, such as the researches of the anisotropic etching process simulations of GaAs, InP, etc.

---

## Conclusions

A novel model is developed for silicon anisotropic etching in alkaline solutions, by relating the microscopic etch rates of surface atoms with the macroscopic etch rates of different crystal planes using a series of equations. Based on these equations, the microscopic etch rates for different surface atoms with various atom configurations are fitted out for different etching conditions, and the microscopic activation energies for some typical surface atoms are further calculated. The model has been extended to a simulation systems based on a dynamic CA approach, and the simulation results using the simulation system demonstrate to be in good agreement with the experimental results. This indicates the effectiveness of the model and the silicon anisotropic etching simulation system, and this is useful for the research of anisotropic etching technology and the development of MEMS design.

---

## References

- Allongue P, Costa-Kieling V, Gerischer H (1993) Etching of silicon in NaOH solutions. *J Electrochem Soc* 140:1009–1026

- Bean KE (1978) Anisotropic etching of silicon. *IEEE Trans Electron Devices* ED-25:1178–1185
- Buser RA, Rooij NF (1991) ASEP: a CAD program for silicon anisotropic etching. *Sens Actuators* 28:71–78
- Camon H, Moktadir Z (1995) Atomic scale simulation of silicon etched in aqueous KOH solution. *Sens Actuators* A46:27–29
- Danel JS, Delapierre G (1992) *Sens Actuators* 31:267
- Fruhauf J, Trautmann K, Wittig J et al (1993) A simulation tool for orientation dependent etching. *J Micromech Microeng* 3:113–115
- Gosalvez MA, Nieminen RM, Kilpinen P et al (2001) Anisotropic wet chemical etching of crystalline silicon: atomistic Monte Carlo simulations and experiments. *Appl Surf Sci* 178:7–26
- Gosalvez MA, Foster AS, Nieminen RM (2002) Atomistic simulations of surface coverage effects in anisotropic wet chemical etching of crystalline silicon. *Appl Surf Sci* 202:160–182
- Gosalvez MA, Xing Y, Sato K (2008) Analytical solution of the continuous cellular automaton for anisotropic etching. *J Microelectromech Syst* 17:410
- Gosalvez MA, Ferrando N, Xing Y, Pal P, Sato K, Cerda J, Gadea R (2011) Simulating anisotropic etching of silicon in any etchant: evolutionary algorithm for the calibration of the continuous cellular automaton. *J Micromech Microeng* 21:065017
- Gosalvez MA, Ferrando N, Fedoryshyn Y, Leuthold J, McPeak KM (2016) Evidence for faster etching at the mask-substrate interface: atomistic simulation of complex cavities at the micron-/submicron-scale by the continuous cellular automaton. *J Micromech Microeng* 26:045013
- Herr E, Baltes H (1991) KOH etch rates of high-index planes from mechanically prepared silicon crystals. In: *IEEE Transducers'91*, pp 807–810
- Hubbard TJ, Antonsson EK (1994) Emergent faces in crystal etching. *J Microelectromech Syst* 3:19–27
- Hubbard T, Antonsson EK (1997) Cellular automata in MEMS design. *Sens Mater* 19:437–448
- Kakinaga K, Tabata O, Baba N et al (2004) Simulation of anisotropic chemical etching of single crystalline silicon using cellular-automata. *IEEJ Trans Sens Micromachines* 124:7–13
- Liu C (2006) *Foundations of MEMS*. Prentice Hall, Upper Saddle River
- Madou M (1997) *Fundamentals of microfabrication*. CRC Press, Boca Raton
- Nishidate K, Baba M, Gayload R (1998) ASECA: a cellular-automata simulation program for silicon anisotropic super-micro-etching process in aqueous KOH. *Comput Phys* 12:88–93
- Pal P, Gosalvez MA, Sato K, Hida H, Xing Y (2014) Anisotropic etching on Si{110}: experiment and simulation for the formation of microstructures with convex corners. *J Micromech Microeng* 24:125001
- Petersen KE (1982) Silicon as a mechanical material. *Proc IEEE* 70:420–457
- Radjenovic B, Radmilovic-Radjenovic M, Mitric M (2006) *Appl Phys Lett* 89:213102
- Sato K, Shikida M, Matsushima Y, Yamashiro T, Asaumi K, Iriye Y, Yamamoto M (1998) Characterization of orientation-dependent etching properties of single-crystal silicon: effect of KOH concentration. *Sens Actuators* A64:87–93
- Schroder H, Obermeier E, Horn A et al (2001) Convex corner undercutting of {100} silicon in anisotropic KOH etching: the new step-flow model of 3-D structuring and first simulation results. *J Microelectromech Syst* 10:88–97
- Seidel H (1990) Anisotropic etching of crystalline silicon in alkaline solutions I. *J Electrochem Soc* 137:3612–3626
- Seidel H, Cseppregi L, Heuberger A, Baumgartel H (1990) Anisotropic etching of crystalline silicon in alkaline solutions II. *J Electrochem Soc* 137:3626–3632
- Sequin CH (1992) Computer simulation of anisotropic crystal etching. *Sens Actuators* 34:225–241
- Tabata O, Asahi R, Funabashi H, Sugiyama S (1992) Anisotropic etching of silicon in TMAH solutions. *Sens Actuators* A34:51–57
- Than O, Buttgenbach S (1994) Simulation of anisotropic chemical etching of crystalline silicon using a cellular automata model. *Sens. Actuators A* 45:85–88
- van Veenendaal E, Nijdam AJ, van Suchtelen J et al (2000) Simulation of anisotropic wet chemical etching using a physical model. *Sens Actuators A* 84:324–329

- van Veenendaal E, Sato K, Shikida M, van Suchtelen J (2001) Micromorphology of single crystal silicon surfaces during anisotropic wet chemical etching in KOH and TMAH. *Sens Actuators A*93:219–231
- Wind RA, Hines MA (2000) Macroscopic etch anisotropies and microscopic reaction mechanisms: a micromachined structure for the rapid assay of etchant anisotropy. *Surf Sci* 460:21–28
- Wind RA, Jones H, Little MJ, Hines MA (2002) Orientation-resolved chemical kinetics: Using microfabrication to unravel the complicated chemistry of KOH/Si etching. *J Phys Chem B* 106:1557–1569
- Xing Y, Gosalvez MA, Sato K (2007) Step flow-based cellular automaton for the simulation of anisotropic etching of complex MEMS structures. *New J Phys* 9:436
- Yun MH, Burrows VA, Kozichi MN (1998) Analysis of KOH etching of (100) silicon on insulator for the fabrication of nanoscale tips. *J Vac Sci Technol B*16:2844–2848
- Zhou ZF, Huang QA, Li WH et al (2007) A cellular automaton based simulator for silicon anisotropic etching processes considering high index planes. *J Micromech Microeng* 17: S38–S49
- Zhou ZF, Huang QA, Li WH et al (2009) Modeling and simulations of anisotropic etching of silicon in alkaline solutions with experimental verification. *J Electrochem Soc* 156:F29–F37
- Zhu Z, Liu C (2000) Micromachining process simulation using a continuous cellular automata method. *J Microelectromech Syst* 9:252–261
- Zubel I (2001) The influence of atomic configuration of (hkl) planes on adsorption processes associated with anisotropic etching of silicon. *Sens Actuators A*94:76–86
- Zubel I (2003) The model of etching of (hkl) planes in monocrystalline silicon. *J Electrochem Soc* 150:C391–C400
- Zubel I, Kramkowska M (2004) Etch rates and morphology of silicon (hkl) surfaces etched in KOH and KOH saturated with isopropanol solutions. *Sens Actuators A*115:549–556



# Modeling and Simulation of Silicon Dry Etching

Yan Xing

## Contents

Introduction .....	28
DRIE Process Simulation System .....	30
The Framework of Simulation System .....	30
DRIE Process Model .....	31
Effects in DRIE Process .....	33
Cellular Automation Method .....	34
Process Parameter Optimization .....	36
Particle Swarm Optimization .....	36
Principle of PSO Algorithm .....	37
Fitness Function for DRIE Model .....	38
Computational Model .....	39
Data Representation .....	39
Implementation of Cellular Automation for DRIE .....	41
Determination of Normal Vector .....	42
Mask Edge Recognition .....	44
PSO-CCA Method .....	45
Parameters in PSO-CCA .....	45
Parameter Calibration for DRIE Model .....	48
The Modified Process Model .....	50
The Parallel Computation for DRIE Simulation .....	52
Simulation of DRIE Applications .....	57
Summary .....	63
References .....	64

---

Y. Xing (✉)  
Southeast University, Nanjing, China  
e-mail: [xingyan@seu.edu.cn](mailto:xingyan@seu.edu.cn)

---

**Abstract**

As a widespread form of dry etching, deep reactive ion etching (DRIE) is a highly anisotropic etch process. It alternates switching the chemistry for etching and passivation cycles, typically leads to characteristic scalloping patterns on the sidewalls with high aspect ratios. Measurements of the etch depth per cycle  $l_d$  and undercut length per cycle  $l_u$  show a strong dependence of the undercut ratio  $l_u/l_d$  on the trench aspect ratio for a wide range of opening sizes. Although various simulation models have been proposed, the determination of the corresponding parameters from experimental data remains unsolved. We present the use of (i) the continuous cellular automaton (CCA), to simulate the process reliably in three dimensions; (ii) the particle swarm optimization (PSO) method, to determine suitable values for the atomistic CCA parameters directly from experimental data; and (iii) a GPU, parallel implementation of the CCA, to increase the computational efficiency of the simulations. The resultant, parameter-optimized CCA simulations show good agreement with the experiments. The approach has a large potential for the simulation of other MEMS processes.

---

**Keywords**

Dry etching · Deep reactive ion etching · Cellular automaton · CUDA

---

**Introduction**

Deep reactive ion etching (DRIE) process is a widely used technique for the manufacture of micro-electro-mechanical systems (MEMS). By repeatedly cycling etching and passivation procedures, this process enables the fabrication of sophisticated high aspect ratio microstructures on silicon (Rangelow 2003; Tsujimoto et al. 1986). The overall etch rate per cycle is a complex function of the polymer deposition rate, the polymer etch rate, and the silicon etch rate. Although the polymer deposition rate and the silicon etch rate are attenuated with increasing aspect ratio (due to the need for the neutral radicals to diffuse into the features), the polymer etch rate is essentially independent of the aspect ratio (due to its ion-driven, directional nature). Due to these features, the DRIE process has a number of well-known artifacts, which can seriously influence the performance of the resulting device. This includes scalloping and the lag effects (Zhibo et al. 2012; Saraf et al. 2011; Ranganathan et al. 2011; Mita et al. 2006; Lai et al. 2006; Chienliu et al. 2005; Knizikevičius 2003).

Scalloping consists of the development of crests on the trench sidewalls, especially for shallow depths (Mita et al. 2006). It is due to the isotropic nature of the silicon etch rate, which provides lateral propagation in addition to vertical advancement. The lag effect refers to the phenomenon where the etch depth of narrower trenches is smaller than that of wider trenches for the same etch time (Lai et al. 2006). It is due to the attenuation of the overall etch rate per cycle with increasing aspect ratio. The inverse-lag effect appears under conditions where the attenuations in the silicon etch rate and polymer deposition rate are out of sync (Knizikevičius 2003).

To predict and avoid these features – and aid the design of MEMS structures – several methods have been developed (Knizikevičius 2003; Ertl and Selberherr 2010; Tan et al. 2006; Wang et al. 2007). In this chapter we combine the continuous cellular automaton (CCA) method for modeling the propagation of the etch front and the particle swarm optimization (PSO) in order to find suitable values for the various parameters required to perform the simulations.

In general, cellular automata (CA) are dynamical systems for which time and space are discrete (Hernández Encinas et al. 2007). They consist of a collection of cells interacting locally with each other. Each cell is associated with a state that changes every time step according to the states of the neighbor cells at the previous time step. Their wide use across many areas reflects the good performance in the modeling of realistic problems (Hernández Encinas et al. 2007; Ferrando and Gosálvez 2012; Ferrando et al. 2011; Xing and Gosálvez 2007; Nagel and Schreckenberg 1992). In the CCA, the etching/growth process is regarded as a continuous decrease/increase of the occupation  $\Pi$ , defined for every cell. For etching,  $\Pi$  has an initial value of ‘1’ and is gradually decreased at each time step on the cells located at the etch front. When  $\Pi$  reaches the value ‘0’ the cell is removed, leading to the emergence of one or several of its bulk neighbors to the surface (etch front). For growth, the reverse procedure is applied. This provides a mechanism to continuously incorporate new cells into the etch/growth front and enables the description of the splitting and/or coalescing of disjoint regions of the etch/growth front without any special programming effort.

Similarly, PSO is an optimization algorithm for the search of parameter values in systems with multiple variables (Kennedy and Eberhart 1995). Inspired on the behavior of foraging swarms, each individual in the swarm, called a particle, follows very simple rules, but the global behavior becomes rather complex due to the communication and cooperation between the particles. In this context, each particle is a candidate solution to the problem, and the complex global behavior is used to optimize the solutions. When the search is finished, the best solution of the swarm is regarded as the optimal solution. PSO has been frequently applied to solve complex optimization problems due to the ease of implementation, minimum mathematical complexity, and good performance (Kennedy and Eberhart 1995; Alam et al. 2014; Shi et al. 2011; Shi and Eberhart 1998; Poli et al. 2007).

To generate realistic DRIE simulations, the actual conditions of the reactor need to be introduced into the simulation environment. In practice, however, most theoretical models consider phenomenological parameters that are only indirectly related to the actual process parameters. Nevertheless, the models typically perform surprisingly well. Correspondingly, in this study we implement DRIE models, thus stressing the flexibility of the CCA to describe different simulation strategies and the ability of the PSO algorithm to intelligently adjust different combinations of model parameters by only using a small set of experimental data. In particular, we consider the process calculation as a complex theoretical description of the DRIE process.

## DRIE Process Simulation System

### The Framework of Simulation System

Figure 1 shows the infrastructure of the proposed DRIE process simulation system in this chapter. The simulation system has three parts: 3D DRIE process simulation, DRIE process model optimization, and CUDA parallel computing acceleration.

The process simulation module performs DRIE process calculation and visualizes the simulation result. The input information of the module includes mask layout, the model settings, and process parameters such as conditions, etching rate, and alternating etching and passivation cycles. The process simulator converts the input information, sets up a 3D cellular automation calculation model, and performs evolution until reaching predefined etch time. Finally, the visualization part outputs the 3D structure as well as the morphology of the etching surface.

Since different DRIE system achieves different etch result, an efficient optimization method is employed to fit the parameters of theoretical simulation model for better accuracy. After the simulation result converges to the experimental data, the system is able to provide accurate result for the specified machine settings. In case of the complex mask layout and high density of mask patterns, parallel computing acceleration is necessary for the system to carry out time-consuming calculation.

By combining the optimization process and DRIE model, this chapter proposes a flexible method for the calibration of DRIE simulations. The DRIE process model

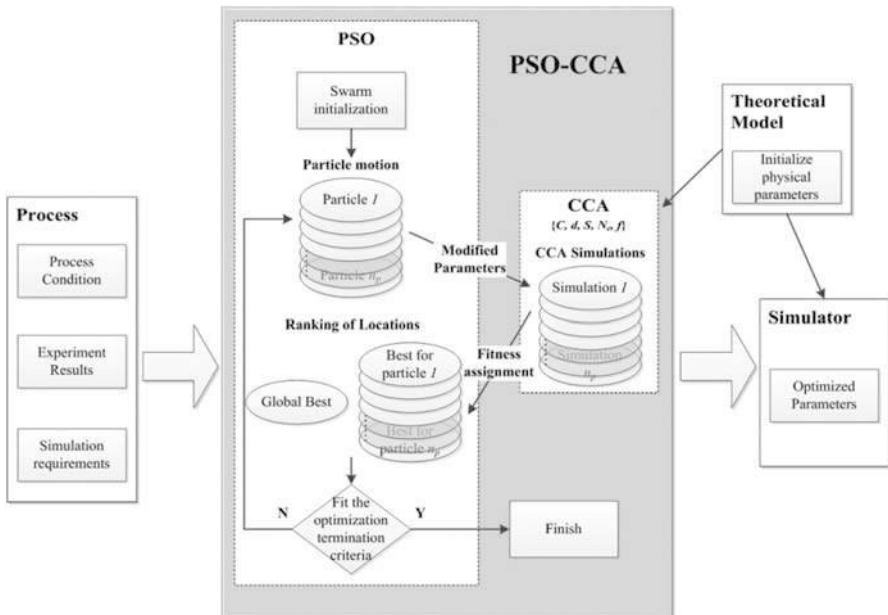


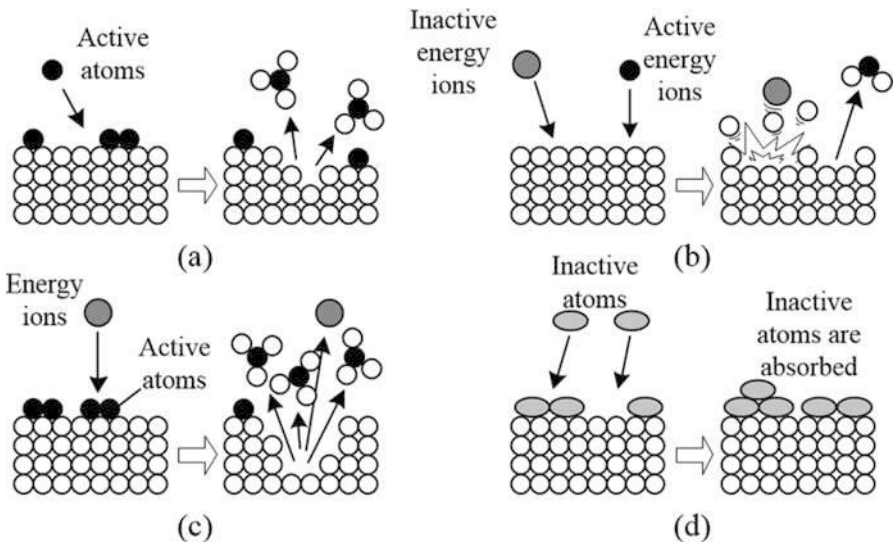
Fig. 1 The PSO-CCA method procedure



offers a suitable framework to propagate the etch/growth front, automatically describing the splitting and/or coalescing of disjoint regions without any special programming effort. The optimization algorithm adjusts the parameters of the underlying physical model to describe any particular process condition, using only a small set of experimental measurements, such as the dependence of the scalloping undercut per cycle and/or the depth per cycle on the number of cycles. Typical artifacts of the DRIE process are reproduced in the simulations, including the scalloping and its attenuation with increasing aspect ratio as well as the lag effects. Good agreement with experiment demonstrates that the model can be used to guide the process design. This indicates that the proposed system can be generalized to model the DRIE process with other physical models, simultaneously fulfilling different needs to achieve optimization.

### DRIE Process Model

Figure 2 shows reactions in the DRIE process. In Fig. 2a it indicates a thermally induced chemical reaction, which is a reaction that takes place by the adsorption of active atoms on the surface of the etched material. Figure 2b shows ion sputtering, involving physical sputtering and chemical sputtering effects. In physical sputtering, the kinetic energy of nonactive ions transfers to atoms on the surface of the material due to collision. In chemical sputtering, the incidence of reactive ions leads to desorption of surface atoms. Figure 2c demonstrates ion assisted etching. This reaction happens when ions with energy bombard active atoms on surface, resulting



**Fig. 2** Surface reactions in DRIE (a) chemical reaction; (b) sputtering; (c) ion assisted etching; (d) deposition

in chemical reaction between active atoms. Figure 2d describes the process of deposition, in which inactive atom groups adhere to the surface of the material gradually.

### Etching Model

Chemical reaction and ion assisted etching are the main reactions on surface of substrate. In this chapter, the etching rate of DRIE has two parts, namely:

$$E = E_{\text{er}} + \text{ER}_{\text{iic}} \quad (1)$$

where  $E_{\text{er}}$  is the chemical etch rate. It is associated with parameters of the process; the expression of  $E_{\text{er}}$  can be as the following:

$$E_{\text{er}} = k_0 F^r \exp\left[-\frac{E_e}{k_B T_s}\right] \cdot R_s \quad (2)$$

where  $E_e$  is active energy,  $K_0$  is constant, and  $F^r$  is the flow of active group; these parameters are determined by the initial process conditions.  $T_s$  is the temperature of the substrate, and  $k_B$  is the Boltzmann constant, the value is  $k_B = 1.38 \times 10^{-23} \text{ J/K}$ .

One of the most important parameters in chemical etching rate is  $R_s$ , which is the contact rate of the reaction gas on the etched surface. Since the etch chamber is filled with gas,  $R_s$  is essentially the volume fraction of the gas in the limited space where one position (cell) of the surface is etched away. This parameter ensures the isotropic nature of the chemical etch.

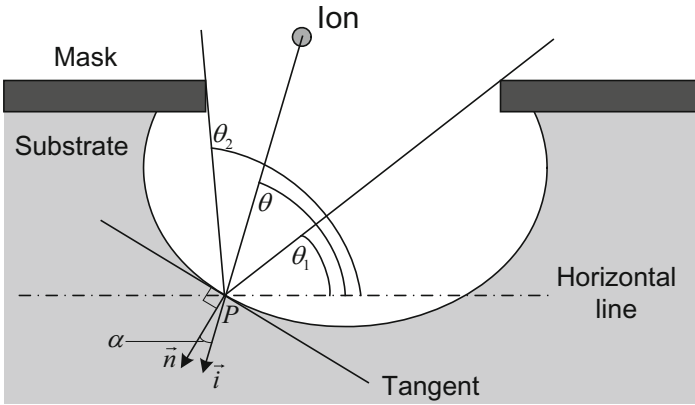
The rate of ion assisted etching  $\text{ER}_{\text{iic}}$  can be expressed as the following:

$$\text{ER}_{\text{iic}} = C_i \int_0^\pi \int_{\theta_1}^{\theta_2} \cos \alpha \cdot J(\varphi, \theta) d\theta d\varphi \quad (3)$$

Here,  $C_i = cJ_i Y_{\text{ast}}$ ,  $c$  is the constant value relating to substrate material to be etched, the value of it is  $c = 1/\rho = 1/(5.0 \times 10^{22}) = 2 \times 10^{-23}$ ;  $J_i$  is the ion flux;  $Y_{\text{ast}}$  is the ion assisted etch rate,  $Y_{\text{ast}} = A(\sqrt{E} - \sqrt{E_{\text{th}}})$ ; and  $Y_{\text{ast}}$  is proportional to the square root of the ion energy.

The distribution function of ion source  $J(\varphi, \theta)$  is a two-dimensional normal distribution that one-dimensional normal distribution  $J(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\theta - \frac{\pi}{2})^2}{2\sigma^2}\right]$  rotates  $180^\circ$  around the axis of  $\theta = \frac{\pi}{2}$ . For a certain angle  $\varphi$ , the diagram of each angle is shown in Fig. 3, where  $\theta$  is the angle between incident ray of ion and horizontal line.

As Fig. 3 shows  $\vec{n}$  is the unit normal vector,  $\vec{i}$  is the vector of incident ion,  $\alpha$  is the angle between normal vector and the vector of incident ion, and  $\theta_1$  and  $\theta_2$  are angles between the edge of opening mask for an incident position on the surface of substrate.



**Fig. 3** The diagram of relative angle of ion assisted etching

In this way, total rate of etching can be expressed as the following:

$$E = k_0 F^r \exp \left[ -\frac{E_e}{K_B T_s} \right] \cdot R_S + C_i \int_0^{\theta_2} \int_0^{\pi} \cos \alpha \cdot J(\varphi, \theta) d\theta d\varphi \quad (4)$$

**Deposition Model**

Deposition for passivation procedures protects the vertical sidewall in DRIE process. Since the deposition rate and the chemical etching rate are both isotropic, the deposition rate is a linearly proportional to the chemical etching rate in this study. We define  $Dr = k_D \cdot E_{\text{ets}}$ , here  $k_D$  is the parameter for deposition.

**Effects in DRIE Process**

**Lag Effect**

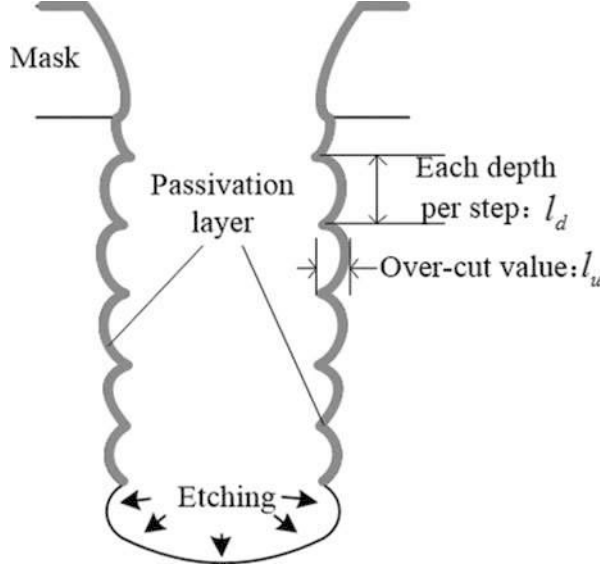
In deep reactive ion etching and reactive ion etching, the effect of the aspect ratio in the etch rate results in lag effect, which is characterized by the fact that the etch rate gradually decreases as the aspect ratio increases. A trench with a small mask opening will etch less than the trench structures under a large mask opening.

In the etch rate equation,  $\theta_2 - \theta_1$  is the angle of the substrate exposed to incident ions in between two sides of the mask opening, the etch rate increases as the angle increases.

**The Scalloping Sidewall**

The etch and deposition alternate in one etch cycle of DRIE operation, which creates another characteristic on etched profile. The sidewalls after etching consist of a series of continuous scalloping shape. Figure 4 shows a typical DRIE sidewall. We define

**Fig. 4** Typical DRIE scalloping sidewall



$l_d$  as the etch depth in one etch cycle; it is the height of the scalloping shape.  $l_u$  indicates the undercut value achieved in one DRIE cycle.

For the MES application, it is necessary to precisely control the ratio of  $l_d$  and  $l_u$  in the each etch cycle to get as smaller as possible  $l_u/l_d$ . The values  $l_u$  and  $l_d$  decrease with the depth of etching, which leads to the decrease of  $l_u/l_d$  with the increase of etching depth.

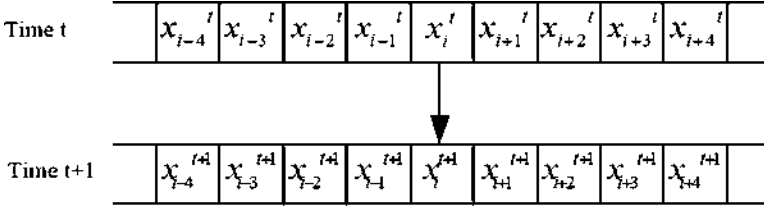
The distribution function of ion source is

$$J(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[ -\frac{(\theta - \frac{\pi}{2})^2}{2\sigma^2} \right] \quad (5)$$

In this normal distribution function, the standard deviation  $\sigma$  affects the shape of distribution. The larger the value of  $\sigma$ , the more dispersed the ions are. Therefore, the larger the standard deviation of the ion source distribution function, the more dispersed the ions and the smaller the ion flux between  $\theta_1$  and  $\theta_2$ . As a result, it leads to decrease of the ion assisted etching rate and etching depth.

## Cellular Automation Method

Cellular automation is a dynamic system that its time and space are discrete. Each point of the system scattered into the regular grid takes a finite discrete state, follows the same rules of reaction, and updates synchronously according to the determined local rules. A large number of grids form the evolution of dynamic systems through



**Fig. 5** One-dimensional cellular automation

simple interactions. The biggest difference between a cell algorithm and a line algorithm is that the grid point of the substitution object is fixed and the shape of the object is determined by the state of the many lattice points that are divided, rather than by the movement of the lattice like the line algorithm.

In the cellular automation, each unit that is scattered in the regular grid is called cell. Other cells that are interacted with object cellular are called neighbors. Each cell only takes a finite discrete state, and this state is determined by the state of neighbors and specific local rules, the rules are known as the transition rule. A large number of cells refresh their statuses through interaction, so as to constitute the dynamic evolution of the system. Figure 5 shows the process of evolution of one-dimensional cellular automation from time  $t$  to  $t + 1$ .

In Fig. 5, it assumes that the one-dimensional cellular automaton is a four-neighbor cellular automaton, so the  $i$ th cell's status at the time  $t + 1$  is in this formula, here function  $f$  is the rule of local transformation:

$$x_i^{t+1} = f(x_{i-2}^t, x_{i-1}^t, x_i^t, x_{i+1}^t, x_{i+2}^t) \tag{6}$$

The continuous cellular automaton (CCA) model for 3D simulation of DRIE can be defined as the set  $\{C, d, S, N_e, f\}$ .  $C$  is the simulation space discretized into cells.  $d$  is a vector that indicates the size along each dimension of  $C$ . When  $d = [n_i, n_j, n_k]$ ,  $C$  is a 3D domain formed by  $n_i \times n_j \times n_k$  cells:  $C = \{(i, j, k), 1 < i < n_i, 1 < j < n_j, 1 < k < n_k\}$ . Each cell  $(i, j, k)$  stores the state of the substrate site located at the corresponding coordinate position. The state of the cell  $(i, j, k)$  at iteration  $n$  is the combination of the material type  $\mu_{ijk}^{(n)}$  and the material occupation  $\pi_{ijk}^{(n)}$ .  $\mu_{ijk}$  is an element of a finite set  $S$  of material types, for example,  $S = \{\text{silicon, polymer, mask, etchant}\}$ , and  $\pi_{ijk}$  is a continuous variable in the range of  $[0, 1]$ . When its value changes over the thresholds (0 or 1), the material type  $\mu_{ijk}$  will also change. The occupation and the material type are updated according to the following functions:

$$\pi_{ijk}^{(n+1)} = \pi_{ijk}^{(n)} + f_{oc}(N_{e(ijk)}^{(n)}) \cdot \Delta t$$

$$\mu_{ijk}^{(n+1)} = f_m(\mu_{(ijk)}^{(n)}, \pi_{(ijk)}^{(n+1)})$$

where  $f_{oc}$  is the rate function of the occupation and  $f_m$  is the material update function. The updated cell occupation is a function of the previous states of the cells constituting its neighborhood  $N_e$ . The updated cell material type is determined by the previous type and the current occupation. For 3D simulation, the 3D Von Neumann neighborhood is considered:

$$N_{e(ijk)} = \{(i-1, j, k), (i, j, k+1), (i, j+1, k), (i, j, k), (i, j-1, k), (i+1, j, k), (i, j, k-1)\}$$

$f = \{f_{oc}, f_m\}$  is the function set that contains the update rules. When the graphical results are presented,  $\mu_{ijk}$  is used to mark the material state of each space part and  $\pi_{ijk}$  is used to reconstruct the interface between different materials inside the cells by the volume-of-fluid method (Hong et al. 2000; Weymouth and Yue 2010; Pilliod and Puckett 2004). The CCA results presented in this study are visualized using open inventor.

---

## Process Parameter Optimization

### Particle Swarm Optimization

To fit the parameters of the theoretical model in simulation to fabricated profile, it is essential to combine an optimization method to determine the parameters in DRIE process model. In recent years, many optimization algorithms have been put into practice, like genetic algorithm, artificial neural network, and population algorithm. The particle swarm optimization (PSO) algorithm is introduced in this chapter to adjust the parameter values of a given DRIE physical model in order to describe different process settings. The method considers a population of candidate solutions, referred as particles, each consisting of a set of values for the tunable parameters. The particles are moved around in the parameter space according to velocities that take into account the actual difference between each particle's position and the best individual (overall) position for each particle (all particles), updating the best individual and global positions as better positions are gradually found. The relative performance of each particle is determined by a fitness function. The key aspects are (i) the definition of the particles, (ii) the cognitive ability and social adaption of the particles, and (iii) their intelligent locomotion (Kennedy and Eberhart 1995; Alam et al. 2014; Shi et al. 2011; Shi and Eberhart 1998; Poli et al. 2007).

PSO has been first put forward by Eberhart and Kennedy. Different from genetic evolution, in PSO, the force that is driving each particle to the solution is adjusted all the time by its experience. The population of these particles is a social group, which makes the PSO algorithm more intelligent and plays an important role in solving the optimization problem. In this chapter, the optimization of DRIE process model is based on particle swarm optimization algorithm.

## Principle of PSO Algorithm

Individuals in PSO are regarded as points and scattered in D-dimensional search space. The  $i$ -th particle can be expressed as  $x_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{iD})$ . The optimal position calculated by previous experience values of the particles is  $p_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{iD})$ , and the moving speed of the particle is denoted by  $v_i = (v_{i1}, v_{i2}, v_{i3}, \dots, v_{iD})$ . The particle with the best position among all particles is marked as  $g$ . The movement of the particles follows the next equation:

$$\begin{aligned} v_i(t+1) &= w \cdot v_i(t) + c_1 \cdot \phi_1 \cdot (p_i(t) - x_i(t)) + c_2 \cdot \phi_2 \cdot (p_g(t) - x_i(t)) \\ x_i(t+1) &= x_i(t) + v_i(t+1) \end{aligned} \quad (7)$$

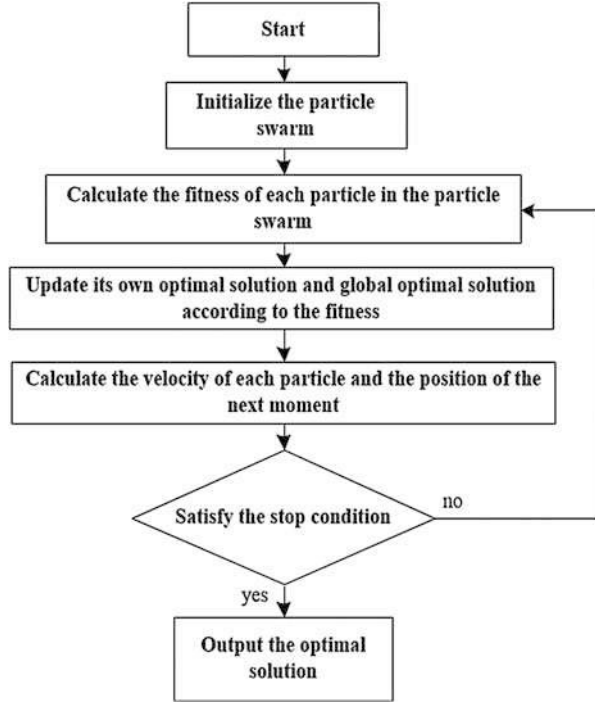
where  $c_1$  and  $c_2$  are positive constants,  $\phi_1$  and  $\phi_2$  are two random numbers between 0 and 1. According to the velocity formula, the velocity of particles is composed of three parts. The first one,  $w \cdot v_i(t)$ , is the inertia of the particle,  $w$  is constant, and  $v_i(t)$  is the speed of the particle in previous time. When  $w$  is large number, the speed in the past contributes more for the next moment so  $w$  is also called the inertial factor. The second part  $c_1 \cdot \phi_1 \cdot (p_i(t) - x_i(t))$  is the recognition of the particle itself, which is obtained by comparing the best position  $p_i(t)$  in the memory and the particle position  $x_i(t)$  in the previous time. The third part  $c_2 \cdot \phi_2 \cdot (p_g(t) - x_i(t))$  is the contribution of the whole population to the individual particle, and the value is obtained by comparing the optimal position  $p_g(t)$  of the whole population and the previous position  $x_i(t)$ .  $c_1$  and  $c_2$  are called learning factor for the process model. The learning factors  $c_1$  and  $c_2$  represent the self-cognitive ability of the particle and the ability of the cognitive society, respectively. In the case that  $c_1$  is smaller than  $c_2$ , the individual particle behaves more obediently to the population, the particle swarm will quickly converge toward the optimal solution. The particle swarms show the status of less collaboration. As a result, the convergence speed becomes slow and the system is even difficult to get the optimal point. When  $c_1$  is smaller than  $c_2$ , the particle behaves more obediently to the population, so that the particle swarm may quickly converge toward the optimal point. In fact, the setting of learning factors is complicated to decide in the optimization problem.

Different from the selection mechanism of genetic algorithm, PSO has no elimination of bad particles. All the individuals in the population work together until they reach the optimal area. And the driving force of the particle is updated according to the best position of the particle. Compared with the survival of the fittest method, the particle swarm has a senior collaborative social behavior.

The basic process of PSO is shown in Fig. 6:

- (a) Determine the number of particles in the particle swarm and the number of iterations, set the range of parameters, and initialize the program.
- (b) Calculate the fitness value of each particle according to the fitness function.
- (c) Comparing the fitness value with optimal solution, update the individual optimal solution and the global optimal solution.
- (d) Update the velocity and position of the particles according to the system function.

**Fig. 6** The structure of PSO program in DRIE process optimization



- (e) Check condition of the iteration. If neither the maximum number of iterations nor the error to the experiment is still large, the process continues to the step 2. Otherwise, the program is terminated and the global optimal solution can be set for the specified process model.

### Fitness Function for DRIE Model

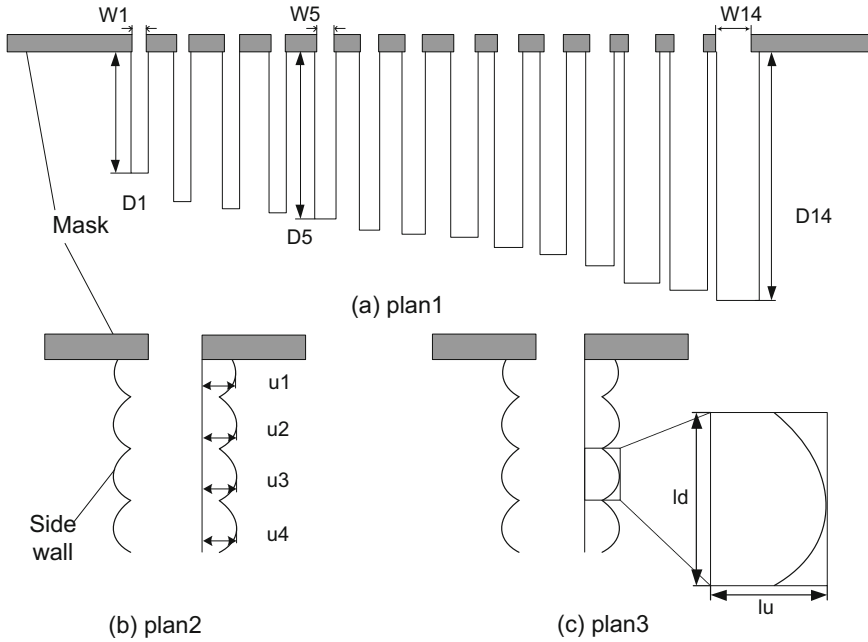
The optimization goal in the DRIE model is to minimize the difference between the simulation results and the experiment. Suppose that in the measurement of profile features from the experiment, the value of  $i$ -th sample is  $l_i$ , the corresponding result for simulation is  $l_{Sk}$ . The total error is

$$\min f = \min \left( \sum_{k=1}^n |l_k - l_{Sk}| \right) \quad (8)$$

In the DRIE, there are two kinds of essential effects, namely lag effect and scalloping undercutting, which have been considered in the fitness function.

In Fig. 7, Plan 1 only aims to lag effect. The DRIE process is performed on a mask with a series of openings. After a certain period of time, the corresponding etching depth  $D_k$  with the width  $W_k$  is measured. The set of  $\{(W_k, D_k) | k = 1, 2, \dots, n\}$  is the optimization target of the process model.





**Fig. 7** Optimization profile features from the experiment

Plan 2 considers undercut per cycle. The  $uk$  is the undercutting size for  $k$ -th cycle is measured, and the set of  $\{(k, uk) | k = 1, 2, \dots, n\}$  is the optimization target for the process model.

Etching rate of DRIE is inversely proportional to the ratio of depth and width of the trench. However, the inverse relationship between the etch rate and aspect ratio is caused by the etch depth in one single-step. In other words, In plan 3, on the basis of measuring each scalloping undercut value  $lu$  and etch depth  $ld$ , we can take the set of  $c\{(lu_k, ld_k) | k = 1, 2, \dots, n\}$  as the optimization target of the process model. This scheme can be optimized for both Lag effect and scalloping undercut.

Fig. 7 demonstrates the etching depth  $l_d$  and the undercutting value  $l_u$  of each etching and passivation circle. The optimization goal is  $\min \left( \sum_{k=1}^n |lu_k - lu_{sk}| + |ld_k - ld_{sk}| \right)$ .

## Computational Model

### Data Representation

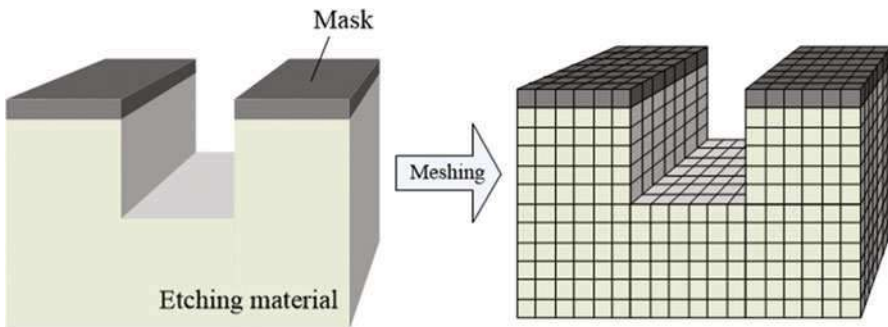
Cellular automaton is composed of basic regular grids. The algorithm defines the substrate of DRIE system into grids by the shape of cellular hexahedron.

As is shown in Fig. 8, the geometry shape of the etched structure is possible to express by cell definition. In order to determine whether the cell is etched, a parameter recording the residual solids in each cell is introduced. During the simulation, the program traverses all the cells and calculates the etch rate, and decides whether the target cell is etched away (set to air) or not.

The representation scheme of data used in a system determines the efficiency of the data storage, the computing costs, and the modeling abilities. In order to describe the dynamic movement of the etched surface and the displacement of substrate, two conventional methods have often been used in previous studies. One of them is the dynamic method. Instead of allocating memory for the whole substrate using a three-dimensional array, the method allocates and de-allocates the memory on a layer-by-layer basis, as required by the system evolution. This method is very useful for the case of vertically propagating surfaces, where allocation of new layers always occurs at the bottom region and de-allocation of already used, empty layers occurs at the top region.

As an alternative to the above methods, there are algorithms which can avoid the representation of empty portions of the system far from the surface (i.e., in the bulk and empty) as well as provide fast simulations even in the case of larger neighborhoods. For the purpose of simulating complex MEMS engineering applications, the etch front is typically a multivalued surface, not a solid-on-solid system. This means that the surface advances towards arbitrary directions depending on the mask patterns and the anisotropy of the process. In this context, the octree structure has the ability to model the dynamic movement using minimal computing resources. An octree-based cellular automata is suitable for a fast simulation program.

As a hierarchical representation, the octree structure is a tree in which every branch contains eight subbranches (or octants), whose status can be one of “empty,” “full,” or “partially full.” The tree is generated by recursively subdividing the cubic space into octants, effectively splitting the outermost cubic region into  $2^n \times 2^n \times 2^n$  unit cubes (or leaves), where  $n$  is the resolution parameter (or number of levels, that is, the number of subdivisions) and  $2^n$  is the one-dimensional size of the tree. The root node of the tree represents the entire space and the last level contains the leaves



**Fig. 8** Data representation of cellular system

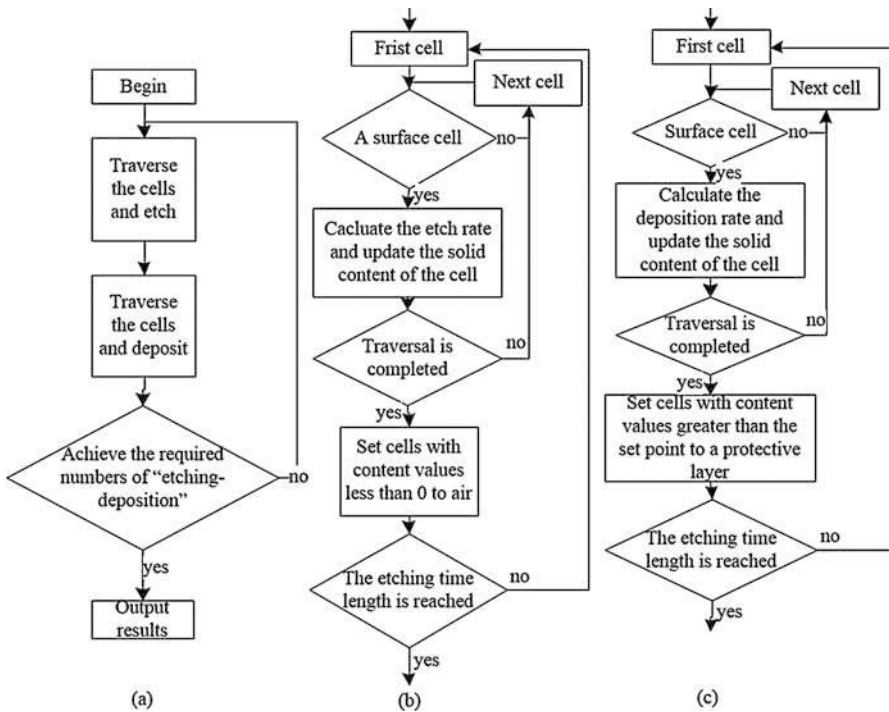
of the tree. In our implementation, each leaf is topologically a unit cell. In this study, each unit cell at the leaves level is also referred to as a voxel.

### Implementation of Cellular Automation for DRIE

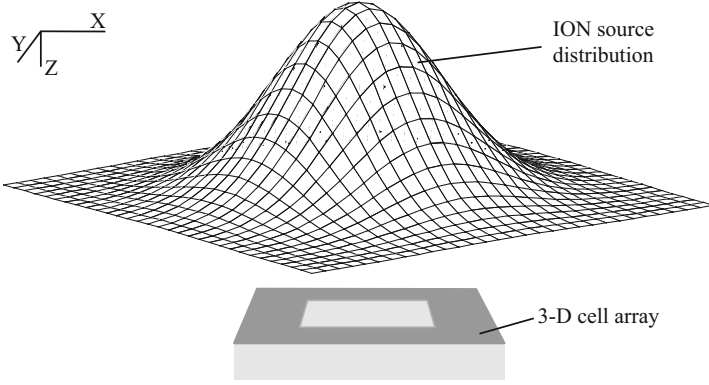
The basic flow chart of the program is shown in Fig. 9. The whole process of “etching-deposition” in DRIE system is described in Fig. 9a. Figure 9b, c demonstrates the flow of etching and deposition stages.

The deposition process is similar to that of the etching step, and there are two differences from etching. Firstly in calculating the rate of surface cells, we use a deposition model instead of an etch model. Secondly in updating cell’s statues value according to the growth rate, the calculation result is added on the original value.

Cellular automata system can be thought as a 3-D cellular array, for example, aCA3D[Dimz][Dimx][Dimy] formed in tree structure; one cell is labeled by aCA3D [k][i][j], whose coordinate is (x,y,z). It has variables members Specie and Amount that represent the cell kind and its solid content. There is a hypothetical ion source above the 3-D cellular array, and its distribution function is a 2-D normal distribution. In the ion assisted etching rate:



**Fig. 9** Flow chart of cellular automation simulation system, (a) total process; (b) the process of etching step; (c) the process of deposition



**Fig. 10** 3-D distribution of ion source

$$ER_{iic} = C_i \int_0^{\pi} \int_{\theta_1}^{\theta_2} \cos \alpha \cdot J(\phi, \theta) d\theta d\phi, \text{ when } \varphi = 0, J(0, \theta) \text{ is the normal distribution}$$

function of the ion source in the cross section of the  $\varphi$ -angle through Z-axis. Since the ions are uniformly distributed in different angular directions, the one-dimensional normal distributions of different cross sections such as  $J(\frac{\pi}{2}, \theta)$  and  $J(\frac{3\pi}{2}, \theta)$  are the same (Fig. 10).

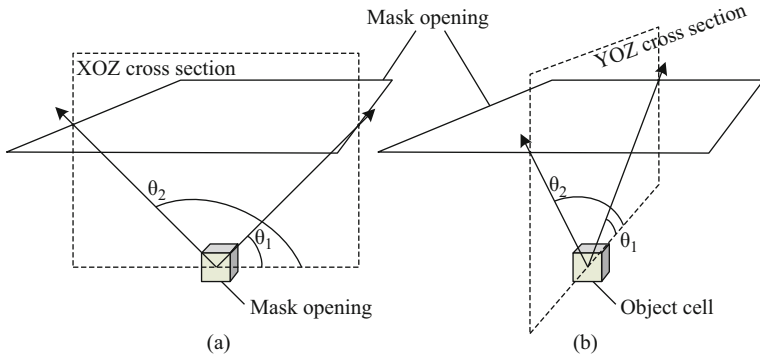
In calculating etch rate of surface cell under a certain mask opening, Z-axis of the cell usually is set as the center axis. The integral from 0 to  $\pi$  has to be divided by  $m$  parts, and the integral from  $\theta_1$  to  $\theta_2$  in each cross section should be divided into  $n$  parts. The calculating steps are as following:

1. Search the opening angles  $\theta_1$  and  $\theta_2$  under the mask in the  $i$ -th section cut as Fig. 11 shows.
2. Calculate the normal vector of the object cell according the neighbors statues of the target cell. Following the Eq. 3, the etching rate component for this section is calculated.  $\alpha$  is the angle between the normal vector and incident vector.
3. Calculate all the etch rate component for every section and multiply it by the step size  $\frac{\pi}{m}$  to obtain the total ion assisted etching rate for the target cell.

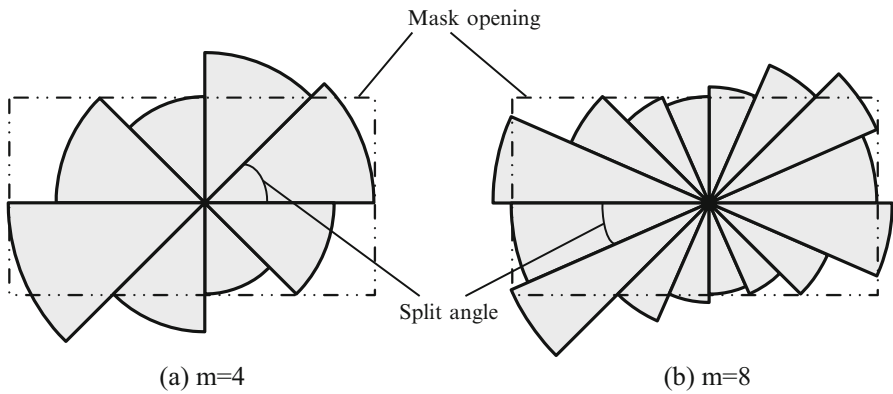
Figure 12 shows the sum steps for the Eq. 3.  $\pi$  is divided into four parts and eight parts. The shape of the mask opening is rectangular by the dashed line. Better accuracy of simulation result can be achieved by considering more section copies in the calculation. However, the increase in the number of sections means increasing computation costs for each surface cell.

## Determination of Normal Vector

In the simulation, another time-consuming part is the calculating the normal vector of the surface cell according to the etched profile. In the ion assisted etching rate



**Fig. 11** The angle of the window for 3-D cellular automata section

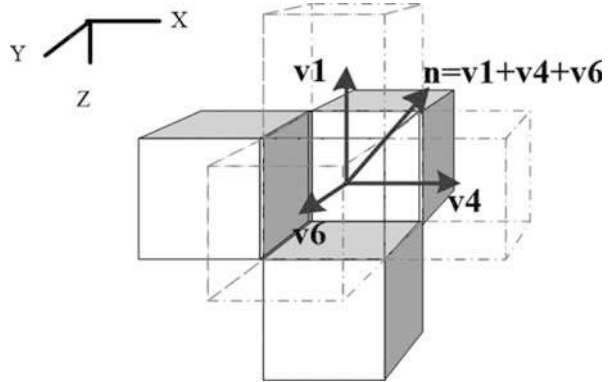


**Fig. 12** Schematic diagram for the summation of 8 and 16 parts of segmentation

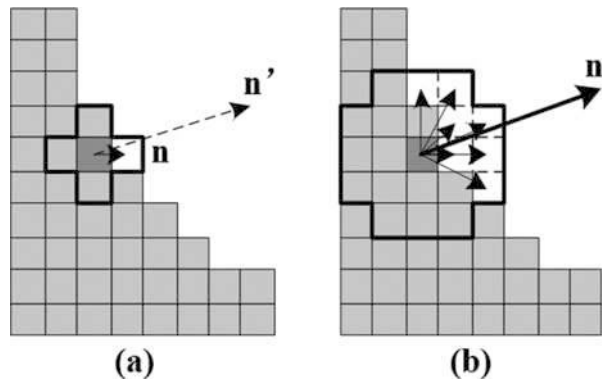
expression,  $\alpha$  is the angle between the normal vector and the incident vector of the object. The normal vector affects etch rate value significantly. Due to the definition of cellular automata, we use the sum of all neighbors' vectors as the normal vector of the target surface cell. The program visits all the neighbors. If the  $i$ -th neighbor is not on substrate, the object cell coordinates  $(x_0, y_0, z_0)$  are subtracted by its coordinates  $(x_i, y_i, z_i)$ . Since the neighbor cells and the object cells are adjacent, the sum of these vectors is approximation of the normal vector. Figure 13 is an example of six-neighbor cellular automata.

In Fig. 13, neighbors of the central target cell are labeled following the order of Z, X, Y direction. In this case, the upper cell is labeled No1, No2 is the bottom cell, the left and right ones are No. 3 and 4, and the back and front cell is labeled by No5 and 6. In this figure, 1, 4, and 6 are in the air (shown by the dotted line),  $\mathbf{v1}$ ,  $\mathbf{v4}$ , and  $\mathbf{v6}$  are for these three vectors of neighbors, the normal vector is  $\mathbf{n} = \mathbf{v1} + \mathbf{v2} + \mathbf{v3}$  and unit normal vector is  $\mathbf{n}_{unit} = \frac{\mathbf{n}}{|\mathbf{n}|} = \frac{\mathbf{v1} + \mathbf{v4} + \mathbf{v6}}{\sqrt{|\mathbf{v1}|^2 + |\mathbf{v4}|^2 + |\mathbf{v6}|^2}}$ .

**Fig. 13** The method of getting the normal vector in 3-D six-neighbor cellular automation



**Fig. 14** The method of calculating the normal vector before (a) and after the improvement (b)

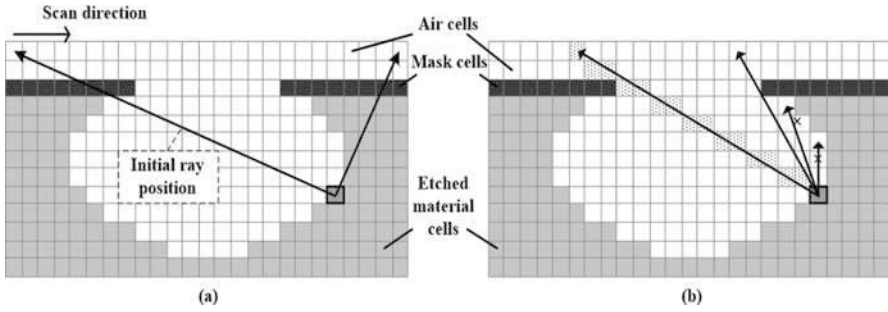


Due to the limited number of neighbors in such six-neighbor cell systems, only combinations of six-neighbor cell states bring errors. In order to obtain accurate normal vector, we can introduce more neighbors as shown in Fig. 14. Figure 14b shows the method of finding the normal vector after considering a wide range of neighbors. The normal vector  $\mathbf{n}$  is obtained by summing more vector components of the surface neighbor cells. In practice, this method has good accuracy.

## Mask Edge Recognition

In calculation of ion assisted etching rate, different angles  $\theta_1$  and  $\theta_2$  create the lag effect, they are decided by mask openings. Searching the open window for  $\theta_1$  and  $\theta_2$  is carried out by a kind of scan line algorithm. Figure 15 shows the method of obtaining the angles of the window  $\theta_1$  and  $\theta_2$  in the simulation system.

For etch object cell, scanning rays range from the left side of the substrate to the right as shown in Fig. 15. In order to represent the size of the opening window for incident ions, the angle between horizontal line and the ray from the left position of



**Fig. 15** The calculating method of the window angles: the initial state (a) and final state (b)

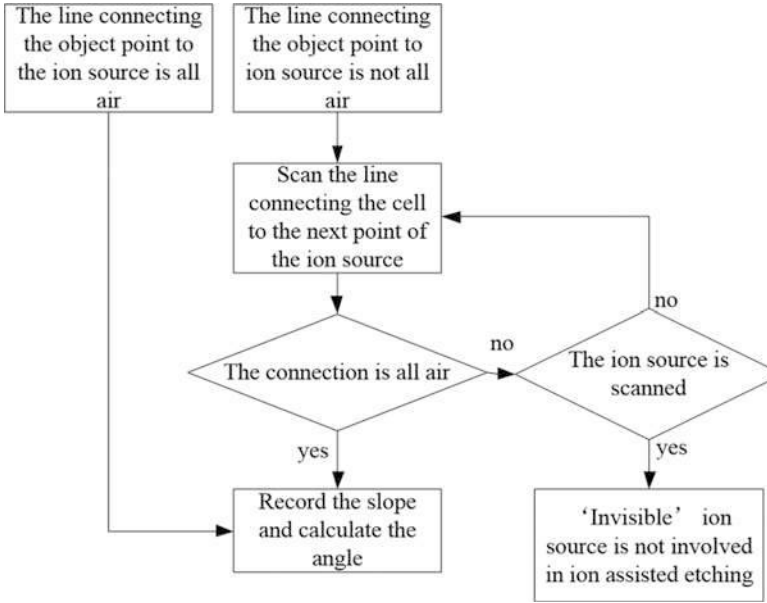
mask edge to the cell of the substrate is defined as  $\theta_1$ , and the angle between horizontal line and the ray from the right position of mask edge to the cell is defined as  $\theta_2$ . Figure 16 is the flow chart of the algorithm.

The process of scanning rays in this chapter is similar to the process of generating grid lines in computer graphics in Fig. 17. The program needs to calculate the coordinates of each cell on the ray, check the type of cell, and then determine whether the ray has passed through the solid cell or not. Linear generation algorithm is on the basis of method drawing complex geometry. This DRIE simulation embedded the Bresenham algorithm to generate a scan line in cellular automaton system. The advantage is the process has good calculation performance; all the time it only uses integer variables in a recursion sequence.

## PSO-CCA Method

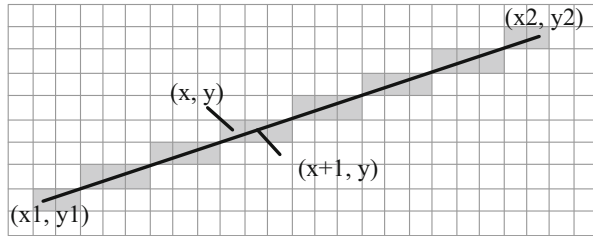
### Parameters in PSO-CCA

The PSO-CCA method is the combination of the PSO and CCA (Poli et al. 2007; Yuan et al. 2013; Li et al. 2015). It provides a means to adjust the parameters of a chosen theoretical model in order to describe the observed features of the corresponding experimental process, which is simulated by the CCA method. On the basis of the chosen experimental condition and theoretical model, a set of tunable parameters is identified in order to describe the actual location of the particles in the PSO procedure. After randomly initializing the swarm, all the particles move intelligently. Every visited location (modified parameters) is tested by performing a CCA simulation, thus comparing the result against the experiment (fitness assignment). The best locations are updated for each particle as well as for the overall swarm (ranking of locations). The procedure is continued, providing new particle locations and ranking them while updating the best individual/global locations, until the best global location meets the termination criterion. At this stage, the best global location is regarded as the optimal solution for the tunable parameters. This solution



**Fig. 16** The calculating flow of window angle

**Fig. 17** The ordinary linear generating algorithm



can then be used to reliably simulate the selected experimental process under the chosen conditions for different situations, such as the use of different mask patterns.

In this chapter, the fitness function of the particle swarm optimization algorithm is defined as

$$f = \sum_{k=1}^n (|lu_k - lu_{sk}| + |ld_k - ld_{sk}|) \tag{9}$$

where  $n$  is the number of “etch-deposition” cycles experienced in the DRIE,  $lu$  is the one step undercutting by etching, and  $ld$  is one step etch depth. The fitness value is the sum of the absolute values of the difference between the experimental results  $lu$  and  $ld$  and the simulated results  $lu_s$  and  $ld_s$ . In DRIE model, in order to adjust the ratio of the isotropic and the anisotropic etch rate, the chemical etch rate  $E_{er}$  and the



ion assisted etch rate  $ER_{iic}$  and two constants  $Mod_{Che}$  and  $Mod_{Ion}$  are introduced in the system; the total rate is expressed as

$$E = Mod_{Che} \cdot E_{er} + Mod_{Ion} \cdot ER_{iic} \quad (10)$$

The chemical etch rate consists of several parameters  $k_0$ ,  $F^r$ ,  $E_e$ ,  $k_B$ , and  $T_s$ .  $k_B$  is the Boltzmann constant,  $k_0$  is the process condition constant,  $F^r$  is the flow rate of active radicals,  $E_e$  is the activation energy, and  $T_s$  is the temperature. These parameters are fixed when the equipment condition is set.

The ion assisted etch rate part contains the constant parameter  $C_i$  which is associated with the spacing of atoms.  $\theta_1$  and  $\theta_2$  and  $\alpha$  are the variables according to the etched profile. The standard deviation of the ion source distribution function  $\sigma$  determines the ion concentration of the ion source. It is a choice as parameters in optimization process. Since the etch profile is affected by the time duration of one step etching, this control variable is represented by the EtchStep of the etch cycle in the simulation program. The larger the number of EtchStep, the longer the duration of one single step etching. All the parameters of the DRIE model used in the optimization are  $Mod_{Che}$ ,  $Mod_{Ion}$ , standard deviation  $\sigma$ , and EtchStep.

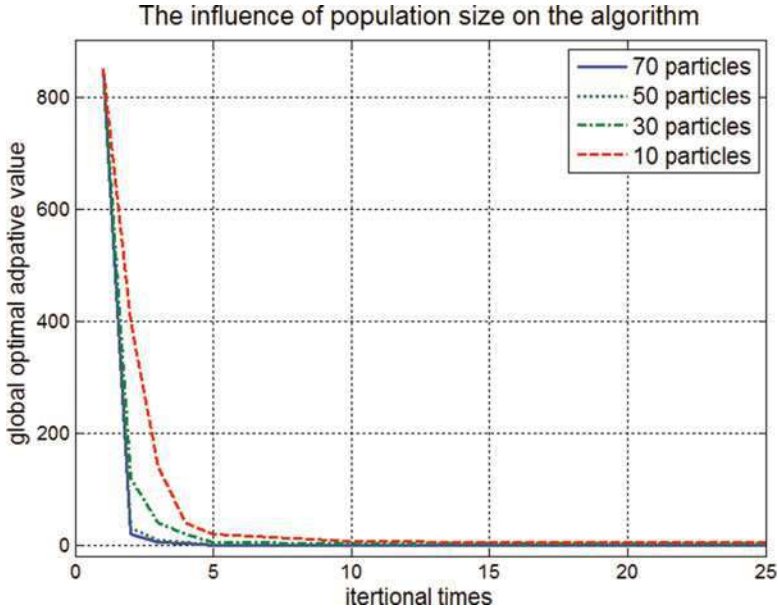
PSO convergence relates to the initial parameter values. The initial parameters also include the numbers of particles in population, momentum factors  $w$ , as well as learning factors  $c_1$  and  $c_2$ . These parameters are usually set differently in different optimization problems.

The fitness function to search optimal parameters of PSO-CCA model is defined as the following equation:

$$\sum_{k=1}^n (|lu_k - lu_{Sk}| + |ld_k - ld_{Sk}|) = f(Mod_{Che}, Mod_{Ion}, \sigma, EtchStep) \quad (11)$$

Population size is the numbers of the particles in population. Because of the collaboration property in PSO method, theoretically a system with more the particles behaves smarter and can fast converge to the optimal solution. But as the number of particles increases, it will take more computation time. For the PSO-CCA method, Fig. 18 shows global optimal fitness value with the population size of 10, 30, 50, 70 in the simulation. As the number of the particles in the population increases, the speed of the convergence increases. Global optimal values will approach the optimum solution after only tens of iterations. As particle number increases up to 50, the convergent speed of this algorithm gets its limitation.

Momentum factor  $w$  describes the particle inertia property. In PSO process, a particle will maintain part of its property when communicating with other particles to determine the direction of next movement. The proportion of the reserved speed is controlled by the momentum factor  $w$  whose value is in between 0 and 1. In case that  $w$  is 0, the particle will totally follow the speed from the external learning. As a result those particles will quickly gather together, which means that the solution will converge to the local optimal solution and on longer being searched again. When  $w$  is close to 1, those particles are difficult to change their directions, which mean



**Fig. 18** The influence of population size on the algorithm

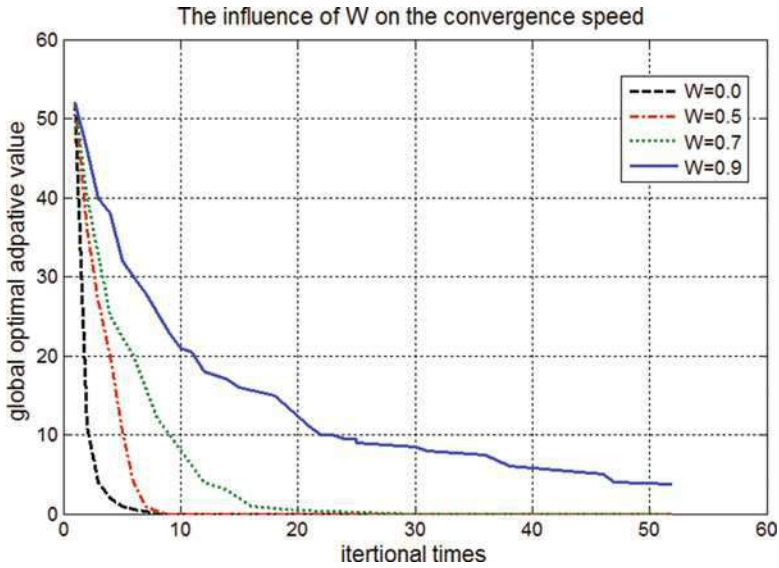
that those particles have no collaboration to get the optimal point. Figure 19 shows the plot of optimal fitness value with different momentum factor  $w$  at (0.0, 0.3, 0.6, and 0.9) for the PSO-CCA system.

Convergent speed of PSO changes with different  $w$  value. In fact, when the momentum factor is a large number, it takes long time to get the optimal resolution.  $c_1$  and  $c_2$  are learning factors.  $c_1$  refers to the memory of the best position that the particles have experienced in the past, which means the study by itself.  $c_2$  refers to the recognition of the particles in the group, which means learning from the social group. The configuration of the learning factors is predefined in this study. The distribution of different configuration of learning factors to get the average global optimal solution is shown in the following figure. According to the simulation result with different parameters in PSO-CCA method, the initial number of population is 30, the momentum factor  $w$  is 0.5, learning factor  $c_1$  is 2, and  $c_2$  is 1.5 (Fig. 20).

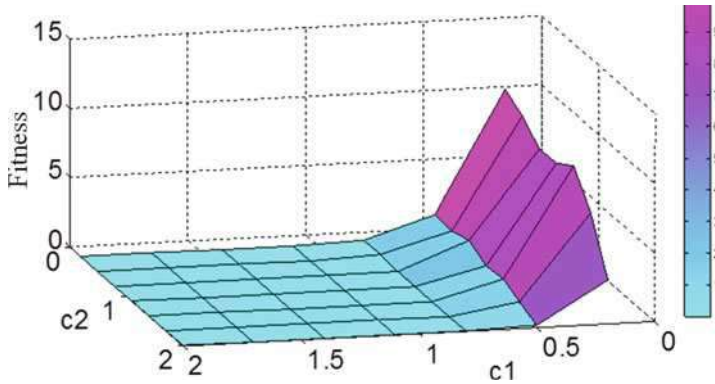
## Parameter Calibration for DRIE Model

In the calibration process, firstly, the program initializes the number of particles, defines search intervals for every parameters as the following:

$$\text{Mod}_{\text{Che}} \in [0.5, 1.5], \quad \text{Mod}_{\text{Ion}} \in [0.001, 0.0025], \quad \sigma \in [0.2, 0.4] \quad \text{and} \\ \text{EtchStep} \in [150, 250].$$



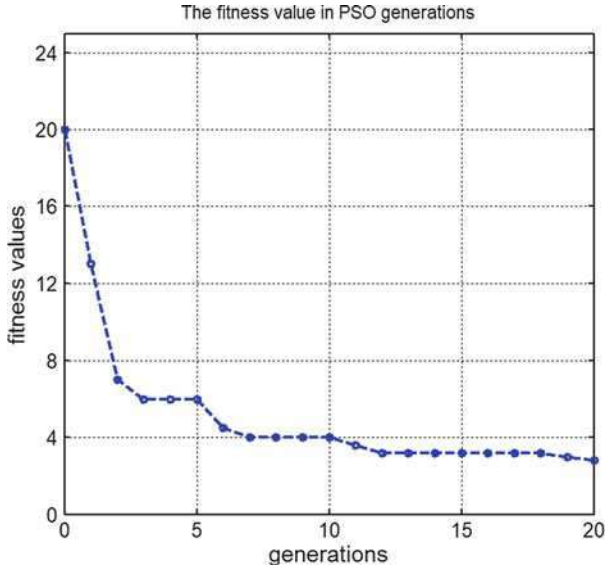
**Fig. 19** The influence of momentum factors on the algorithm



**Fig. 20** The influence of learning factors on the algorithm

Then the program uses random value to calculate the position and velocity for every particles. The max number of iterations to achieve the optimal solution is set as the termination condition of the algorithm if it fails to search the optimal parameter.

Secondly, every particle runs the program of DRIE process with its own parameter set. After certain steps of gas chopping process, the etch depth  $l_d$  and undercut dimension  $l_u$  of every etching step are retrieved from the etch profile so that the fitness values are able to be calculated by Eq. 11.



**Fig. 21** The fitness value in PSO generations

Finally according to the experimental data of  $ld$  and  $lu$ , the global optimal solution can be achieved for specified DRIE process. By using the set of tuned parameters, the complex structures by a wide range of mask patterns can be simulated with high accuracy.

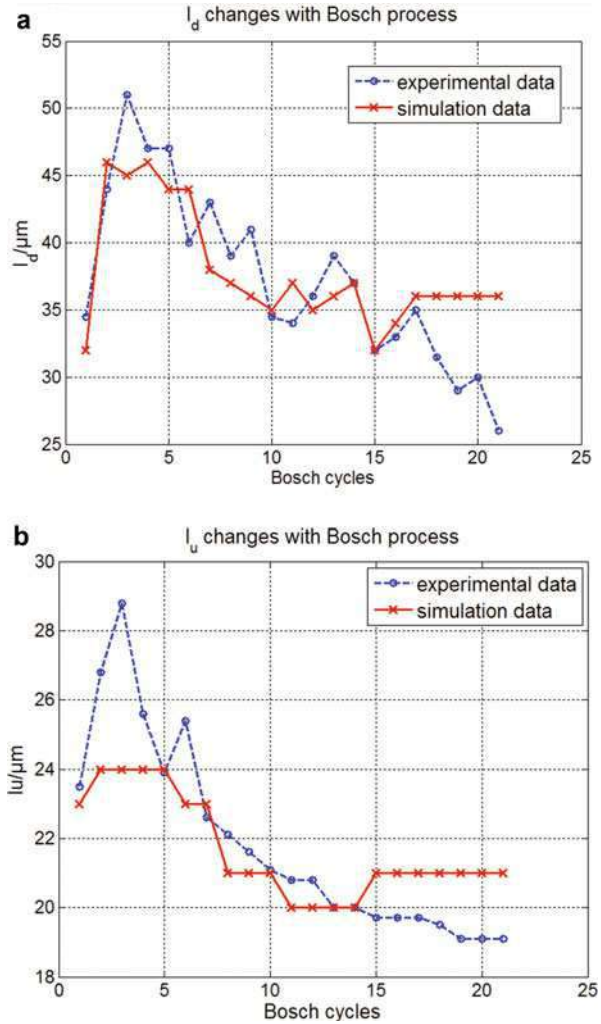
The fitness value in each generation is shown in Fig. 21. It shows that the program quickly approaches to the optimal solution, particles move less, and the simulation result gets converged and close to the experiment data.

Figure 22 compares  $ld$  and  $lu$  in simulation with the experiment within 20 gas chopping cycles of DRIE etching. The optimization results are consistent with the experimental data. The average error of global optimal solution is less than 34 nm. During the optimization process, errors of etching depth and undercutting values calculated by simulation program reduce significantly at the beginning, but after 20 steps errors of simulation stop decreasing. From Fig. 22, it is observed that errors of the undercutting size  $lu$  are relatively large at the beginning and the end of DRIE fabrication.

## The Modified Process Model

Since after tens of etching steps the reduction of the undercutting and etch depth are less than expectation, a modified process model is considered. One etching step includes isotropic chemical etching and anisotropic ion assisted etching. Since isotropic chemical etch rate is equal in all directions, so it contributes the etching on both sidewalls and bottoms. While the anisotropic ion assisted etching mainly removes the material on the bottom.

**Fig. 22** The curve of the  $l_d$  and  $l_u$  that changes with the loops



In current simulation model, the reduction of  $l_d$  value mainly results from effect of the ion assisted etching and the size of  $l_u$  relates to the chemical etching in each cycle. As the etching depth increases, ion assisted etch rate decreases and the calculation results of  $l_d$  and  $l_u$  in simulation begin to show larger errors. In this research, a modified model is proposed for better accuracy. A parameter  $\beta$  is introduced to the present model, which relates to the aspect ratio of the etch profile.  $\beta$  decreases as the aspect ratio of microstructure increases, which bring the reduction effect of the etch rate of chemical etching. The relationship of the concentration of etchant at the trench bottom with the concentration of etchant on the mask opening can be expressed by Knudsen diffusion function.

$$\frac{Tb}{Tt} = \frac{1}{\cosh\phi_t + \frac{\phi_t w}{2L_d} \sinh\phi_t} \quad (12)$$

where  $\phi_t = \frac{L_d}{w} \sqrt{\frac{3}{2}} \eta$ .  $Tb$  is the concentration of the etchant at the bottom of etched profile, and  $Tt$  is the concentration of the etchant on the top of opening position.  $W$  is the width of the opening window,  $ld$  is the depth of the trench, and  $\eta$  is the coefficient of the transmission. When the opening size on substrate surface is constant, the higher the viscosity coefficient, the lower the concentration at the bottom of the trench. The smaller viscosity coefficient will lead to more slowly decrease in the concentration of etchant at the bottom.

The concentration of etchant reduces depends on two parameters. One is the aspect ratio  $\frac{L_d}{w}$ , another is the coefficient  $\eta$ . The definition of the formula for  $\beta$  is as following equation:

$$\beta = \min\left(1, \left(\frac{w}{L_d}\right)^\eta\right) \quad (13)$$

$\beta$  is the in between  $\left(\frac{w}{L_d}\right)^\eta$  and 1, because the maximum value of  $\beta$  is 1. When the etching depth is more than the width of the opening on substrate, this parameter will start to work. As a result the etch rate of chemical etching is modified as

$$E_{er} = k_0 F^r \exp\left[-\frac{E_e}{k_B T_s}\right] \cdot R_s \cdot \beta \quad (14)$$

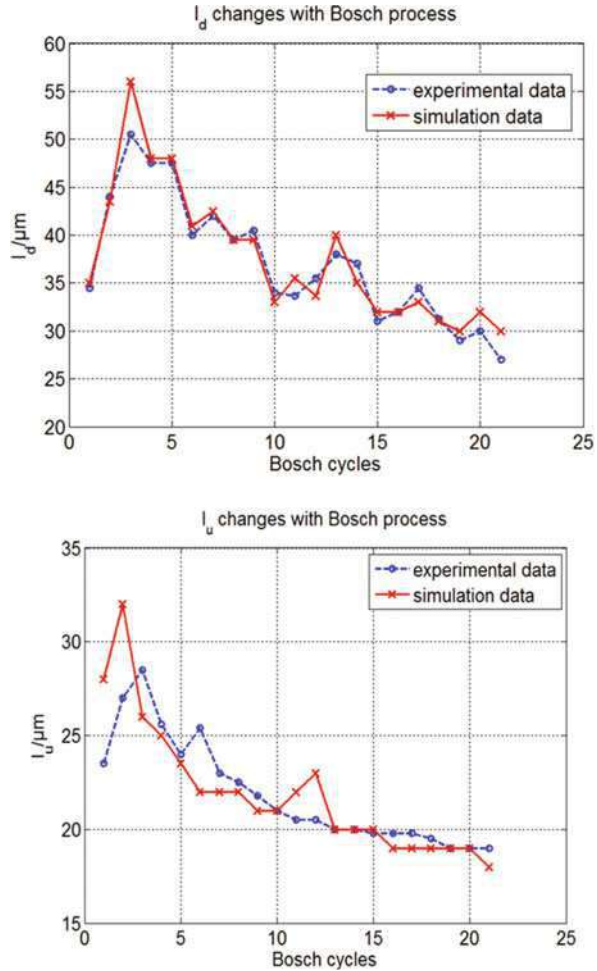
By using the modified etch model, the simulation result of PSO-CCA is compared with the experimental results in Fig. 23. The final global optimal solutions from DRIE model is less than 21 nm. And comparing with the experiment, the average error of undercutting for each etching cycle is about 20 nm. The relative error is less than 0.105%. Obviously this modified etching model improved the accuracy of DRIE model significantly. The reduction trend of  $ld$  and  $lu$  value during DRIE is consistent with the experimental data.

---

## The Parallel Computation for DRIE Simulation

DRIE is a technology that is now getting more and more useful in MEMS. There has been commercial software that includes the module of DRIE technology. However, the methods with complex physical simulation model run slowly that cannot satisfy the needs of users. As a result, one has to drop out physical model and get more simple geometry model instead. For the simulation of DRIE model in this chapter, the PSO uses tens of particles. Usually the CCA calculation converges after tens of generations in the optimization process. The time used for optimization is hundred times longer than the DRIE process, which means that we cannot use much time in

**Fig. 23** Comparison between improved model and experimental data

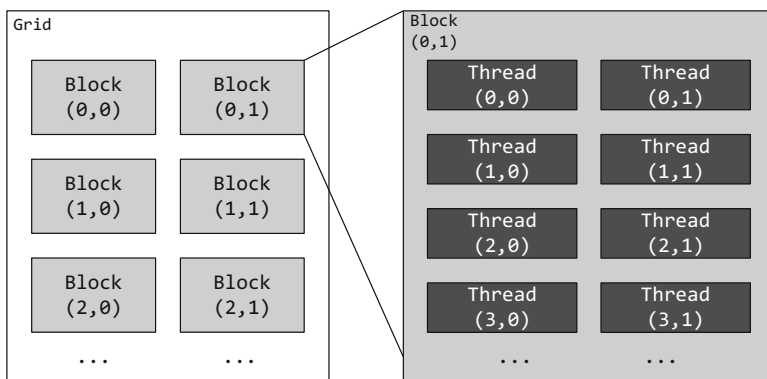
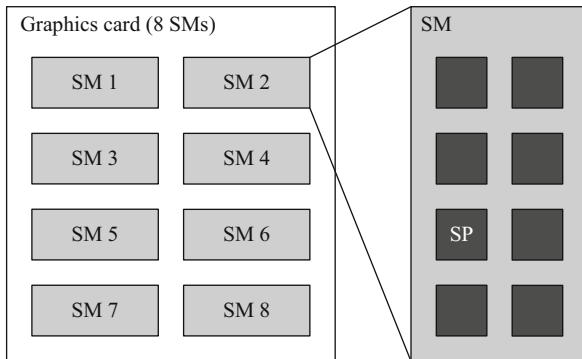


DRIE process. This chapter is going to introduce how to use Nvidia CUDA device in detail. The full name of CUDA is Compute Device Architecture. It is launched by Nvidia. We can use parallel computing engine of Nvidia graphics card to solve complex DRIE computing tasks. It is more efficient than the CPU implementation.

CUDA architecture in general is based on an extensible array of multithreaded streaming multiprocessors; every streaming multiprocessor is called SM. In each SM, there are many stream processors (SP). Taking Nvidia Geforce 9600GT for example, as shown in Fig. 24, there are eight SM in graphics, so the total SP is  $8 \times 8 = 64$ .

In CUDA programming, all SPs in all SMs can perform calculating tasks at the same time. In order to manage multithreads, CUDA uses the execution mode called SIMT. SIMT refers that multiple threads execute instructions at the same time as a single one. Different threads through their own ID access different storage addresses.

**Fig. 24** The CUDA architecture of Geforce 9600GT



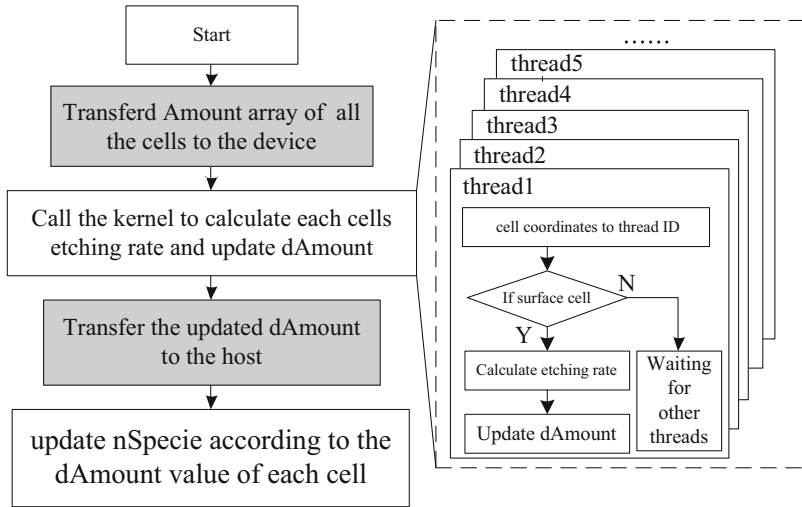
**Fig. 25** Organization schematic of two-dimensional thread

In CUDA program, the core part of parallel computing is defined as kernel function. When a kernel function is run, it will be executed by N threads each time. Threads can be organized into one-dimensional, two-dimensional, or three-dimensional blocks, and multiple thread blocks form a complete thread grid. Figure 25 shows a set of two-dimensional thread grid.

Usually, CUDA program includes both the serial and parallel codes, with serial code executing on the CPU and parallel code executing on the GPU. This programming model is known as heterogeneous programming. CUDA program refers CPU as host and GPU as device. Host and device manage their own storage space – host memory and device memory. So whether a CUDA program can perform calculating tasks efficiently depends on not only the distribution of threads but also the control of data transmitting between host and device.

The difference between GPU calculation and CPU calculation is that GPU can launch hundreds of threads at the same time to do the calculation, while the single core CPU can only launch one thread at one time to handle the task. Thus the parallelization of PSO-CCA for DRIE process is to change the serial code using the





**Fig. 26** Simple parallelization of the simulated etching serial algorithm

index loop to the parallel calculation by GPU’s thread index. Because when the program runs, the host and device can only access to their own register, host register and device register should change their date before and after the kernel code runs. For the typical serial solution of one step etching in DRIE process cellular automaton program traverses all the cells twice. One loop is for etch rate calculation and changing the status of the cells, the other loop is for updating and propagation of the etch front.

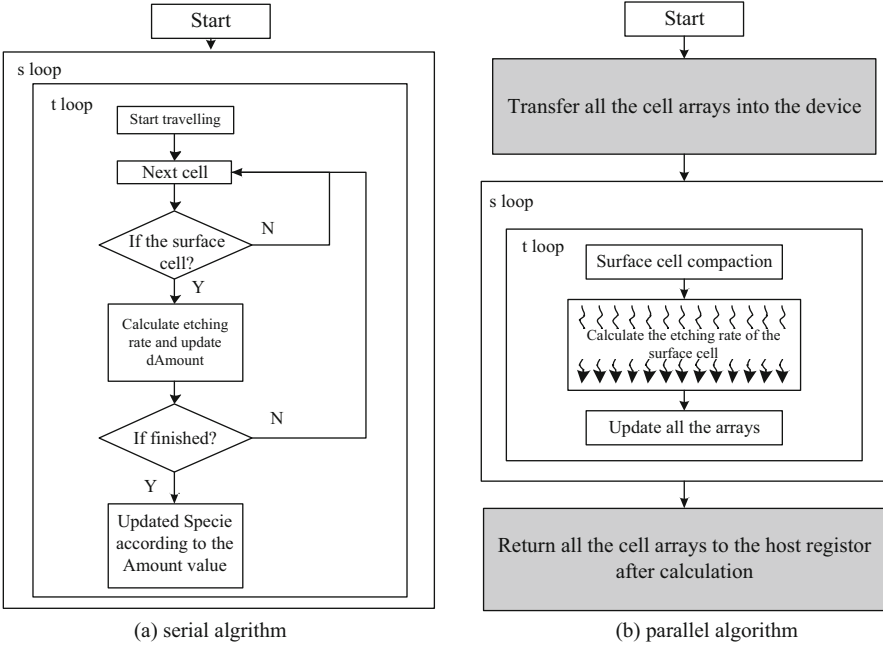
Figure 26 describes the implementation of parallel etching process. Firstly, the program transfers the whole cell array to the device register. After it calls the kernel code, the kernel will launch multiple threads. Each thread has its own ID and deals with only one cell. Thread ID is mapped into the geometric coordinate of the cell. The mapping rule is as the following:

$$z = \text{threadId}/(\text{Dimx}*\text{Dimy});$$

$$x = (\text{threadId} - z*\text{Dimx}*\text{Dimy})/\text{Dimy};$$

$$y = \text{threadId} - z*\text{Dimx}*\text{Dimy} - x*\text{Dimy};$$

In each thread, the program firstly decides whether the cell is on surface or inside the bulk, then it calculates the etch rate of surface cell according to the normal vector of the etched profile, finally the program updates the cell status and changes the etched profile. Finally, the worker thread waits for other threads to complete the calculation and submits the data to host register. This parallelization implementation is an easy solution for PSO-CCA program and it gets good acceleration in time-consuming DRIE simulation model.

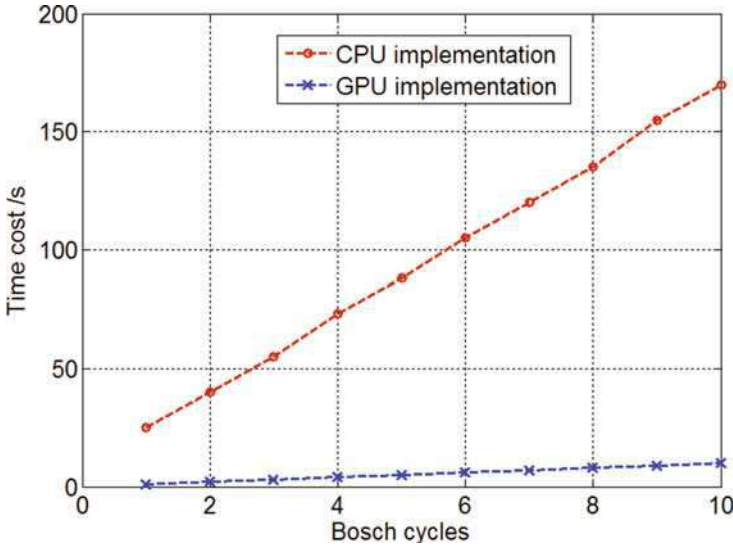


**Fig. 27** The comparison between the serial algorithm and the parallel algorithm of DRIE simulation

In this study, both serial and parallel algorithms for PSO-CCA program are performed by CPU and GPU, respectively. The frameworks of the two implementations are shown in Fig. 27. The parallel algorithm has accelerated the most time-consuming part of the DRIE process model. The parallel algorithm uses the serial code to transmit and exchange the data only at the beginning and the very end of the program. The comparison is run on the desktop computer with an Intel Dual-Core 2.5 GHz CPU and Nvidia GeForce GTX 560 Ti GPU. This GPU contains 1GB memory and 384 CUDA core SP.

Since each particle in PSO optimization process calls the DRIE simulation program for the calculation of fitness function, the parallel algorithm helps simulation model quickly converge to the optimal parameter setting.

The DRIE simulation in the following figures runs by a square mask with an opening size of 10,000 nm\*10,000 nm in three dimension; the program terminates at the etching depth of 30,000 nm. The resolution of cellular automaton cell system is  $10^7$  (there are  $10^7$  cells totally). The calculation time of each etching process for the comparison of serial and parallel algorithm is shown in Fig. 28. As the etch depth after several etch and deposition cycles increases, computation time cost in each DRIE process keeps almost constant for the parallel algorithm. The computation time increases linearly with the depth of the microstructures by the serial implementation of simulation.



**Fig. 28** Time cost in each DRIE process

In order to simulate complex and large openings of mask patterns, the resolution of PSO-CCA should be high enough to show the details of microstructures. Figure 29 demonstrates the simulation time consumed for different scales of cell system. Our parallel implementation of the CCA method accelerates the computation significantly, reducing the time to perform a complex simulation to several tens of seconds.

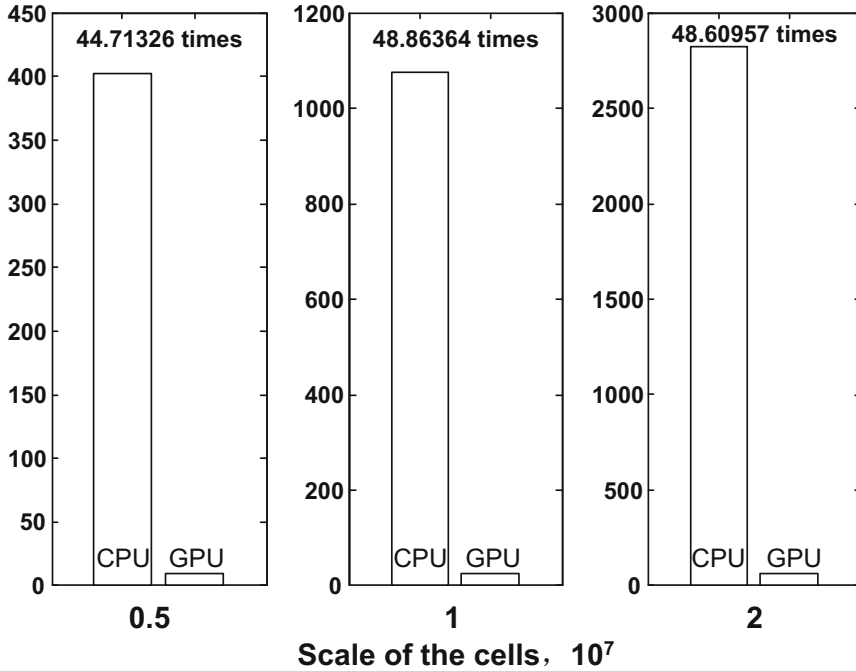
## Simulation of DRIE Applications

The influence on the etch profile by the standard deviation of the ion source distribution function is described in Fig. 30. From the simulation result, the larger standard deviation of the ion source distribution function leads to less etching at the bottom. The undercutting of sidewall gets the reverse result. The etching profile becomes isotropic and smooth as standard deviation increases. In this section, different  $\sigma$  (0.02, 0.2, 0.4, 0.8) were selected.

The etching rate in DRIE model has isotropic chemical etching and the anisotropic assisted etching rate parts. The etching profiles by three different ratio of two types etch style in one step are shown in Fig. 31.

The ratio of isotropic and anisotropic etching part is 1:5, 1:1, and 5:1 for a trench structure. With the increasing of the anisotropic etching part, the etch depth increases (Fig. 32).

DRIE fabrication process is able to fabricate high aspect ratio microstructures. The typical method to test lag effect is to etch different sizes of mask opening on

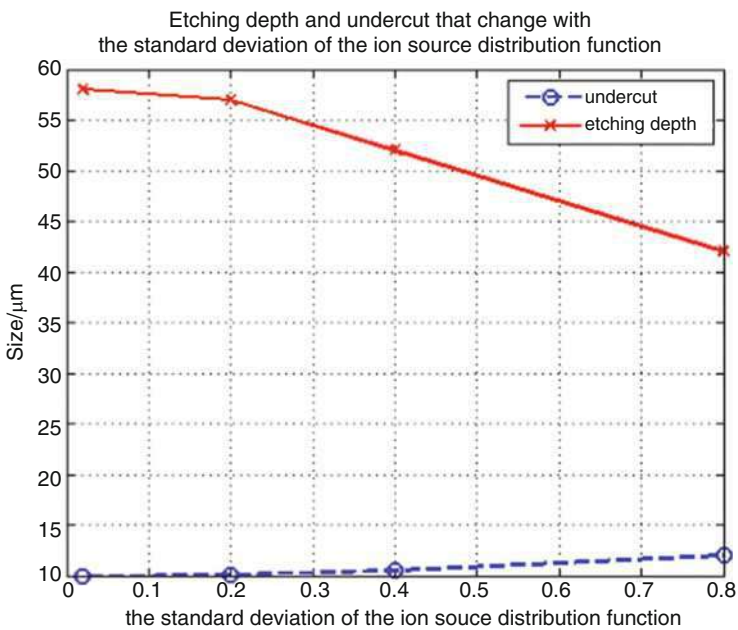
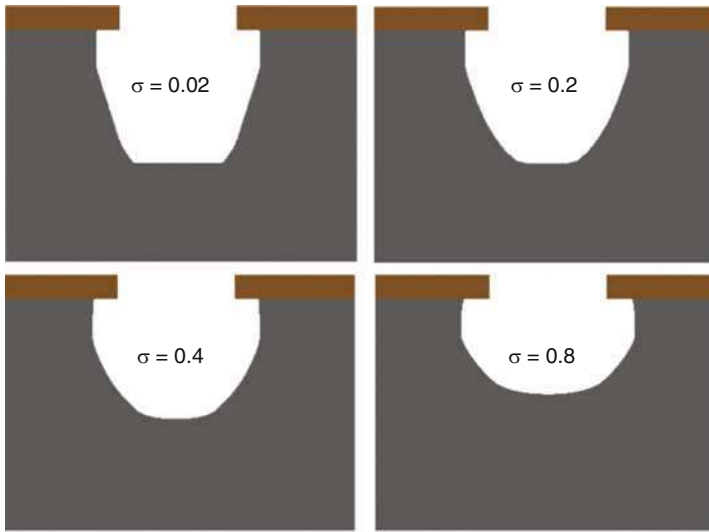


**Fig. 29** Time used in different scales of cell system

the wafer and measure the etch depths after DRIE etching process. The experimental conditions are Si wafer, with reactant  $\text{SF}_6$  and  $\text{C}_4\text{F}_8$ . The flow rates of reactant are 180 and 130 sccm, respectively. The cycle time interval for  $\text{SF}_6$  and  $\text{C}_4\text{F}_8$  are 12 and 5 s. The pressure is 15 mTorr. Total process time for aspect-ratio-dependent etching is 69 min. The etched profile shows lag effect and scalloping sidewalls. The lag effect is also considered during the PSO search. Figure 33 shows that the simulated trench depth agrees well with the experiment for different opening sizes.

The linear fit curve for aspect-ratio-dependent etching is shown in Fig. 33. The etching depth of different mask opening size follows the form of exponential curve. The logarithmic scales for etching depth on both the horizontal and vertical axes are shown in Fig. 33. The fitted line has slope 0.14345. The simulation of PSO-CCA is carried out using the same size mask pattern in the experimental work. The slope of the double logarithmic curve was 0.14459, which is close to the experiment. The simulation result provides strong evidence that both the passivation and etching steps of the DRIE process are well simulated (Fig. 34).

Figure 35 is the cross section of the DRIE process simulation. Scalloping structure caused by etch cycle is identical to the experiment. Scalloping sidewall is an important issue in DRIE. Some applications of MEMS, such as optical



**Fig. 30** The influence of the standard deviation of the ion source distribution function on etching profile

MEMS, require the size of scalloping small enough to obtain the smooth sidewall.

Usually the number of gas chopping cycles in a DRIE process is the number of scalloping shape. The width and height of each scalloping shape in the SEM image

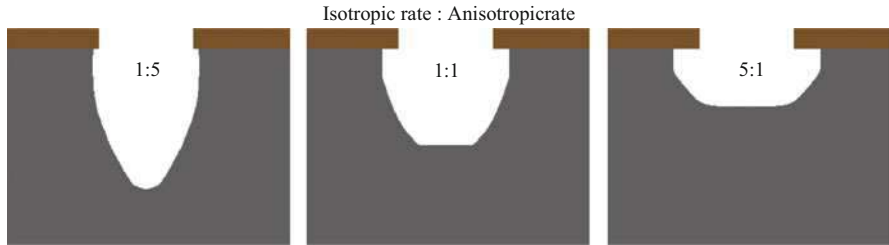


Fig. 31 Profile simulated by different etching rate

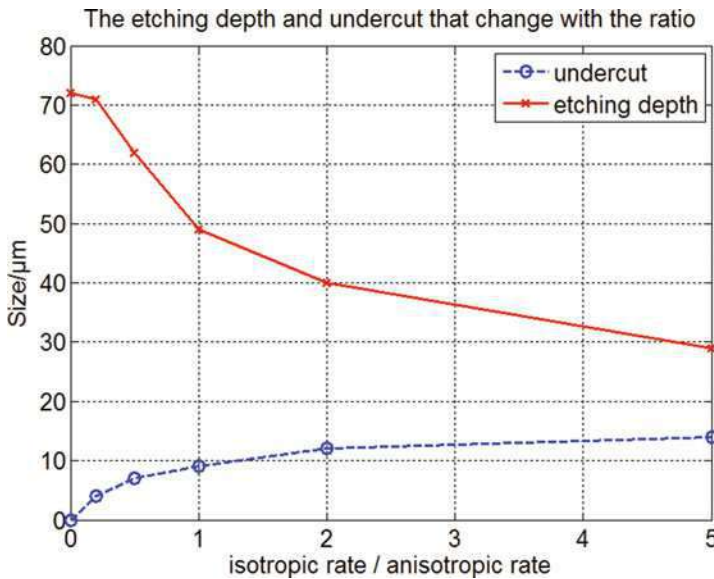
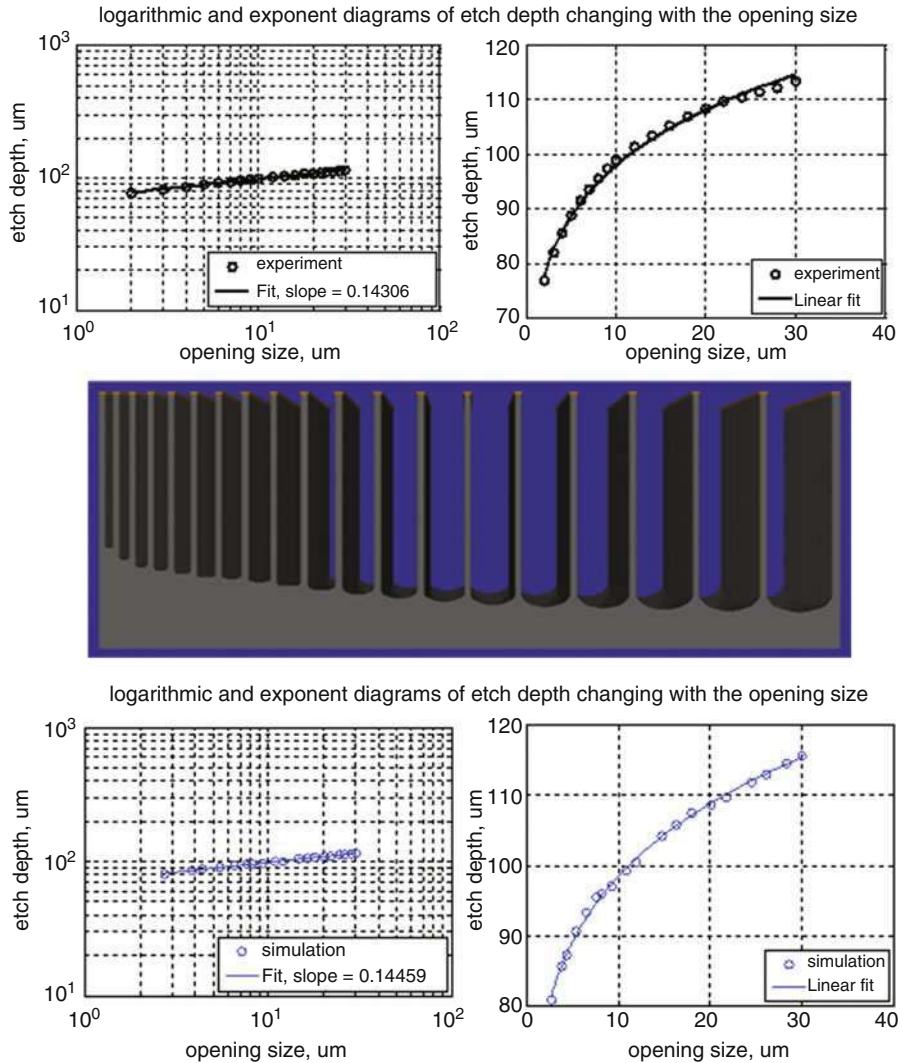


Fig. 32 The etching depth and undercut that change with the ratio

were measured. A total of 23 cycles were performed in the experiment, besides the first and last scalloping shape, there are totally 21 sets of valid data for depth per cycle  $l_d$  and undercut length per cycle  $l_u$ .

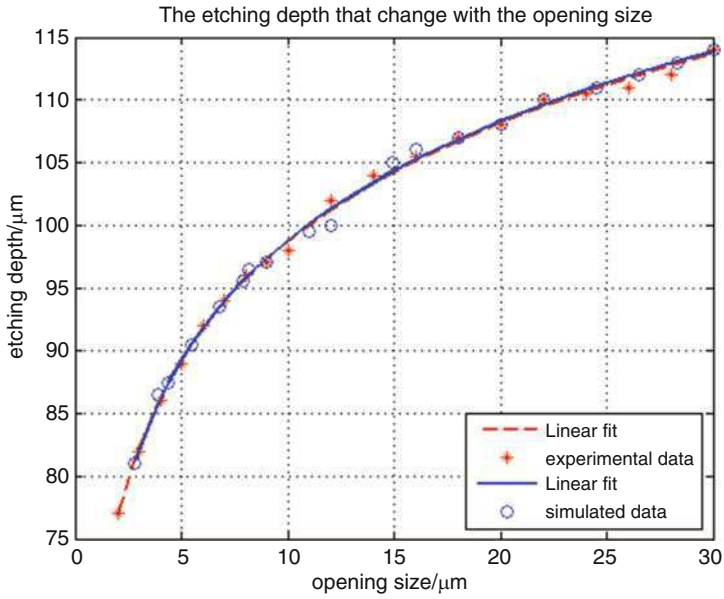
Due to the variability of  $l_d$  and  $l_u$ , we also used double logarithmic curve for the comparison of the experimental and simulation data with the change of etching depth. As shown in Fig. 36 the slope of the double logarithmic curve by simulation fits well with the experiment.

3D simulations with complex mask patterns are also fabricated and simulated in the program. Since the time cost for the 3D simulations is excessively long, the resolution (number of cells per dimension, which determines the physical size of each cell) is determined by vector  $d$ . Larger  $d$  will provide more details at the cost of



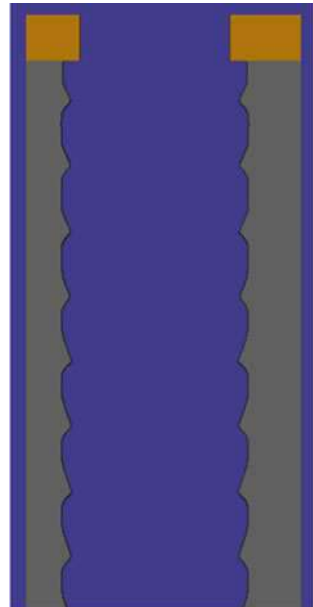
**Fig. 33** The comparison of the lag effect in simulation and experimental data

more time-consuming simulations. The etching depth of the narrower opening part is smaller than that of the wider part, while the overlapping portions are influenced by a mixture of both effects. Figure 37 presents a collection of experimental and simulated results for different masks. Such simulations can help to design the masks and predict the fabrication results. The presented 3D simulations were performed on a standard PC with a 3.10 GHz CPU and CUDA acceleration on an Nvidia GeForce 560GTX GPU.

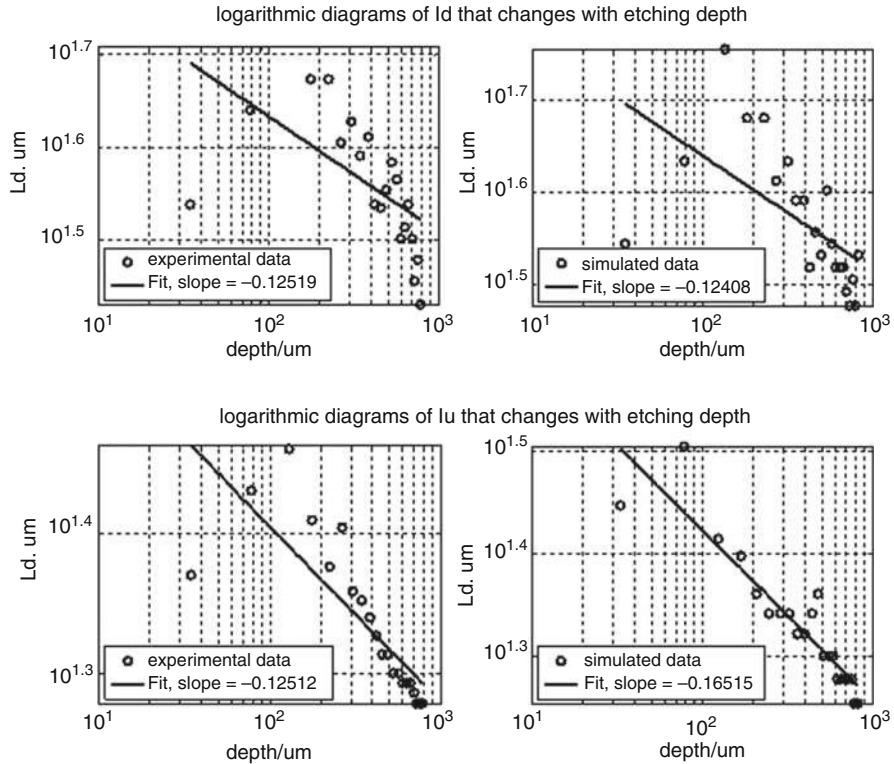


**Fig. 34** The comparison between the simulation and experiment in the etching depth that changes with the opening size

**Fig. 35** Simulation of trenches fabricated by DRIE process





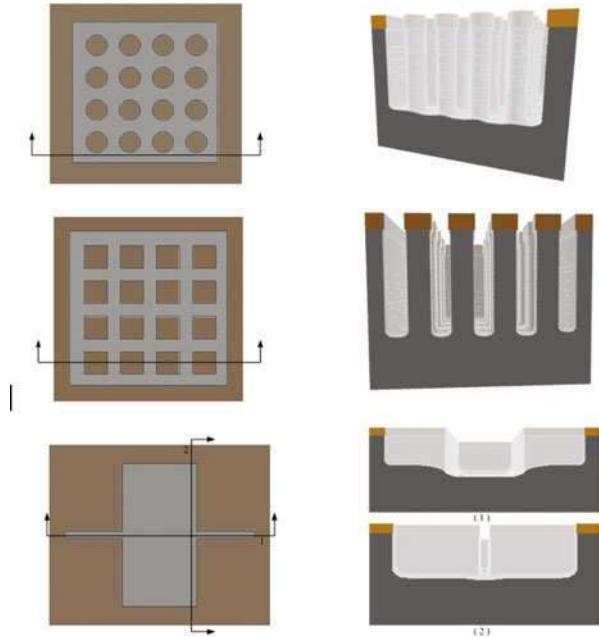


**Fig. 36** Comparison of the simulation result of  $l_d$  and  $l_u$  with the experiment

## Summary

As a widespread form of deep reactive ion etching (DRIE), the process alternates etching and passivation cycles, typically leading to characteristic scalloping patterns on the sidewalls. Measurements of the etch depth per cycle  $l_d$  and undercut length per cycle  $l_u$  show a strong dependence of the undercut ratio  $l_u/l_d$  on the trench aspect ratio for a wide range of opening sizes. Although various simulation models have been proposed, the determination of the corresponding parameters from experimental data remains unsolved. We present the use of (i) the continuous cellular automaton (CCA), to simulate the process reliably in three dimensions; (ii) the particle swarm optimization (PSO) method, to determine suitable values for the atomistic CCA parameters directly from experimental data; and (iii) a GPU, parallel implementation of the CCA, to increase the computational efficiency of the simulations. The resultant parameter-optimized CCA simulations show good agreement with the experiments. The approach has a large potential for the simulation of other MEMS processes.

**Fig. 37** The simulated results for complex mask patterns



## References

- Alam S, Dobbie G, Koh YS, Riddle P, Ur Rehman S (2014) Research on particle swarm optimization based clustering: a systematic review of literature and techniques. *Swarm Evol Comput* 17:1–13
- Chienliu C, Yeong-Feng W, Yoshiaki K, Ji-Jheng S, Yusuke K, Chih-Kung L, Kuang-Chong W, Masayoshi E (2005) Etching submicrometer trenches by using the Bosch process and its application to the fabrication of antireflection structures. *J Micromech Microeng* 15:580
- Ertl O, Selberherr S (2010) Three-dimensional level set based Bosch process simulations using ray tracing for flux calculation. *Microelectron Eng* 87:20–29
- Ferrando N, Gosálvez MA (2012) Evolutionary continuous cellular automaton for the simulation of wet etching of quartz. *J Micromech Microeng* 22:025021
- Ferrando N, Gosálvez MA, Cerdá J, Gadea R, Sato K (2011) Octree-based, GPU implementation of a continuous cellular automaton for the simulation of complex, evolving surfaces. *Comput Phys Commun* 182:628–640
- Hernández Encinas A, Hernández Encinas L, Hoya White S, Martín del Rey A, Rodríguez Sánchez G (2007) Simulation of forest fire fronts using cellular automata. *Adv Eng Softw* 38:372–378
- Hong L-S, Shimogaki Y, Komjyama H (2000) Macro/microcavity method and its application in modeling chemical vapor deposition reaction systems. *Thin Solid Films* 365:176–188
- Kennedy J, Eberhart R (1995) Particle swarm optimization. In: *Proceedings of IEEE international conference on neural networks, Perth, vol 4*, pp 1942–1948
- Knizikevičius R (2003) Simulation of inverse reactive ion etching lag. *Vacuum* 72:53–57
- Lai S, Johnson D, Westerman R (2006) Aspect ratio dependent etching lag reduction in deep silicon etch processes. *J Vac Sci Technol A* 24:1283–1288

- Li Y, Gosálvez MA, Pal P, Sato K, Xing Y (2015) Particle swarm optimization-based continuous cellular automaton for the simulation of deep reactive ion etching. *J Micromech Microeng* 25:450–456
- Mita Y, Sugiyama M, Kubota M, Marty F, Bourouina T, Shibata T (2006) Aspect ratio dependent scalloping attenuation in drie and an application to low-loss fiber-optical switches. In: Proceedings of IEEE international conference on micro electro mechanical systems, Istanbul, pp 114–117
- Nagel K, Schreckenberg M (1992) A cellular automaton model for freeway traffic. *J Phys I* 2:2221–2229
- Pilliod JE Jr, Puckett EG (2004) Second-order accurate volume-of-fluid algorithms for tracking material interfaces. *J Comput Phys* 199:465–502
- Poli R, Kennedy J, Blackwell T (2007) Particle swarm optimization. *Swarm Intell* 1:33–57
- Ranganathan N, Lee DY, Youhe L, Lo G-Q, Prasad K, Pey KL (2011) Influence of Bosch etch process on electrical isolation of TSV structures. *IEEE Trans Components Packag Manufact Technol* 1:1497–1507
- Rangelow IW (2003) Critical tasks in high aspect ratio silicon dry etching for micro-electromechanical systems. *J Vac Sci Technol A Vac Surf* 21:1550–1562
- Saraf I, Goeckner M, Goodlin B, Kirmse K, Overzet L (2011) Mask undercut in deep silicon etch. *Appl Phys Lett* 98:161502
- Shi Y, Eberhart R (1998) A modified particle swarm optimizer. In: Proceedings of IEEE international conference on evolutionary computation, Anchorage, pp 69–73
- Shi Y, Liu H, Gao L, Zhang G (2011) Cellular particle swarm optimization inform. *Science* 181:4460–4493
- Tan Y, Zhou R, Zhang H, Li Z (2006) Modeling and simulation of the lag effect in a deep reactive ion etching process. *J Micromech Microeng* 16(12):2570–2575
- Tsujimoto K, Tachi S, Ninomiya K, Suzuki K, Okudaira S, Nishimatsu S (1986) A new side wall protection technique in microwave plasma etching using a chopping method. International conference on solid state devices & materials, Tokyo, pp 229–232
- Wang Y, Guo Y, Zhang H (2007) Modeling and simulation of footing effect in DRIE process. IEEE conference on nanotechnology, pp 1135–1138
- Weymouth G, Yue DK-P (2010) Conservative volume-of-fluid method for free-surface simulations on Cartesian-grids. *J Comput Phys* 229:2853–2865
- Xing Y, Gosálvez MA (2007) Step flow-based cellular automaton for the simulation of anisotropic etching of complex MEMS structures. *New J Phys* 9:436
- Yuan L, Yan X, Gosálvez MA, Pal P, Zhou Y (2013) Particle swarm optimization of model parameters: simulation of deep reactive ion etching by the continuous cellular automaton. In: Proceedings of IEEE international conference on transducers & eurosensors, Barcelona, pp 1087–1090
- Zhibo M, Chengyu J, Weizheng Y (2012) A triple-layer protection process for high-aspect-ratio silicon micromachining by DRIE of SOI substrates. *J Micromech Microeng* 22:055028



# Modeling and Simulation of SU-8 Thick Photoresist Lithography

Zai-Fa Zhou and Qing-An Huang

## Contents

Overview .....	68
UV Lithography of SU-8 Photoresist .....	69
Simulation Models .....	72
Aerial Image Simulation Models .....	73
Exposure Simulation Models .....	76
PEB Simulation Models .....	78
Development Simulation Models .....	79
Algorithms for Etching Surface Advancement Simulation .....	79
Fast-Marching Method .....	81
Hash Fast-Marching Method .....	83
Tests and Discussions .....	87
Simulations and Discussions .....	90
Conclusions .....	92
References .....	92

## Abstract

SU-8 photoresist can be used to produce high aspect ratio and three-dimensional (3D) lithographic patterning based on standard contact lithography equipment due to its excellent coating, planarization, and processing properties and thus has become the favorite photoresist material for the fabrication of various micro-electromechanical system (MEMS) structures and devices. However, as feature sizes get smaller and pattern complexity increases, particular difficulties arise and need to be carefully considered. The accuracy and precision, with which a feature on a mask can be reproduced throughout a thick resist structure, will depend on key parameters in the setup, the material properties of the SU-8 resist, and the

---

Z.-F. Zhou (✉) · Q.-A. Huang  
Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, Jiangsu, China  
e-mail: [zfzhou@seu.edu.cn](mailto:zfzhou@seu.edu.cn); [hqa@seu.edu.cn](mailto:hqa@seu.edu.cn)

thickness of the resist structure. Modeling and simulation studies may help improve our understanding and process design of the SU-8 lithography, thereby allowing rapid product and process development. In this chapter, the basic process and mechanism of UV lithography of the SU-8 are introduced briefly. Various models for the lithography, including the aerial image model, exposure model, post-exposure bake model, and development model, are presented and discussed. Main algorithms for the etching surface advancement simulation, including the string, ray-tracing, cellular automaton, and fast-marching algorithms, are then compared and analyzed. Simulations of the UV lithography of the SU-8 are presented, and a series of experiments have been performed for SU-8 2000 series photoresists under UV source with 365 nm (2.6 mW/cm<sup>2</sup>) radiation. The simulation results demonstrate to be in agreement with the experimental results. This is useful to optimize the inclined UV lithography processes of SU-8 photoresists and to accurately design and control the dimensions of some MEMS microstructures.

---

**Keywords**

SU-8 photoresist · Cellular automaton · Lithography simulation · Modeling · Fast marching

---

**Overview**

SU-8 photoresist has excellent physical, mechanical, and chemical corrosion resistance capacity and thus has become an ideal photoresist material for the fabrication of various MEMS structures and devices (Han et al. 2004; Hung et al. 2004; Balslev and Romanato 2005; Becnel et al. 2005; Yoon et al. 2006; Sato et al. 2006; Campo and Greiner 2007; Hung and Liang 2008; Cheng et al. 2009; Colin et al. 2010; Moser et al. 2011; Rouabah et al. 2011), such as optical pickup head (Hung et al. 2004), microlens (Cheng et al. 2009), microfluidic channels (Colin et al. 2010), micropump (Rouabah et al. 2011), etc. The SU-8 can be efficiently used to produce microstructures with aspect ratio larger than 20 based on standard contact lithography equipment, for its low optical absorption in the ultraviolet (UV) range, especially the UV light with 365 nm. Although X-ray LIGA (Lithografie, Galvanoformung, Abformung) (Becker et al. 1996) using SU-8 is well suited for fabricating very high aspect ratio structures, very precise shape definition, high straightness, and planarity of sidewalls (Balslev and Romanato 2005; Becnel et al. 2005), this approach has not made its way into large-scale industrial applications, because of the expensive synchrotron source needed to expose the photoresist layer and the rather small patterned area. On the contrary, UV lithography utilizes an inexpensive UV light source to expose the SU-8. Because heating and transmittance are not an issue in optical masks, a simple chromium mask can be substituted for the technically sophisticated X-ray mask. These reductions in complexity make the UV lithography of SU-8 much cheaper and more accessible than its X-ray counterpart.

Various UV lithography processes of SU-8, including vertical/inclined (tilted) lithography (Han et al. 2004; Hung et al. 2004; Sato et al. 2006; Campo and Greiner 2007;

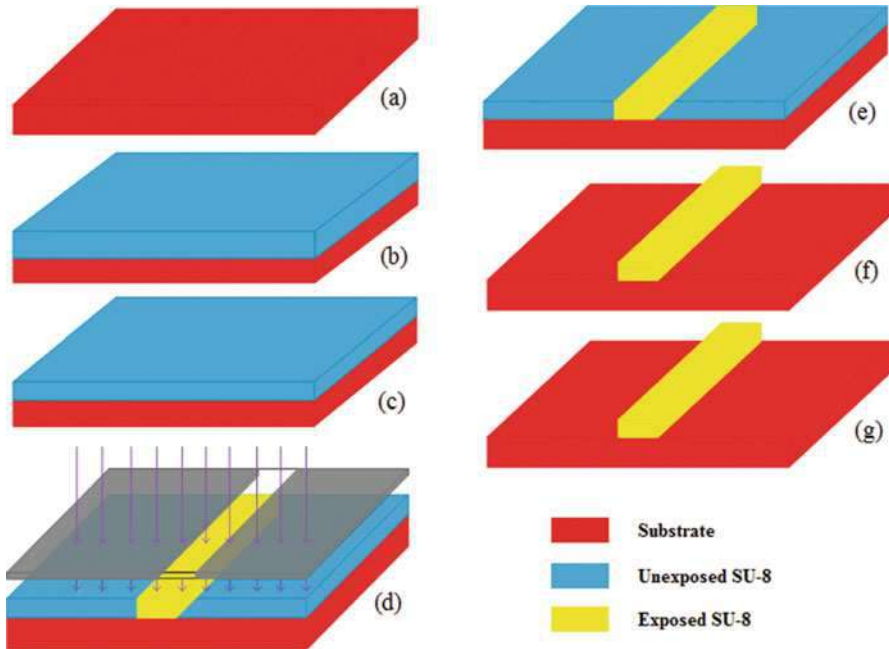
Hung and Liang 2008), multidirectional lithography (Yoon et al. 2006), interference lithography (Cheng et al. 2009), etc., have been widely used for fabricating MEMS structures and devices for over 15 years. Lots of lithography parameters, such as gap between the mask and SU-8 layer, exposure time, light intensity, SU-8 thickness, mask shape and development time, etc., affect the final lithography profiles. Thus accurate control of the dimensions of microstructures and optimization of the lithography processes become complicated and time-consuming by using conventional repeated-experiment method. Simulation is widely acknowledged to be ideally suited for the optimization of the lithography process (Dill et al. 1975; Mack 1985; Henke et al. 1990; Neureuther and Mack 1997; Cole et al. 2001; Karafyllidis et al. 2000; Zhu et al. 2004), since it is efficient to understand the fundamental effects of individual process parameters. For many years, lithography simulations have been employed to optimize the design of micron and submicron devices and related devices in integrated circuit (IC) fields, and simulation software for thin photoresist lithography processes in the IC field has been commercialized. However, the lithography process simulation of thick photoresists such as SU-8 cannot be accurately implemented by the software specified for thin photoresist, since the SU-8 has properties different from the thin photoresist (Tang et al. 2007). To solve this problem, some specified researches have been implemented during the past several years for the UV lithography simulations of thick photoresists (Tang et al. 2005, 2007; Cheng et al. 1999; Chuang et al. 2002; Tian et al. 2005; Sensu et al. 2005; Yang and Wang 2005; Rumpf and Johnson 2005; Kang et al. 2006; Feng et al. 2006; Zhou et al. 2007a, b, 2010, 2011, 2014; Zhu et al. 2008, 2009; Shi et al. 2011; Miao 2013; Yang et al. 2013; Huang and Hsu 2014).

As the complexity of MEMS devices increases, the UV lithography of the SU-8 is increasingly utilized to produce microstructures and devices. Simulations play an important role in timely market entry of some MEMS products. In this chapter, the basic process and mechanism of UV lithography of the SU-8 are introduced briefly. Various models for the lithography, including the aerial image model, exposure model, postexposure bake model, and development model, are presented and discussed. Main algorithms for the etching surface advancement simulation, including the string, ray-tracing, cellular automaton, and fast-marching algorithms, are then compared and analyzed. Simulations of the UV lithography of the SU-8 are presented, and a series of experiments have been performed for SU-8 2000 series photoresists under UV source with 365 nm (2.6 mW/cm<sup>2</sup>) radiation. The simulation results demonstrate to be in agreement with the experimental results. This is useful to optimize the inclined UV lithography processes of SU-8 photoresists and to accurately design and control the dimensions of some MEMS microstructures.

---

## UV Lithography of SU-8 Photoresist

SU-8 is a negative chemical amplification photoresist based on EPON SU-8 epoxy resin (from Shell Chemical) that has been originally developed, and patented (US Patent No. 4882245 (1989)) by IBM (Shaw et al. 1997). The SU-8 is a very viscous polymer that can be spun or spread over a thickness ranging from below 1  $\mu\text{m}$  up to



**Fig. 1** Fabrication flowchart of SU-8 photoresist: (a) Substrate pretreatment, (b) Spin-coating of SU-8, (c) Soft bake, (d) Exposure, (e) Postexposure bake, (f) Development, and (g) Rinse and dry

above 1000  $\mu\text{m}$ . When exposed, the cross-linking of the long molecular chains of SU-8 causes the solidification of the material. The original SU-8 resist series (SU-8 5, SU-8 50, SU-8100, etc.) and its newer formulations (SU-8 2000 series and SU-8 3000 series) have been now commercialized by the company MicroChem Inc. (previously Microlithography Chemical Corp.) (<http://www.microchem.com>). Gersteltec Sarl (<http://www.gersteltec.ch>), another company, has now also bought a license from IBM to produce and sell the SU-8. Since the products from the two companies are similar, this paper will only discuss the SU-8 from MicroChem Inc. for the simplification of description.

Although the contents for various SU-8 are a little different, the fabrication flowchart for all SU-8 is basically the same, including the following main steps: substrate pretreatment, spin-coating of SU-8, soft bake, exposure, postexposure bake, development, and rinse and dry, as shown in Fig. 1.

For the substrate pretreatment, the silicon wafer (with or without antireflection thin films such as  $\text{TiO}_2$ , etc.) is subsequently cleaned with acetone, isopropyl alcohol (IPA), and deionized water. The wafer is then dehydrated on a hot plate or an oven. After that, the SU-8 is dispensed onto the wafer using a conventional photoresist spinner, with the film thickness controlled by the spin speed and the solid content of the epoxy solution (SU-8 types). After spin-coating, the wafer coated with SU-8 is soft baked in order to remove the solvent and to promote the

adhesion of the SU-8 layer to the substrate. The wafer is soft baked on a well-leveled hot plate. To reduce the stress in thick SU-8 layer, soft bake process with multistep ramping and stepping of temperature should be used. The soft bake temperature needs to be well controlled, because a lower temperature will leave residual solvent content, while a higher temperature may initiate thermal cross-linking (Ong et al. 2006). For thick SU-8 layer, the soft bake time is about 1 h per 100  $\mu\text{m}$ .

During the exposure process, the wafer coated with SU-8 can be directly exposed to UV light (365 or 405 nm for some special cases) on a standard UV aligner for conventional vertical lithography process. However, for various special lithography processes, such as inclined, multidirectional, and backside lithography, etc., the implementation of exposure process needs typical devices or equipment (Han et al. 2004; Hung et al. 2004; Zhou et al. 2011; Jiang et al. 2012; Hirai et al. 2007). The photoreaction initiator decomposes, and a strong acid ( $\text{HSbF}_6$ ) is generated within the SU-8 during exposure process. After exposure, the SU-8 layer should be directly baked, to induce the cross-linking reaction in the SU-8 layer. Since the cross-linking in the SU-8 layer is very slow at ambient temperature, heating process is adopted to increase the molecular motion and accelerate the cross-linking. Similar to the soft bake process, multiple steps of temperature changes are used to reduce the stress in PEB process. Usually, two steps of temperature changes are adopted (Zhu et al. 2009; <http://www.microchem.com>). After PEB, the wafer is immersed in a SU-8 developer (PGMEA, propylene glycol methyl ether acetate). Besides the cross-linking degree of the SU-8, the required development time depends on development conditions. Generally speaking, increasing the temperature and adding an agitation during development will accelerate the development rate. Finally, the wafer with development profiles should be rinsed briefly with IPA and then dried with a gentle stream of air or nitrogen. SU-8 has good mechanical properties; therefore, hard bakes are normally not required. For some special applications where the SU-8 is to be left as part of the final device, the SU-8 may be ramp/step hard baked between 150 °C and 200 °C on a hot plate or in a convection oven to further cross-link the material (<http://www.microchem.com>).

For above lithography process, many lithography parameters should be determined and optimized according to different design purposes. This makes it a little bit difficult to determine reasonable lithography parameters for various users with different background of knowledge. Fortunately, MicroChem Corp. provides process sheets for valuable reference concerning this issue (<http://www.microchem.com>). For example, Table 1 gives a practical process sheet for the UV lithography process of SU-8 2075 during experiments (Zhu et al. 2008), referring to the suggestions from MicroChem Corp. However, these process sheets seem to be not enough, with the increasing demand for high-accuracy and high-yield fabrication of microstructures based on SU-8. For micromachining community, various lithography methods and lithography conditions are expected to be investigated and optimized for the fabrication of different microstructures and devices. Simulation can offer a solid foundation to efficiently optimize this complicated process, making working in the fabrication facility more effective.



**Table 1** Process parameters for the UV lithography process of SU-8 2075 layer with 140- $\mu\text{m}$  thickness

Step	Process and parameters
Pretreat	Clean and 180 °C bake for 4 h
Spin coat	0–1700 rpm for 25 s (speedup) and 1700 rpm for 30 s (speed constant)
Soft bake	65 °C for 12 min at the first step bake and 95 °C for 48 min the second step bake
Expose	2.6 mW/cm <sup>2</sup> for 240 s
PEB	65 °C for 12 min at the first step PEB and 95 °C for 35 min at the second step PEB
Develop	Megasonic bath for 12 min at 20 °C
Rinse and dry	Isopropyl alcohol at 1 min and spin-dry

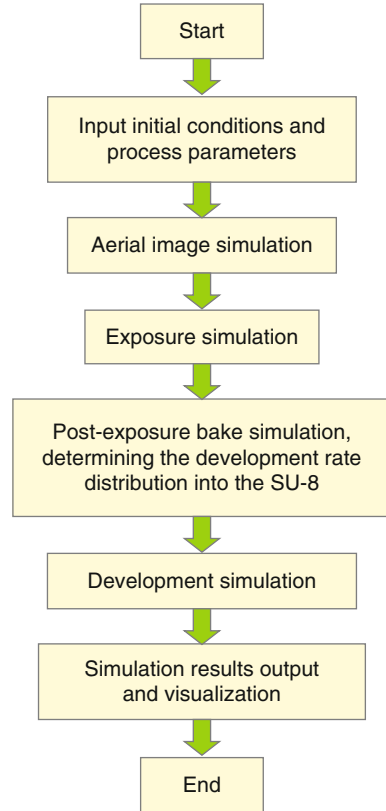
## Simulation Models

Currently, simulating the UV lithography of the SU-8 has been studied by constructing separate models for exposure, PEB, and development processes, similar to other thin negative chemical amplification photoresists in the IC fabrication field (Karafyllidis 1997; Fuhner et al. 2007; Karafyllidis et al. 1999). For the evaluation of effects from some lithography steps such as substrate pretreatment, spin-coating, soft bake, and rinse and dry, experimental characterizations are usually adopted, since the models for these lithography steps have not been implemented efficiently.

Although various lithography methods have been employed for the UV lithography of the SU-8, in a typical comprehensive lithography simulation, the aerial image simulation, exposure simulation, PEB simulation, and development simulation will determine the final development profiles of the SU-8. Figure 2 shows the simulation flowchart for the UV lithography of the SU-8. Generally speaking, the aerial image simulation and development simulation are the most time-consuming steps.

At the beginning of the simulation, the initial conditions and process parameters, including mask data (CIF or GDSII format), exposure time, UV light intensity, SU-8 layer thickness, UV light incident angle, gap between mask and SU-8 layer, reflectivity of substrate, light wavelength, PEB time and temperature for both two steps, development time and temperature, etc., should be input into the simulation system. The aerial image simulation is then to calculate the illumination of the mask by the incident UV light source, resulting in an aerial image on top of the SU-8. Following that, the exposure simulation calculates the UV light propagation as well as the generation of the photoacid  $\text{HSbF}_6$  within the SU-8 according to the light intensity distribution. After the exposure simulation, the PEB simulation is to describe the photoacid-catalyzed cross-linking reaction, obtaining the cross-linked site concentration in the SU-8. The PEB step completes the condensation reaction as well as to amplify the cross-linking yield to enhance sensitivity and improve image contrast (Berry et al. 1989; Reichmanis et al. 1994). Finally, the development simulation is to calculate the development (etching) rate distribution into the SU-8 according to the obtained cross-linked site concentration during the PEB simulation. With the

**Fig. 2** Simulation flowchart for the UV lithography of SU-8



development rate distribution, the propagation of the development profiles of the SU-8 can be obtained by appropriate etching surface advancement algorithms introduced in section “[Simulations and Discussions](#).” The final development profiles can be visualized and output for further analysis for various purposes.

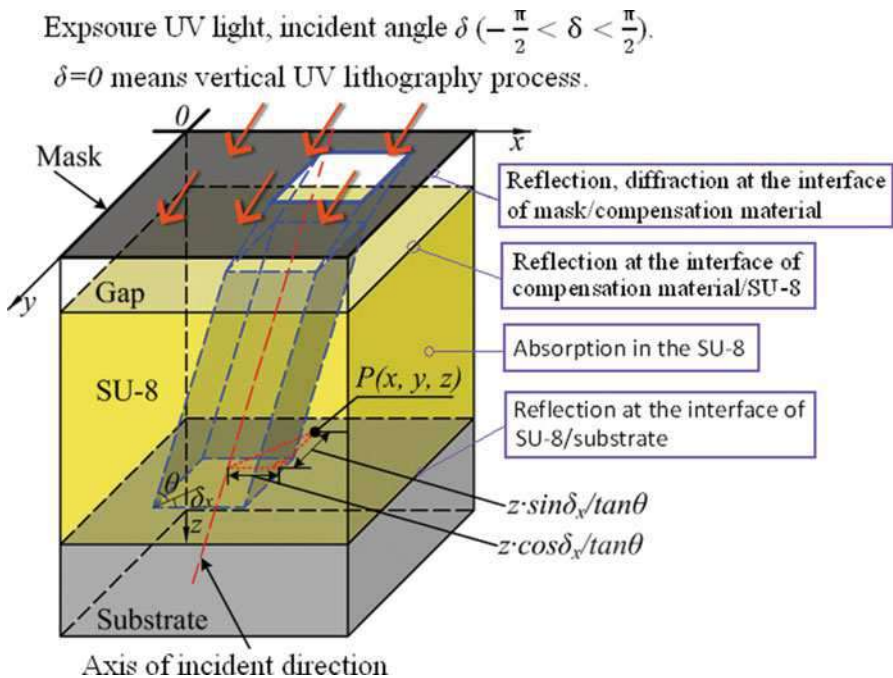
## Aerial Image Simulation Models

Aerial image simulation includes models describing illumination, transfer of the optical field through the mask and optical system of the lithographic equipment, and the propagation inside the SU-8. Its ultimate goal is to obtain the light intensity distribution into the SU-8 for the following simulation steps. Currently, contact and proximity exposure methods of the SU-8 using a so-called mask aligner are still cost-efficient alternatives to an optical projection exposure method in current MEMS field (<http://www.synopsys.com>; Bourdillon et al. 2000). At present, there are two main categories of models used for the aerial image simulations. One is based on scalar diffractive theory (Cheng et al. 1999; Chuang et al. 2002; Tian et al. 2005; Tang et al. 2005; Zhu et al. 2008, 2009; Zhou et al. 2014; Bourdillon et al. 2000), and the other

is based on rigorous electromagnetic field theory (Wong and Neureuther 1994; Erdmann et al. 2006, 2009; Nyysönen 1982; Lucas et al. 1996; Evanschitzky and Erdmann 2005; Schellenberg et al. 2005; Shao et al. 2008). The former can be further divided into models based on the Fresnel diffraction (Cheng et al. 1999; Chuang et al. 2002; Tian et al. 2005; Bourdillon et al. 2000), Fresnel–Kirchhoff diffraction (Zhou et al. 2007a, 2014; Zhu et al. 2008, 2009), and diffractive angular spectrum theory (Yang and Wang 2005), etc. At the same time, the latter can be mainly divided into models based on the finite difference time domain method (FDTD) (Wong and Neureuther 1994; Erdmann et al. 2006, 2009) and waveguide method (Nyysönen 1982; Lucas et al. 1996; Evanschitzky and Erdmann 2005; Schellenberg et al. 2005; Shao et al. 2008). Compared with the models based on rigorous electromagnetic field theory (Wong and Neureuther 1994; Erdmann et al. 2006, 2009; Nyysönen 1982; Lucas et al. 1996; Evanschitzky and Erdmann 2005; Schellenberg et al. 2005; Shao et al. 2008), including the waveguide method (Erdmann et al. 2009), models based on the scalar diffractive theory are faster and more practical (Born and Wolf 1999). In this chapter, a comprehensive aerial image model based on Fresnel diffraction is specifically developed to simulate the 3D inclined/vertical UV light intensity distribution into the SU-8 with the diffraction, refraction, absorbance, and reflection during light transmission efficiently considered simultaneously.

Ideal contact exposure without any gap between the mask and the thick SU-8 is not practical for inevitable errors such as surface flatness, surface roughness, etc. To reduce the diffraction effects, different materials such as glycerol have been adopted as compensation materials to fill the air gap between the mask and the resist during conventional contact/proximity lithography (Chuang et al. 2002). Because of the large refractive index mismatch between air and SU-8, the application of inclined UV lithography is restrained to fabricate structures with inclined angles lower than  $53.2^\circ$  (Zhu et al. 2008). To overcome this limitation, glycerol and water have been employed in an exposure process for index-matching materials (Chuang et al. 2002). For example, glycerol is employed as an index-matching material, extending the possible inclined angles from  $54^\circ$  (in air) down to  $19^\circ$  (in glycerol). As shown in Fig. 3, to model the 3D light intensity distributions during the thick SU-8 lithography for above cases, we define some parameters associated with the calculations of the UV light intensity distribution.  $\delta$  and  $\theta$  represent the incident angle in the air and the refraction angle of inclined UV light in SU-8, respectively, and  $\delta x$  is the angle between x-axis and the projection of incident direction on the xy-plane.  $n_1$ ,  $n_2$  and  $n_3$  are defined as the refractive indices of air, compensation materials and SU-8 photoresists, respectively.  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  stand for the UV light wavelength in air, compensation materials and SU-8 photoresists, respectively.  $R_1$ ,  $R_2$ , and  $R_3$  are the reflection coefficients at the interfaces of mask/compensation material, compensation material/SU-8 and SU-8/substrate interfaces, respectively.

Since the propagation of the incident UV light will be refracted on the compensation material/SU-8 interface, it is difficult to deal with the refraction in Fresnel–Kirchhoff diffraction integral equation. To solve this problem, the mask-shifting approach is utilized to handle the diffraction and refraction effects simultaneously. Based on previous researches (Zhou et al. 2007a, 2011), the light intensity of any



**Fig. 3** Schematic modeling of the 3D inclined and vertical UV light intensity distribution into the SU-8 photoresist using the improved paraxial approximation

calculation point  $p$  in photoresist for inclined and vertical UV lithography can be derived as

$$I_p = \frac{(1 - R_1)(1 - R_2) \cdot I_{\text{lamp}}}{2} \left( \{ [c(u_2) - c(u_1)]^2 + [s(u_2) - s(u_1)]^2 \} \cdot \{ [c(v_2) - c(v_1)]^2 + [s(v_2) - s(v_3)]^2 \} + R_3 \{ [c(u_4) - c(u_3)]^2 + [s(u_4) - s(u_3)]^2 \} \cdot \{ [c(v_4) - c(v_3)]^2 + [s(v_4) - s(v_3)]^2 \} \right) \quad (1)$$

where  $I_{\text{lamp}}$  is the original intensity of the incident UV light and  $c(u)$  and  $s(u)$  are the Fresnel integrals in  $x$ -direction, while  $c(v)$  and  $s(v)$  are those in  $y$ -direction.  $u_i$  ( $i = 1, 2, 3, 4$ ) and  $v_i$  ( $i = 1, 2, 3, 4$ ) represent the Fresnel numbers in  $x$ -direction and  $y$ -direction, respectively, written as

$$u_i^2 = \frac{2(n_3/n_2)}{\lambda_2(z - p_1 + p_2)/\cos \theta} \cdot (x_i - x - (z - p_1 + p_2) \cdot \tan \theta \cdot \cos \delta_x)^2, \quad i = 1, 2 \quad (2)$$

$$v_i^2 = \frac{2(n_3/n_2)}{\lambda_2(z - p_1 + p_2)/\cos \theta} \cdot (x_i - x - (z - p_1 + p_2) \cdot \tan \theta \cdot \cos \delta_x)^2, \quad i = 1, 2 \quad (3)$$

$$u_i^2 = \frac{2(n_3/n_2)}{\lambda_2(2d + p_1 - {}^{1z+p_2})/\cos\theta} \cdot (x_i - x - (2d + p_1 - z + p_2) \cdot \tan\theta \cdot \cos\delta_x)^2, \quad i = 3, 4 \quad (4)$$

$$\nu_i^2 = \frac{2(n_3/n_2)}{\lambda_2(2d + p_1 - {}^{1z+p_2})/\cos\theta} \cdot (x_i - x - (2d + p_1 - z + p_2) \cdot \tan\theta \cos\delta_x)^2, \quad i = 3, 4 \quad (5)$$

where  $x_i$  and  $y_i$  represent the latitudinal coordinates and longitudinal coordinates of the mask aperture, respectively,  $d$  represents the thickness of the SU-8 layer,  $z$  stands for the vertical distance from the mask plane to the calculation point in SU-8,  $p_1$  is the primal gap thickness between the mask and SU-8, and  $p_2$  represents the gap thickness after shifting the mask.

It should be noted that if no compensation material is employed to reduce the diffraction effects, our aerial image model is still corrected, by defining  $n_2 = n_1$  and  $\lambda_2 = \lambda_1$  in above equations. Furthermore, since the relationship between the incident and refractive angle always satisfies Snell's law for each medium, no matter whether compensation materials are utilized or not,  $\theta$  in Eqs. (2), (3), (4), and (5) is obtained by

$$\theta = \sin^{-1}(n_1 \sin\delta/n_3) \quad (6)$$

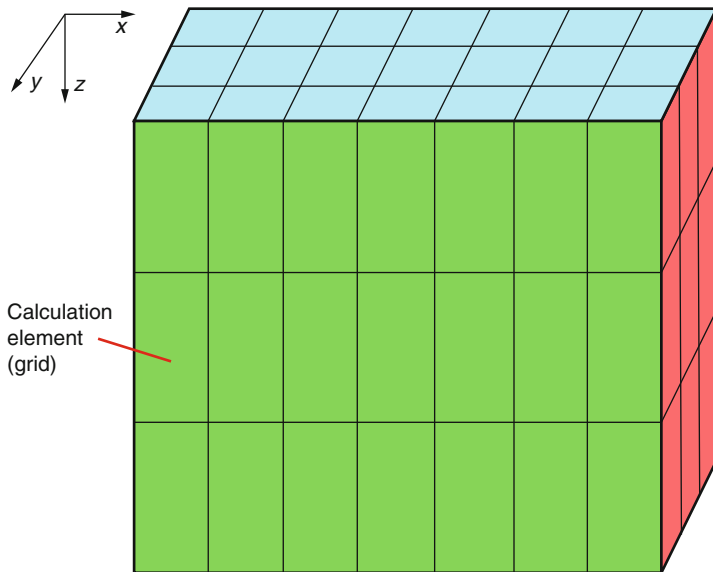
For the case that index-matching materials are employed during the lithography process, the presented aerial image model is also effective, only by defining  $n_1$  equal to the refractive index of the specified index-matching material.

For the 3D light intensity distribution calculation, the whole calculation domain in SU-8 is divided into a matrix of identical cubic elements (grids) with the same side length in our previous study. In this paper, adaptable element size of the cubic elements in x-, y-, and z-direction is adopted in the whole calculation domain, as shown in Fig. 4. Thus the total elements needed to be calculated are obviously reduced, so does the calculation time. After the light intensity distribution calculation, redivision of the whole calculation domain in SU-8 is carried out for the following simulation step.

## Exposure Simulation Models

According to the obtained UV light intensity distribution into the SU-8 by above aerial image simulation, an exposure model will be employed to describe the exposure kinetics in the SU-8. The original *Dill* model (Dill et al. 1975) plays an important role in conventional optical lithography simulations (Mack 2005). The simulation results are usually compatible with the experimental ones for thin photoresists with the original *Dill* model.

Unlike thin photoresists, there are many effects caused by the nonlinear factors in the exposure process of thick photoresists like SU-8, such as the concentration distribution of contents, the change of the refractive index, etc. Thus, the original *Dill* model has to be modified for dealing with these problems resulting from the



**Fig. 4** Schematic diagram of adaptable element size in x-, y-, and z-direction to speed up the aerial image model

nonlinear characteristics of thick photoresists (Zhou et al. 2007a, 2014; Sohn et al. 2001; Liu et al. 2005). Liu et al. presented an enhanced *Dill* model by using a set of modeling parameters, which will vary with the photoresist thicknesses and the different processing conditions (Sohn et al. 2001). Zhou et al. developed an improved *Dill* model (Zhou et al. 2014), accurately describing the nonlinear effects in the SU-8 photoresist, to simulate the exposure kinetics during exposure process (Zhou et al. 2007a, 2014) by

$$\frac{\partial m(x, y, z, t)}{\partial t} = -I(x, y, z, t)m(x, y, z, t)C(z) \quad (7)$$

$$\alpha_i(x, y, z, t) = A(z)m(x, y, z, t) + B(z) \quad (8)$$

$$I(x, y, z, t) = I_0(x, y, z) \cdot \exp\left(-\int_0^l \alpha_i(x, y, z, t) dz\right) \quad (9)$$

$$C_A(x, y, z, t) = 1 - m(x, y, z, t) \quad (10)$$

where the SU-8 is divided into several layers,  $\alpha_i$  is the absorption coefficient for the  $i$ th layer,  $I_0(x, y, z)$  is the light intensity from the aerial image simulation,  $m(x, y, z, t)$  and  $C_A(x, y, z, t)$  denote the normalized photoacid generator concentration and HSBF<sub>6</sub> concentration, respectively,  $z(\mu\text{m})$  is the variable to indicate the photoresist depth (the distance from the top to the bottom of the SU-8 layer), and  $l$  indicates the

distance that the UV light passes through in the SU-8. The *Dill* parameters vary with the SU-8 thickness:

$$A(z) = i_1 + i_2 \cdot z + i_3 \cdot z^2 (\mu\text{m}^{-1}) \quad (11)$$

$$B(z) = j_1 + j_2 \cdot z + j_3 \cdot z^2 (\mu\text{m}^{-1}) \quad (12)$$

$$C(z) = l_1 + l_2 \cdot z + l_3 \cdot z^2 (\text{cm}^2/\text{mJ}) \quad (13)$$

$i_1, i_2, i_3, j_1, j_2, j_3, l_1, l_2,$  and  $l_3$  are the constants obtained by fitting experimental results using the conventional “*Dill* graphical method” (Zhou et al. 2007a; Henderson et al. 1997). For SU-8 2000 series, the values for these constants are fitted to be (Zhou et al. 2007a, 2014):  $i_1 = -0.0031, i_2 = 4.7878 \times 10^{-6}, i_3 = -3.0860 \times 10^{-9}, j_1 = 0.0079, j_2 = -7.5510 \times 10^{-6}, j_3 = 4.0971 \times 10^{-9}, l_1 = 0.0865, l_2 = -1.3078 \times 10^{-4},$  and  $l_3 = 6.1193 \times 10^{-8}$ .

## PEB Simulation Models

During the PEB process, the photoacid HSbF<sub>6</sub> will act as a catalyst to induce the cross-linking of epoxy resin. HSbF<sub>6</sub> catalyzes ring opening of epoxies and turns them into cross-linked sites. In this way, SU-8 resin transfers from a low-molecular-weight material to a highly cross-linked 3D polymer network. This polymer network is significantly less soluble than the polymer resin without cross-linking reaction. The details of catalytic chemical reactions of epoxy can be found in Reference (Bobbitt 2001).

Since the HSbF<sub>6</sub> concentration is not uniform in SU-8, there is a diffusion process of the HSbF<sub>6</sub> when the catalytic conversion is undergoing. The chemical reactions and diffusion of species happen simultaneously and couple with each other, so they have to be considered simultaneously during the PEB process simulation. As for the PEB process simulation of thin photoresists in IC field, some complicated and accurate models have been presented (Neureuther and Mack 1997; Cole et al. 2001; Erdmann et al. 2009). However, if these models are expected to be adopted for the PEB process simulation of the SU-8, much efforts should be made to measure many model parameters. Currently, the coupled reaction–diffusion kinetics (Zuniga et al. 1993) is used to describe the PEB process for the SU-8. Since the reaction order for HSbF<sub>6</sub> and cross-linked sites is 1 for the lithography of the SU-8, the equations for the PEB simulation can be simplified as (Zhou et al. 2007a, 2014; <http://www.microchem.com>)

$$\frac{\partial C_{cs}}{\partial t} = k_1(1 - C_{cs})C_A^q \quad (14)$$

$$\frac{\partial C_A}{\partial t} = \nabla \cdot (D_{\text{acid}} \nabla C_A) - k_2 C_A \quad (15)$$

where  $C_{cs}$  is the normalized concentration of the cross-linked sites. One cross-linked site is created when one epoxy ring is open.  $q$  is a fitting parameter,  $D_{\text{acid}}$  is the

diffusivity of  $\text{HSbF}_6$ ,  $k_1$  is the site cross-linking reaction coefficient, and  $k_2$  is the photoacid loss reaction coefficient.  $k_1$ ,  $k_2$ , and  $D_{\text{acid}}$  are possible to fit the temperature dependence by an Arrhenius law. Sensu et al. have determined the values of these parameters using experimental methods (Sensu et al. 2005).

## Development Simulation Models

The acid-catalyzed cross-linking reaction determines the development (etching) rate distribution within the SU-8, and various development rate models have been proposed for many years, such as *Mack* model (Mack 1985, 1992), *Weiss* model (Weiss et al. 1995), *Notch* model (Arthur et al. 1998; Mack and Arthur 1998), and *Enhanced Notch* model (Arthur et al. 1998; Mack and Arthur 1998). These models can be adopted to describe the relationship between the development rate and the cross-linked site concentration. Namely, they can be adopted to calculate the development rate for each grid in the SU-8, according to the cross-linked site concentration of each grid. Namely, the parameter values for development rate models are dependent of the PEB conditions (<http://www.synopsys.com>). However, some other parameters, besides the cross-linked site concentration, should also be measured for these models. Agitation methods are usually adopted for high aspect ratio and/or thick film structures during the development processes, and the development rate can be increased over several times, compared with the development without agitation method (Zhou et al. 2007a, 2014; <http://www.microchem.com>).

For the SU-8 development simulation in this chapter, the *Notch* model (Arthur et al. 1998; Mack and Arthur 1998) is extended to three dimensions to describe the relationship between the dissolution rate and cross-linked site concentration, obtaining the photoresist etching rate distribution for the development simulation. Zhou et al. provided some *Notch* model parameters for SU-8 2000 series (Feng et al. 2006),  $R_{\text{max}} = 1.02 \mu\text{m/s}$ ,  $R_{\text{min}} = 0.00009 \mu\text{m/s}$ ,  $n_1 = 1.2$ ,  $n_{n\text{-notch}} = 20.0$ , and  $M_{\text{th-notch}} = 0.43$ , for the development process with agitation methods. Although sometimes it is necessary to adopt the depth-dependent dissolution rate effect (empirical model) and the swelling model (semiempirical model) to improve the simulations (Zhou et al. 2007a; Zhu et al. 2009; Zanghellini et al. 1998), it should be noted that the swelling effect seems more significant for original SU-8 resist series than its newer formulations (SU-8 2000 series and SU-8 3000 series) (Zhou et al. 2011; <http://www.microchem.com>). After obtaining the development rate distribution into the SU-8, the propagation of the development profiles can be obtained by appropriate etching surface advancement algorithms introduced in section “[Algorithms for Etching Surface Advancement Simulation](#).”

---

## Algorithms for Etching Surface Advancement Simulation

To compute the evolution of the lithography profiles of SU-8, the development (etching) surface should be first represented by any one of algorithms (techniques). Thus the etching surface advancement (movement) with time can be computed based



on the etching rate distribution into the SU-8. Etching surface advancement algorithms are a key component of the profile simulations as they must accurately reflect the advancement without excessive computation. Currently, main algorithms for etching surface advancement simulation, including the string algorithms (Jewett et al. 1977; Hamaguchi and Rosnagel 1995; Levinson et al. 2000; Leunissen et al. 2003; Vyvoda et al. 2000; Arnold et al. 1994), ray-tracing algorithms (Hagouel 1976; Toh et al. 1994; Cooperberg 2002), cellular automaton algorithms (Neumann 1966; Hoang et al. 2008; Hsu et al. 2013; Stout et al. 2006; Zhou et al. 2005, 2007c,d, 2009; Gosálvez et al. 2008; Scheckler et al. 1993; Karafyllidis and Thanailakis 1995; Karafyllidis 1999; Strasser and Selberherr 1995), and fast-marching algorithms (Sethian 1995, 1996, 1999, 2001; Osher and Sethian 1988; Adalsteinsson and Sethian 1995; Chopp 2001; Zhao 2005; Gremaud and Kuster 2006; Yatziv et al. 2006; Hassouna and Farag 2007), have been presented for the lithography simulations.

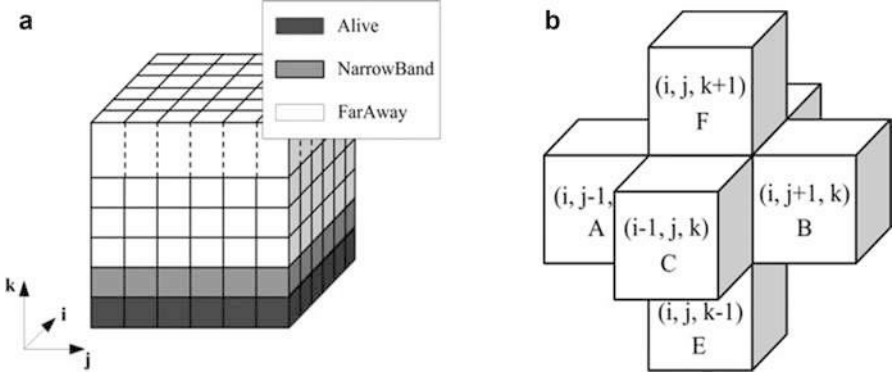
The string method was introduced by Jewett et al. (1977) and was initially used to describe photoresist development for thin photoresists. However, the delooping algorithms must be incorporated to assure accuracy and correctness accuracy for the string method. In the case of 3D simulations, the straight line segments are changed into triangles and polygons, and the delooping in 3D is much more computationally expensive, since it involves all triangle or polygon intersections in a mesh containing thousands of triangles or polygons. The ray-tracing algorithm began to be used for etching surface advancement during lithography simulations in the 1970s (Hagouel 1976). In the ray-tracing algorithms, the developed and undeveloped regions are defined with different development rates. The propagation of the etching vectors is calculated similar to the propagation of light using Snell's law of refraction. The ray-tracing algorithm is easy to implement in 2D cases. However, the initial rays must be chosen carefully; otherwise, some regions may not be reached. Theoretically speaking, the ray-tracing algorithms are probably attractive for 3D case, considering their advantages in speed and computational simplicity in 2D cases. However, the difficulties in reconstruction of the surface from independent rays make the ray-tracing algorithms not easy to be implemented in 3D and not very attractive for 3D etching simulations. The cellular automaton methods (algorithms), first presented by von Neumann following Ulam's suggestions (Neumann 1966), discretize the computation domain (either 2D or 3D) into a set of cells that are associated with different materials. The states at each cell are updated simultaneously based on the states of the cell and its neighboring cells at discrete time steps, according to a definite set of "local rules" (Neumann 1966; Zhou et al. 2005, 2007c, 2007d, 2009; Gosálvez et al. 2008; Scheckler et al. 1993; Karafyllidis and Thanailakis 1995; Karafyllidis 1999; Strasser and Selberherr 1995). Since the "local rule" is defined to govern only local relationships among the neighboring cells, the governing equation for the whole domain is not necessary. Surface advancement is represented by removal or addition of cells or some parts of cells. The cellular automaton algorithms have been used to simulate the shallow trench isolation etch with chlorine plasma (Hoang et al. 2008; Hsu et al. 2013), plasma etching of vias and trenches in SiO<sub>2</sub> (Stout et al. 2006), anisotropic etching of silicon

(Zhou et al. 2007c, 2009; Gosálvez et al. 2008), photoresist development (Scheckler et al. 1993; Karafyllidis and Thanailakis 1995; Karafyllidis 1999; Zhou et al. 2005, 2007d), etc. The advantages of the cellular automaton algorithms include the ease to handle topological changes for arbitrary geometries, the ease to be implemented in 3D, etc. The accuracy of cellular automaton algorithms depends on the number of cells. A large number of cells improve accuracy but at the same time cost more computation time. The simulation using the 3D dynamic cellular automaton method is slow. For a simulation array with  $500 \times 500 \times 500$  grids, the simulation process usually needs several hours for a typical personal computer configuration: OS Windows XP SP3, CPU Intel Core2 @2GHz, and DRAM 2GB. For a relatively large 3D simulation array, the simulation using current fast-marching algorithm (Sethian 1995, 1996, 1999, 2001; Osher and Sethian 1988; Adalsteinsson and Sethian 1995; Chopp 2001; Zhao 2005; Gremaud and Kuster 2006; Yatziv et al. 2006; Hassouna and Farag 2007) is about eight times faster than the 3D dynamical cellular automaton methods but needs more memory elements (Sethian 1995, 1996, 1999, 2001). For a simulation array with  $500 \times 500 \times 500$  grids, the lithography simulations can't be implemented using current fast-marching algorithms with a typical personal computer configuration, OS Windows XP SP3, CPU Intel Core2 @2GHz, and DRAM 2GB, since the memory limit is exceeded. This will limit the application of 3D simulation tools for thick photoresist lithography. To solve this problem, the fast-marching method was modified by citing hash table to store the location in the min-heap, reducing the memory element of current fast-marching method (Sethian 1995, 1996, 1999).

## Fast-Marching Method

The fast-marching method (algorithm), briefly cited as FMM, originally introduced by Sethian (Sethian 1995, 1996), is a stationary version of the general level set method (Osher and Sethian 1988; Adalsteinsson and Sethian 1995). The FMM is a powerful numerical technique for analyzing and computing moving fronts, because it is accurate, extremely fast, and easy to implement in three dimensions. Furthermore, the FMM can handle topological changes without special logic. To simulate the etching surface advancement, the FMM computes the time  $T_{(x, y, z)}$  when the resist-developer interface passes through each point in a grid. The development rate of the photoresist,  $R'_{(x, y, z)}$ , is only a function of position, and the Eikonal equation  $|\nabla T_{(x, y, z)}| R'_{(x, y, z)} = 1$  must be satisfied. Observing the resist-developer interface may only progress in one direction, an “upwind” scheme for viscosity solutions of Hamilton Jacobi equations can be devised to solve the Eikonal equation.

The flow of FMM can be divided into two parts: initialization and marching forward. In the initialization, the object is meshed. Arrays of FarAway, NarrowBand, and Alive points are then established, as in Fig. 5a. When marching forward, neighbors of points which have the minimum T value need to be updated by solving the Eikonal equation using the upwind Scheme (Adalsteinsson and Sethian 1995; Sethian 1999):



**Fig. 5** (a) Initialization of FMM and (b) Update of neighbors

$$F_{ijk} \left[ \begin{array}{l} \max\left(\max\left(D_{ijk}^{-x}T, 0\right), -\min\left(D_{ijk}^{+x}, 0\right)\right)^2 \\ \max\left(\max\left(D_{ijk}^{-y}T, 0\right), -\min\left(D_{ijk}^{+y}, 0\right)\right)^2 \\ \max\left(\max\left(D_{ijk}^{-z}T, 0\right), -\min\left(D_{ijk}^{+z}, 0\right)\right)^2 \end{array} \right]^{1/2} = 1 \quad (16)$$

Assume the point that to be recalculated locates at the center  $(i, j, k)$ , and its  $T$  value is  $x$  after updating; rewrite Eq. (16) as follows:

$$\begin{aligned} & \max(\max(x - A, 0), -\min(B - x, 0))^2 \\ & + \max(\max(x - C, 0), -\min(D - x, 0))^2 = h^2 / F_{ijk}^2 = f_{ijk}^2 \quad (17) \\ & + \max(\max(x - E, 0), -\min(F - x, 0))^2 \end{aligned}$$

Suppose  $A = \min(A, B, C, D, E, F)$ ,  $\min(C, D) \leq \min(E, F)$ , and  $A \leq x \leq A + f$ , so

when  $A + f \leq \min(C, D)$ ,  $(x - A)^2 = f^2$

when  $\min(C, D) \leq A + f \leq \min(E, F)$ ,  $(x - A)^2 + (x - \min(C, D))^2 = f^2$

and when  $A + f \geq \min(E, F)$ ,

$(x - A)^2 + (x - \min(C, D))^2 + (x - \min(E, F))^2 = f^2$ , and then the new  $T$  value of neighbors is obtained by solving the above equations based on the relationship between  $A + f$  and  $\min(C, D)$ .

Min-heap data structure is used to search the minimum  $T$  value point in the NarrowBand. The pointers point to the NarrowBand points which are stored in the min-heap structure according to the magnitude of their time value, and the features of the heap are maintained when the NarrowBand is marching forward, so that the time complexity of searching minimum value is  $O(1)$  while creating and maintaining the min-heap structure cost  $O(\log(n))$  time complexity;  $n$  is the number of points in the NarrowBand.

## Hash Fast-Marching Method

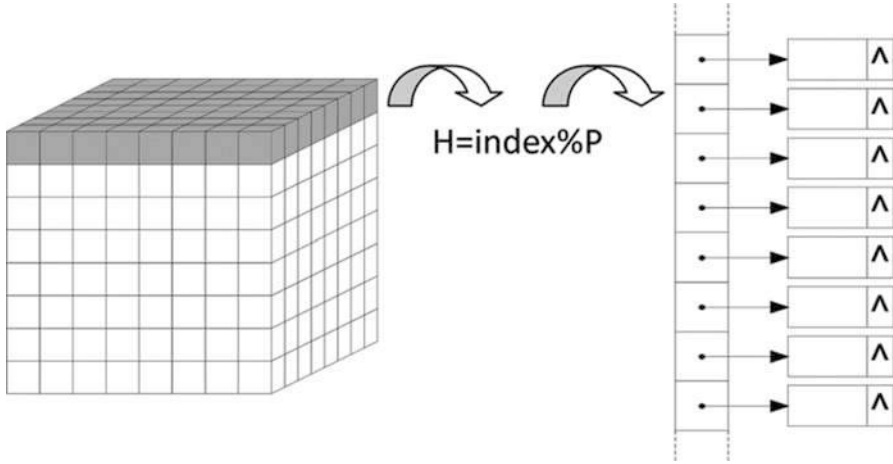
For a cube meshed by  $N \times N \times N$  grids, the original FMM needs four arrays to store:  $4N^3$  bytes to store the T values along with  $4N^3$  bytes to store the etch-rate array,  $0.1 \times 4N^3$  bytes to store the min-heap structure of NarrowBand points, and  $4N^3$  byte companion array to point out the index of nodes in the min-heap, occupying  $12.4N^3$  bytes of memory in total. All calculations in the FMM are concentrated in the Narrowband, and the Alive and FarAway points don't need to be calculated. Thus the rate values in the rate array are also concentrated in current NarrowBand, without other two regions. We propose an improved fast-marching method: a hash table is established to extract the NarrowBand, a data structure is designed to store the nodes' information in NarrowBand, and the rate function is reused to store the Alive nodes' time value, so that the space complexity drops from  $12.4N^3$  to  $4.8N^3 + O(N^2)$ , without computational speed loss. The improved algorithm is called the hash fast-marching method (HFMM).

### Data Structure of HFMM

In the HFMM, the hash Table (Zhou et al. 2014) is established by calculating the remainder of the index, which is a 1D value corresponding to the 3D coordinates. The 1D index will not overflow, because the HFMM requires about 40% memory space of the original FMM (see in the next section), that is,  $40\% \times 12.4N^3 = 4.96N^3$  bytes,  $N^3$  is the total number of grids, suppose the HFMM is running on a 32-bit 4GB DRAM personal computer, and the total grid number  $N^3$  can be  $4 \times 109/4.96 = 8.06 \times 10^8$  at most theoretically, which is far less than the "int" data range in a 32-bit computer. Therefore, it is feasible to map the 3D coordinates into 1D values.

Figure 6 describes the principle of the hash table. Set the hash function as  $H = \text{index} \% P$ ; the sign of % represents the modulo operation, and P is a big prime number, so the indexes of the 3D matrix are mapped to hash table through the modulo operation. The effective memory of hash table is desired to approach to 100%, which means that the load factor approximately equals 1.0, so separate chains are used to deal with conflicts (Zhou et al. 2014). Every chain links nodes sharing the same hash value, and the search operation is constant time complexity because only the modulo operation is needed. The average search length is  $1 + 0.5$ .

The HFMM needs three arrays to work: the original rate array, the NarrowBand hash table, and the min-heap based on the nodes' time value in the NarrowBand. The original rate array stores speed value in normal direction of each node. The NarrowBand hash table data structure contains the index value of each node in the NarrowBand, the position in the min-heap, time value when the front reaches the point, six pointers which point to the adjacent nodes in space, and the pointer point to the next node in the chain list of the hash table, totally ten values of float type. The function of the min-heap is the same as mentioned above, which is extracting the node that has the minimum time value in the NarrowBand.



**Fig. 6** Initialization of the NarrowBand hash table

### Reuse of Etch Rate Array

In the HFMM, the rate array needs to be reused. Though the FMM only calculate values of nodes in the NarrowBand, adjacent nodes in Alive and FarAway are useful, but the hash table does not store the time value of nodes outside the NarrowBand, so we try to reuse the original rate array to store the Alive nodes' time.

It is very simple to reuse: when the NarrowBand node on top of the min-heap becomes "Alive," the rate value in the corresponding point will be set to the opposite of the minimum value (current value), that is,

$$F(i \text{ min}, j \text{ min}, k \text{ min}) = -(\text{min} \rightarrow \text{time}) \quad (18)$$

where  $(i \text{ min}, j \text{ min}, k \text{ min})$  is the index of the current minimum point and min is the corresponding node in the hash table.

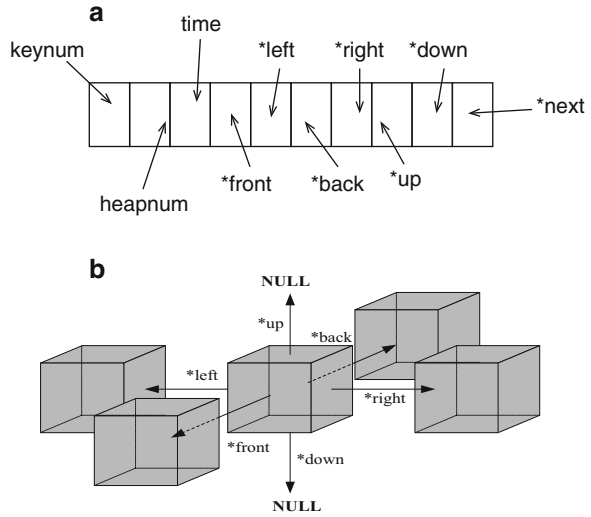
Why it can work? The reasons are:

Firstly, in the FMM, the time values of Alive point are fixed; there is no need to calculate the points' time values. And from Eq. (17), the points' rate values are useless in the Alive area, so the points' rate value can be reused.

Secondly, one of the premise conditions in the FMM is that the rate values are sign-preserving in the whole field. Therefore, nodes which have the opposite sign of the original rate value can be considered to be Alive points. Combining with the NarrowBand hash table, we can easily distinguish the three states of the nodes.

As a result, the reuse of the rate array not only saves memory but also stores the Alive points' time values and status easily without increasing time and space complexity.

**Fig. 7** The data structure of the NarrowBand hash table node



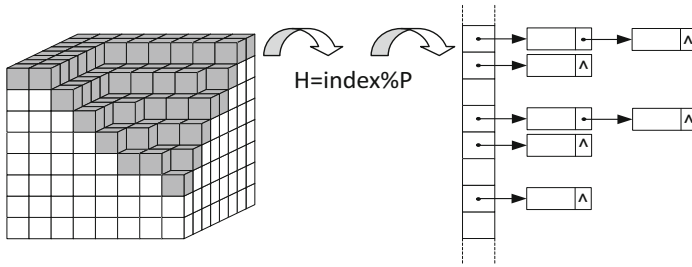
**NarrowBand Hash Table**

The NarrowBand hash table stores the information of the NarrowBand nodes, as shown in Fig. 7a; “keynum” is the node’s coordinate, “heapnum” is the position of the node in the min-heap, “time” records the time value when front reaches this node, and six pointers \*front, \*back, \*left, \*right \*up, and \*down point to the adjacent front, back, left, right, up and down nodes in space, respectively. If the adjacent nodes are available in NarrowBand, the pointers point to the corresponding nodes. If not, point to NULL. Assume that the current surface is like the Narrowband shown in Fig. 5; the two adjacent up and down points are not existing in NarrowBand, so \*up and \*down are pointing to NULL; the spatial relationship of one node in Narrow-Band is shown in Fig. 7b. In addition, the \*next pointer is used to point to the next node sharing the same hash value in the hash table; if the front is like the case in Fig. 4, every \*next points to NULL when P is big enough and there is no conflict in the hash table.

To be noticed, in the propagation of the HFMM, the hash table should update with the evolving of the NarrowBand, matching with the front (see in Fig. 8). Hence, when the FarAway point is brought into the NarrowBand, this node should be added into to the right location in the hash table assigned by hash value. While the time value of one node in the NarrowBand becomes the minimum (the root node of the min-heap), the node should be removed from the hash table.

**Initialization and Propagation**

After dividing the computational area into grids and calculating the rate values of each grid, the magnitude of P is willing to be determined. In general, P is the largest



**Fig. 8** One moment of the propagation

prime number which is less than the length of the hash table. Then the next initialization steps are:

- (i) Build a hash table according to the coordinate values with periodic boundary conditions, and assign keynum time, six adjacent pointers, and the \*next pointer for each node in NarrowBand hash table.
- (ii) Build a min-heap by the time values of each node and set the heapnum as the node's position in the min-heap.

The HFMM is overall the same with the original algorithm in the process of propagation, but creating and deleting hash node operations are frequently required to maintain synchronized changes of the hash table and the Narrowband. In order to minimize the resource consumption caused by the system application and free of memory, a FIFO (first in, first out) queue is established to recycle the node deleted from the hash table, so that new hash node can be directly fetched from the queue. The basic propagation flow of HFMM is as follows:

- (i) Remove the root node of the min-heap, put the corresponding node in the NarrowBand hash table into the recycling queue, and set the pointer pointing to the min-node as NULL and the corresponding rate value as the opposite of the current minimum time value.
- (ii) Determine the coordinate values and the hash values of non-Alive adjacent points of the minimum point based on the space relationship.
- (iii) Find the adjacent points in NarrowBand hash table according to the hash value acquired from the last step; if the adjacent points can be found in NarrowBand, recalculate the time value, and update the location in the min-heap; if cannot, it means the point lies in the FarAway; insert this node into the NarrowBand hash table according to the hash value, and insert it into the min-heap according to the calculated time value for the first time.
- (iv) Repeat step (i) to step (iii), until NarrowBand hash table is empty.
- (v) If the target time is known, the termination condition could be the time when the minimum time value is greater than the target time value, instead of the time when the NarrowBand hash table is empty.

## Tests and Discussions

To investigate the performance of the 3D HFMM, the Gauss rate function has been used to investigate the performance of HFMM. Suppose the initial profile is also a plane at  $z = 0$  in a unit cubic, another rate function given by Gaussian rate function has been applied in the HFMM. This rate function can efficiently express the very high contrast of etch rate near the final resist edge. The Gauss rate function is given as

$$F(x, y, z) = e^{-64[(x-0.5)^2+(y-0.5)^2]}(\cos^2(12z) + 0.01) \quad (19)$$

where  $x$  and  $y$  represent the two coordinates of two mutually orthogonal directions in a horizontal plane and  $z$  represents the coordinate in the height direction. Figure 9 compares the testing results with function (19) under different mesh grids at the same time. The testing result is accurate when the grid number is  $50 \times 50 \times 50$ , and the accuracy increases when the grid number increases.

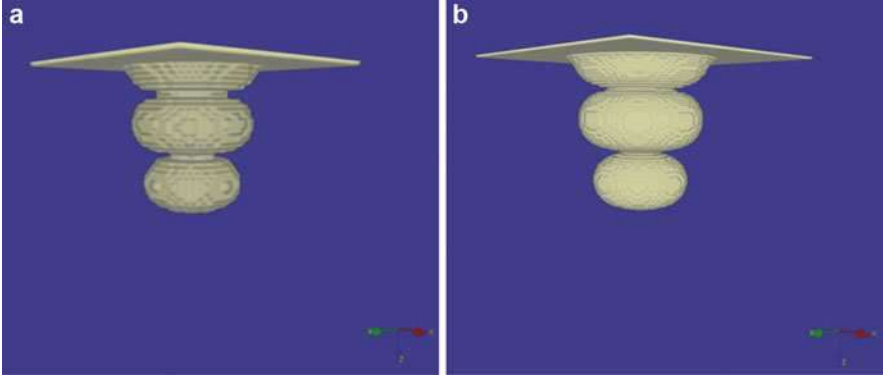
Table 2 is the comparison of the running time and memory consumption between the HFMM and FMM when calculating the profile at  $t = 10$  using Gauss rate function (19). We can see from this table that the FMM consumes 99.2 MB when the grid number is  $200 \times 200 \times 200$ , while in the HFMM, only 42.5 MB of memory are required, which is 57.2% less than the original method. When the grid number reaches  $400 \times 400 \times 400$ , original FMM needs 793.6 MB, while HFMM uses only 323.7 MB of memory, 59.2% less than the original one, and also the saved proportion will increase with the increasing of total grid number. Consider the running time; the HFMM needs basically the same time as the FMM; the reason is that the presented hash table has a favorable distribution so that the searching operation nearly occupies nearly no extra time.

Figure 10 demonstrates the surface appearance under the promotion of Gauss rate function at  $t = 6, 10, \text{ and } 12$  when the grid number is  $200 \times 200 \times 200$ . Simulation results propagate outward with the time so that the HFMM can easily deal with this high contrast function and get continuous and clear surface profiles.

As the memory consumption in the 3D HFMM changes in the marching process, memory values in Table 2 are the maximum value in this marching process. Table 3 compares the precision of the HFMM and the original FMM by comparing their breakthrough time and their relative errors under Gaussian rate function on different grid size. The breakthrough time of the original FMM comes from [www.intellisense.com](http://www.intellisense.com). The breakthrough time is defined as the time when the front of the photoresists reaches its bottom for the first time. Measured with a fourth-order Runge–Kutta scheme, the exact breakthrough time is 10.36761, while the breakthrough time calculated by HFMM on  $200 \times 200 \times 200$  is 10.3771, which has only 0.091% relative error.

Figure 11 indicates that the HFMM and the original FMM have nearly the same computational speed. In the memory consumption, for an object divided by  $N \times N \times N$  grids, the original 3D FMM needs three whole arrays as mentioned

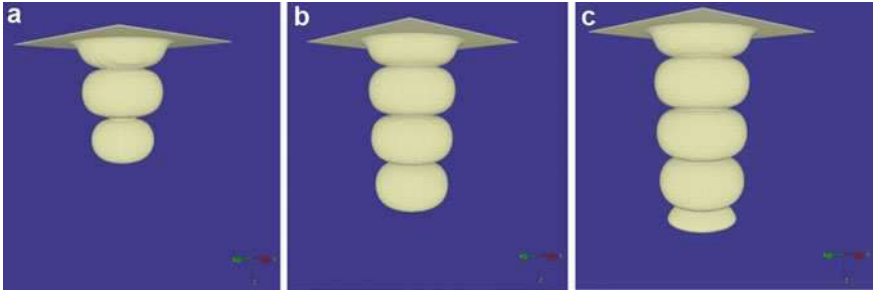




**Fig. 9** Profiles with Eq. (19) as the rate function at  $t = 6$  under (a)  $50 \times 50 \times 50$  and (b)  $100 \times 100 \times 100$  mesh grids

**Table 2** Comparison of the running time and memory usage of the HFMM and FMM at  $t = 10$  when using Eq. (19)

Methods	Grid no.	$50^3$	$100^3$	$200^3$	$400^3$
HFMM	Initialization(s)	0.047	0.375	2.89	24.50
	Propagation(s)	0.047	0.329	4.125	45.530
	Total(s)	0.094	0.704	7.015	70.031
	Memory (MB)	0.84	5.8	42.5	323.7
FMM	Initialization(s)	0.063	0.422	3.594	29.234
	Propagation(s)	0.016	0.266	3.672	41.985
	Total(s)	0.079	0.688	7.266	71.219
	Memory (MB)	1.55	12.4	99.2	793.6

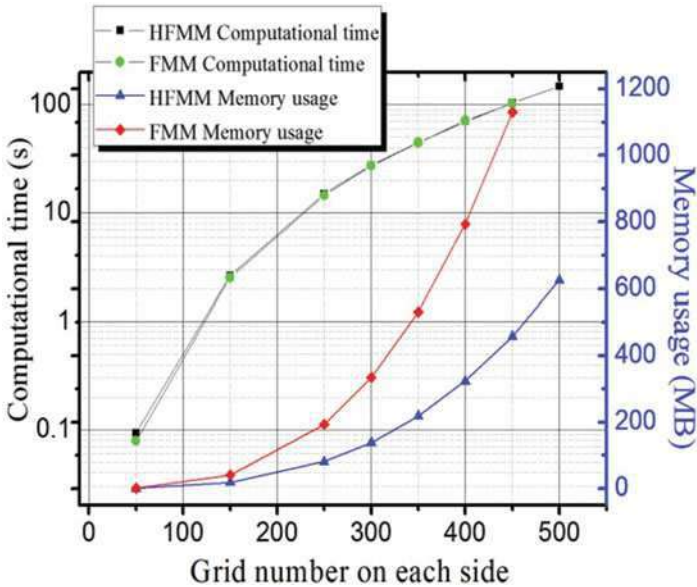


**Fig. 10** Simulation profiles with Eq. (19) at (a)  $t = 6$ , (b)  $t = 10$ , and (c)  $t = 12$  under  $200 \times 200 \times 200$  mesh grids

before, occupying  $3 \times 4 N^3 = 12 N^3$  bytes, and a  $0.1 \times 4 N^3 = 0.4 N^3$  byte min-heap in addition,  $12.4N^3$  bytes totally. While in the HFMM, there are one whole  $4N^3$  byte rate array and one  $0.4N^3$  byte min-heap which are the same as the original method, and one hash table occupying  $0.4N^3 + 40c$  bytes,  $4.8N^3 + 40c$

**Table 3** Comparison of the breakthrough time under different grid sizes when using Eq. (19)

Grid number	$50^3$	$100^3$	$200^3$	Relative error
FMM	9.8091	10.3398	10.3667	Max = 5.39% Min = 0.009%
HFMM	10.8883	10.45457	10.3771	Max = 5.02% Min = 0.091%



**Fig. 11** Comparisons of the running time and memory usage of the original FMM (Sethian 2001) and the HFMM using Eq. (19) on different grid sizes

bytes in total, where  $c$  is the number of nodes in the NarrowBand, so  $c = O(N^2)$  in general case, affected by the different test rate function and different target time, so the HFMM can save  $7.6N^3 - 40c$  bytes compared with the original FMM, the relative saved memory is  $(7.6N^3 - 40c) / 12.4N^3 \rightarrow 7.6 / 12.4 = 61.3\%$  when the value of  $N$  is large enough. Therefore, both from the theoretical calculation and experimental measurement data, our 3D HFMM can reduce about 60% of the memory space compared with the original FMM. In 3D applications, the advantage of low spatial complexity becomes greater in a more refined mesh grid. When the number of grids comes to  $500 \times 500 \times 500$  in the above two examples, the original FMM is unable to run as the limitation of physical memory in the 2 GB personal computer, while for HFMM, only less than 650 MB of memory are required. This also means the HFMM can obtain a more accurate result by refining the mesh grid.

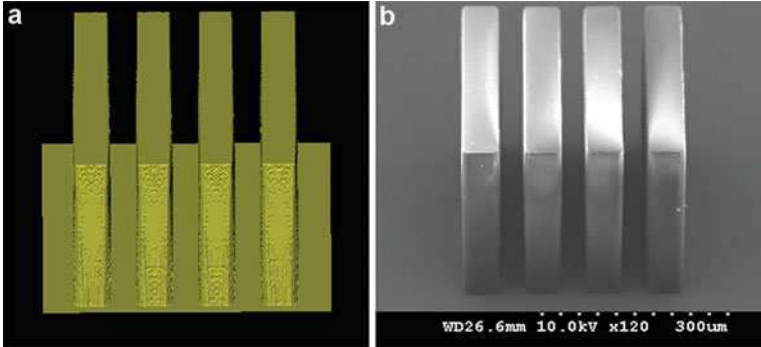
## Simulations and Discussions

To investigate above approaches, some simulations under different lithography conditions have been conducted using various simulation grid arrays and mask shapes. In this paper, some experimental results without compensation materials are presented to verify the simulation results. All studies are carried out on SU-8 2075 under UV source with 365 nm (2.6 mW/cm<sup>2</sup>) radiation. The SU-8 2075 is spin-coated on 3 in. diameter silicon or glass wafers (thicknesses =  $430 \pm 10 \mu\text{m}$ ) to form SU-8 layers with thickness of  $300 \pm 15 \mu\text{m}$  and  $140 \pm 12 \mu\text{m}$ . For the convenience of comparisons of simulation and experimental results, only simulation profiles for relative simple masks are presented. More results for complex masks can be found in the supplemental materials.

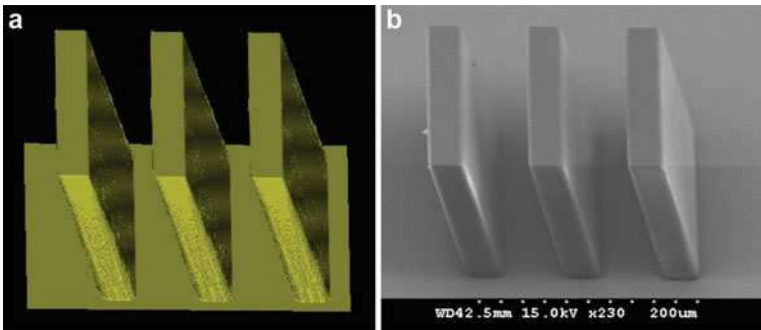
Figure 12 presents the simulation and experimental results of vertical UV lithography with a “|||” shape mask. Thick SU-8 with  $300 \pm 15 \mu\text{m}$  is formed on the silicon wafer, and the exposure time is 450 s, and the development time is 15 min. The line/space of the mask shape is 50/50  $\mu\text{m}$ . During the simulation, 293.5 MB memory elements are adopted for the UV light distribution simulation using our aerial image model, while a  $750 \times 650 \times 450$  simulation grid array is adopted for the following exposure simulation, PEB simulation, and development simulation. The simulation time is about 35 min for the whole process. The maximum error between the simulation and experimental profiles is found to be less than 7.1% in Fig. 12.

For the inclined UV lithography process, a “|||” shape mask has been adopted, and the line/space of the mask shape is 30/60  $\mu\text{m}$ . The SU-8 layer with  $140 \pm 12 \mu\text{m}$  is spin-coated on the glass wafer with thin TiO<sub>2</sub> film formed to eliminate the reflection effects at the SU-8/substrate interface. The exposure time is 240 s and the development time is 12 min. Figure 13 presents the simulation and experimental results of inclined UV lithography thick SU-8 for 30.0° incident angle. In Fig. 13, the maximum error between the simulation and experimental profiles is less than 7.3%. Here 263.4 MB memory elements are adopted for the UV light distribution simulation using our aerial image model, while a  $720 \times 720 \times 280$  simulation grid array is adopted for the following exposure simulation, PEB simulation, and development simulation. The simulation time is about 39 min for the whole process. This indicates that the calculation during inclined UV lithography simulation is a little more complex than the vertical UV lithography simulation.

Although many efforts have been made, the applications of UV lithography simulations of the SU-8 are significantly limited, compared with current lithography simulations of thin photoresists in the IC field. To some degree, UV lithography simulations of SU-8 are behind of the lithography technology development. Up to now, few simulation tools have been released for the UV lithography of SU-8 or other thick photoresists ([www.intellisense.com](http://www.intellisense.com)). To extend the application of lithography simulations of the SU-8, several challenging issues should be concerned or solved. As we all know, 3D simulations for complex masks, as shown in Fig. 14, even full-chip level simulations, are useful for the lithography process optimization and dimensional control of microstructures. However, the simulations need too much computer resources, and the computer resources for certain PC are limited.



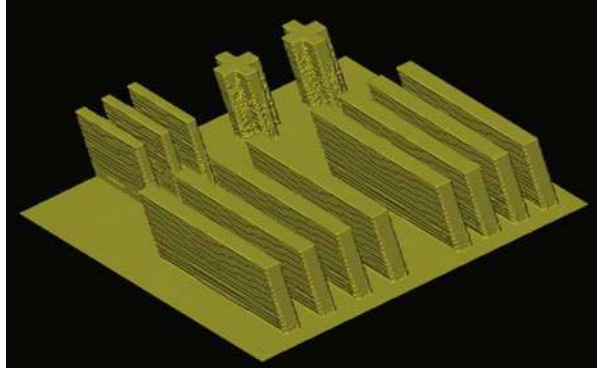
**Fig. 12** (a) Simulation result and (b) experimental result of the vertical UV lithography for  $300 \pm 15 \mu\text{m}$  thick SU-8 with a “|||” shape mask: exposure time = 450 s, development time = 15 min, line/space = 50/50  $\mu\text{m}$ , the first step PEB condition was 65 °C/15 min, and the second step PEB condition was 95 °C/45 min



**Fig. 13** (a) Simulation result and (b) experimental result of the inclined UV lithography for  $140 \pm 12 \mu\text{m}$  thick SU-8 with a “|||” shape mask for 30°UV incident angle with  $\text{TiO}_2$  film as antireflection layer: exposure time = 240 s, development time = 12 min, line/space = 30/60  $\mu\text{m}$ , the first step PEB condition was 65 °C/12 min, and the second step PEB condition was 95 °C/35 min

To solve this problem, fast and efficient algorithms are greatly expected. Furthermore, the mainstream algorithms (cellular automaton and fast-marching algorithms) have an intrinsic parallel updating nature, and the corresponding simulations are highly inefficient when performed on classical central processing units (CPUs), which are designed for the sequential execution of tasks. So, parallelization methods for the UV lithography simulation of the SU-8 on graphics processing units (GPUs) will be helpful in the very near future. A GPU has been used as a cost-efficient example of the parallel architectures for silicon anisotropic etching simulations recently (Ferrando et al. 2011). As mentioned above, the aerial image simulations based on rigorous electromagnetic field (Wong and Neureuther 1994; Erdmann et al. 2006, 2009; Nyssonen 1982; Lucas et al. 1996; Evanschitzky and Erdmann 2005; Schellenberg et al. 2005; Shao et al. 2008) are highly flexible, accurate, and relatively easy to deal with various problems in lithography simulation. However,

**Fig. 14** 3D simulation results of 30° inclined UV lithography for 240 μm thick SU-8 with a complex mask shape on glass wafer with TiO<sub>2</sub> film as antireflection layer: Exposure time = 380 s, development time = 13 min, the first step PEB condition was 65 °C/15 min, and the second step PEB condition was 95 °C/45 min (Zhou et al. 2014)



the computer calculation ability and technology limit their applications in UV lithography simulation of the SU-8. The usage of GPUs is also a possible solution to this problem by using massively parallel computing hardware. Finally, an important requirement for a lithography simulator is accuracy. More accurate and faster (if possible) physically based models for various lithography simulation steps are expected to be incorporated into the simulations, providing much better accuracy over a wider range of process conditions. At the same time, more efforts are expected for the measurement and optimization of the model parameters.

---

## Conclusions

In this chapter, the essential aspects in the modeling and simulation of UV lithography of the SU-8 are introduced. The presented various lithography models, such as the aerial image, exposure, PEB, and development models, are discussed and analyzed. Main algorithms for the etching surface advancement simulation, including the string, ray-tracing, cellular automaton, and fast-marching algorithms, are discussed with the advantages and challenges. Simulations of the UV lithography of the SU-8 and the corresponding experimental results are also presented. The simulation results demonstrate to be in agreement with the experimental results. This is useful to optimize the inclined UV lithography processes of SU-8 photoresists and to accurately design and control the dimensions of some MEMS microstructures. To extend the applications of the UV lithography simulations of the SU-8, several challenging issues are expected to be solved regarding the simulation speed and accuracy.

---

## References

- Adalsteinsson D, Sethian JA (1995) A level set approach to a unified model for etching, deposition, and lithography II: three-dimensional simulations. *J Comput Phys* 122:348–366

- Arnold JC, Sawin HH, Dalvie M et al (1994) Simulation of surface topography evolution during plasma etching by the method of characteristics. *J Vac Sci Technol A* 12:620–635
- Arthur G, Mack CA, Eilbeck N et al (1998) Analyzing the dissolution characteristics of deep UV chemically amplified photoresist. *Microelectron Eng* 41–42:311
- Balslev S, Romanato F (2005) Functionalized SU-8 patterned with x-ray lithography. *J Vac Sci Technol B* 23:2910–2918
- Becker EW, Ehrfeld W, Hagmann P et al (1996) Fabrication of microstructures with high aspect ratios and great structural heights by synchrotron radiation lithography, galvanofarming, and plastic moulding (LIGA process). *Microelectron Eng* 4:35–56
- Becnel C, Desta Y, Kelly K (2005) Ultra-deep x-ray lithography of densely packed SU-8 features: II. Process performance as a function of dose, feature height and post exposure bake temperature. *J Micromech Microeng* 15:1249–1259
- Berry AK, Graziano KA, Bogan LE et al (1989) Polymers in microlithography. In: Proceedings of ACS symposium series, vol 412. American Chemical Society, Washington, DC, pp 87–99
- Bobbitt MM (2001) The effect of chemistry and network structure on morphological and mechanical properties of diepoxide precursors and poly (hydroxyethers) [PH.D dissertation] Virginia Polytechnic Institute and State University
- Born M, Wolf E (1999) Principles of optics, 7th edn. Cambridge University Press, London (Chapters 8, 11 and 13)
- Bourdillon AJ, Boothroyd CB, Kong JR et al (2000) A critical condition in Fresnel diffraction used for ultra-high resolution lithographic printing. *J Phys D Appl Phys* 33:2133–2141
- Campo A, Greiner C (2007) SU-8: a photoresist for high-aspect-ratio and 3D submicron lithography. *J Micromech Microeng* 17:R81–R95
- Cheng Y, Lin CY, Wei DH et al (1999) Wall profile of thick photoresist generated via contact printing. *J Microelectromech Syst* 8:18–26
- Cheng YW, Chiang TH, Ngoc DL et al (2009) Fabrication of microlens arrays based on the mass transport effect of SU-8 photoresist using a multiexposure two-beam interference technique. *Appl Opt* 48:2473–2479
- Chopp DL (2001) Some improvements on the fast marching method. *SIAM J Sci Comp* 23:230–244
- Chuang YJ, Tseng FG, Lin WK (2002) Reduction of diffraction effect of UV exposure on SU-8 negative thick photoresist by air gap elimination. *Microsyst Technol* 8:308–303
- Cole DC, Barouch E, Conrad ED et al (2001) Using advanced simulation to aid microlithography development. *Proc IEEE* 89:1194–1113
- Colin DJ, Calame JP, Morag G et al (2010) UV-LIGA microfabrication of 220 GHz sheet beam amplifier gratings with SU-8 photoresists. *J Micromech Microeng* 20:125016
- Cooperberg DJ (2002) Semiempirical profile simulation of aluminum etching in a Cl<sub>2</sub>/BCl<sub>3</sub> plasma. *J Vac Sci Technol A* 20:1536–1556
- Dill FH, Neureuther AR, Tuttle JA (1975) Modeling projection printing of positive photoresists. *IEEE Trans Electron Devices* ED-22:456–464
- Erdmann A, Evanschitzky P, Citarella G et al (2006) Rigorous mask modeling using waveguide and FDTD methods: an assessment for typical hyper NA imaging problems. *Proc SPIE* 6283:628319
- Erdmann A, Fuhner T, Shao F et al (2009) Lithography simulation: modeling techniques and selected applications. *Proc SPIE* 7390:739002–739001
- Evanschitzky P, Erdmann A (2005) Three dimensional EUV simulations: a new mask near field and imaging simulation system. *Proc SPIE* 5992:1546
- Feng M, Hang QA, Zhou ZF et al (2006) Three-dimensional simulation of the deep UV light intensity distribution in SU-8 photoresists. *Proc ICSICT* 2006:664–666
- Ferrando N, Gosalvez MA, Cerda J et al (2011) Octree-based, GPU implementation of a continuous cellular automaton for the simulation of complex, evolving surfaces. *Comput Phys Comm* 182:628–640
- Fuhner T, Schnattinger T, Ardelean G et al (2007) Dr. LiTHO—a development and research lithography simulator. *Proc SPIE* 6520:65203F

- Gosálvez MA, Xing Y, Sato K (2008) Analytical solution of the continuous cellular automaton for anisotropic etching. *J Microelectromech Syst* 17:410–431
- Gremaud PA, Kuster CM (2006) Computational study of fast methods for the eikonal equation. *SIAM J Sci Comp* 27:1803–1814
- Hagouel PI (1976) X-ray lithographic fabrication of blazed diffraction gratings [Ph.D. Dissertation] University of California, Berkeley
- Hamaguchi S, Rossnagel SM (1995) Simulations of trench-filling profiles under ionized magnetron sputter metal deposition. *J Vac Sci Technol B* 13:183–191
- Han M, Lee W, Lee SK et al (2004) 3D microfabrication with inclined/rotated UV lithography. *Sens Actuators A* 111:14–20
- Hassouna MS, Farag AA (2007) Multistencils fast marching methods: a highly accurate solution to the eikonal equation on cartesian domains. *IEEE Trans Pat Ana Mach Int* 29:1563–1574
- Henderson CL, Pancholi SN, Chowdhury SA et al (1997) Photoresist characterization for lithography simulation part 2: exposure parameter measurements. *Proc SPIE* 3049:816–828
- Henke W, Weiss M, Schwalm R et al (1990) Simulation of proximity printing. *Microelectron Eng* 10:127–152
- Hirai Y, Inamoto Y, Sugano K et al (2007) Moving mask UV lithography for three-dimensional structuring. *J Micromech Microeng* 17:199–206
- Hoang J, Hsu CC, Chang JP (2008) Feature profile evolution during shallow trench isolation etch in chlorine-based plasmas I feature scale modeling. *J Vac Sci Technol B* 26:1911–1918
- Hsu CC, Marchack N, Martin R et al (2013) Feature profile evolution during shallow trench isolation etching in chlorine-based plasmas III the effect of oxygen addition. *J Vac Sci Technol* 31:042201
- Huang YT, Hsu WY (2014) A simulation model on photoresist SU-8 thickness after development under partial exposure with reflection effect. *J J Appl Phys* 53:036505
- Hung KY, Liang TH (2008) Application of inclined-exposure and thick film process for high aspect-ratio micro structures on polymer optic devices. *Microsyst Technol* 14:1217–1222
- Hung KY, Hu HT, Tseng FG (2004) Application of 3D glycerol-compensated inclined-exposure technology to an integrated optical pick-up head. *J Micromech Microeng* 14:975–983
- Jewett RE, Hagouel PI, Neureuther AR et al (1977) Line-profile resist development simulation techniques. *Polym Eng Sci* 17:381–384
- Jiang GM, Baig S, Wang MR (2012) Prism-assisted inclined UV lithography for 3D microstructure fabrication. *J Micromech Microeng* 22:085022
- Kang WJ, Rabe E, Kopetz S (2006) Exposure methods based on reflection and refraction effects in the field of SU-8 lithography. *J Micromech Microeng* 16:821–831
- Karafyllidis I (1997) Simulation of the negative chemical amplification deep-ultraviolet process in integrated circuit fabrication. *Microelectron Eng* 34:155–170
- Karafyllidis I (1999) A three-dimensional photoresist etching simulator for TCAD. *Modeling Simul Mater Sci Eng* 7:157–167
- Karafyllidis I, Thanailakis A (1995) Simulation of two-dimensional photoresist etching process in integrated circuit fabrication using cellular automata. *Modeling Simul Mater Sci Eng* 3:629–642
- Karafyllidis I, Hagouel PI, Neureuther AR et al (1999) Negative resist profiles of close-spaced parallel and isolated lines: experiment, modelling and simulation. *Microelectron Eng* 45:71–84
- Karafyllidis I, Haguel PI, Thanailakis A et al (2000) An efficient photoresist development simulator based on cellular automata with experimental verification. *IEEE Trans Semicond Manuf* 13:61–75
- Leunissen LHA, Jonckheere R, Ronse K, Derksen GB et al (2003) Influence of gate patterning on line edge roughness. *J Vac Sci Technol B* 21:3140–3143
- Levinson JA, Shaqfeh ESG, Balooch M et al (2000) Ion-assisted etching and profile development of silicon in molecular and atomic chlorine. *J Vac Sci Technol B* 18:172–190
- Liu SJ, Du JL, Duan X (2005) Enhanced dill exposure model for thick photoresist lithography. *Microelectron Eng* 78/79:490–495
- Lucas K, Tanabe H, Strojwas AJ (1996) Efficient and rigorous three-dimensional model for optical lithography simulation. *J Opt Soc Am A* 13:2187–2199

- Mack CA (1985) PROLISH: a comprehensive optical lithography model. *Proc SPIE* 538:207–220
- Mack CA (1992) New kinetic model for resist dissolution. *J Electrochem Soc* 139:L35
- Mack CA (2005) 30 years of lithography simulation. *Proc SPIE* 5754:1–12
- Mack A, Arthur G (1998) Notch model for photoresist dissolution. *Electrochem Solid-State Lett* 1:86–87
- Miao ZY (2013) Modeling and simulation of surface profile formation process of microlenses and their application in optical interconnection devices [PH.D Dissertation] Louisiana State University
- Moser Y, Forti R, Lehnert JS (2011) Suspended SU-8 structures for monolithic microfluidic channels. *Microfluid Nanofluid* 10:219–224
- Neumann J (1966) *Theory of self-reproducing automata*. University of Illinois Press, Urbana
- Neureuther AR, Mack CA (1997) *Handbook of microlithography, micromaching, and micro-fabrication*. SPIE Optical Engineering Press, Bellingham
- Nyyssonen D (1982) The theory of optical edge detection and imaging of thick layers. *J Opt Soc Am* 72:1425–1436
- Ong BH, Yuan X, Tao S et al (2006) Photothermally enabled lithography for refractive-index modulation in SU-8 photoresist. *Opt Lett* 31:1367–1369
- Osher S, Sethian JA (1988) Fronts propagating with curvature dependent speed: algorithms based on the Hamilton-Jacobi formulation. *J Comput Phys* 79:12–49
- Reichmanis E, Houlihan FM, Nalamasu O et al (1994) Polymers for microelectronics. In: proceedings of ACS symposium series, vol 537. American Chemical Society, Washington, DC, pp 3–24
- Rouabah HA, Park BY, Zaouk RB et al (2011) Design and fabrication of an ac-electro-osmosis micropump with 3D high-aspect-ratio electrodes using only SU-8. *J Micromech Microeng* 21:125016
- Rumpf RC, Johnson EG (2005) Comprehensive modeling of near-field nanopatterning. *Opt Express* 13:7198–7208
- Sato H, Yagyu D, Ito S et al (2006) Improved inclined multi-lithography using water as exposure medium and its 3D mixing microchannel application. *Sens Actuators A* 128:183–190
- Scheckler EW, Tam NN, Pfau AK et al (1993) An efficient volume-removal algorithm for practical three-dimensional lithography simulation with experimental verification. *IEEE Trans Comput-Aided Design Integr Circuit Syst* 12:1345–1356
- Schellenberg FM, Adam K, Matteo J et al (2005) Electromagnetic phenomena in advanced photomasks. *J Vac Sci Technol B* 23:3106–3115
- Sensu Y, Sekiguchi A, Mori S (2005) Profiles simulation of SU-8 thick film resist. *J Photopolym Sci Tech* 18:125–132
- Sethian JA (1995) A fast marching level set method for monotonically advancing fronts. *Proc Natl Acad Sci* 93:1591–1595
- Sethian JA (1996) Fast marching level-set methods for three-dimensional photolithography development. *Proc SPIE* 2726:262–272
- Sethian JA (1999) *Level sets methods and fast marching methods*, 2nd edn. Cambridge University Press, Cambridge, pp 75–99
- Sethian JA (2001) Evolution, implementation, and application of level set and fast marching methods for advancing fronts. *J Comput Phys* 169:503–555
- Shao F, Evanschitzky P, Reibold D et al (2008) Fast rigorous simulation of mask diffraction using the waveguide method with parallelized decomposition technique. *Proc SPIE* 6792:679206
- Shaw JM, Gelorme JD, LaBianca NC et al (1997) Negative photoresists for optical lithography. *IBM J Res Develop* 41:81–94
- Shi LL, Zhou ZF, Li WH et al (2011) A modified 3D fast marching simulation for thick photoresists lithography. In: *Proceedings of 10th IEEE sensors*, pp 1550–1553
- Sohn DS, Sohn YS, Bak HJ et al (2001) Analysis of the relation between exposure parameters and critical dimension by response surface model. *Proc SPIE* 4345:973–982



- Stout PJ, Rauf S, Nagy A et al (2006) Modeling dual inlaid feature construction. *J Vac Sci Technol B* 24:1344–1352
- Strasser E, Selberherr S (1995) Algorithms and models for cellular based topography simulation. *IEEE Trans Comput-Aided Design Integr Circuit Syst* 14:1104–1114
- Tang XG, Gao FH, Guo YK (2005) Analysis and simulation of diffractive imaging field in thick film photoresist by using angular spectrum theory. *Opt Commun* 244:123–130
- Tang XG, Yang XY, Gao FH et al (2007) Simulation and analysis for microstructure profile of optical lithography based on SU-8 thick resist. *Microelectron Eng* 84:1100–1103
- Tian X, Liu G, Tian Y et al (2005) Simulation of deep UV lithography with SU-8 resist by using 365 nm light source. *Microsyst Technol* 11:265–270
- Toh KKH, Neureuther AR, Scheckler EW (1994) Algorithms for simulation of three-dimensional etching. *IEEE Trans Comput-Aided Design Integr Circuit Syst* 13:616–624
- Vyvoda MA, Li M, Graves DB, Lee H et al (2000) Role of sidewall scattering in feature profile evolution during  $\text{Cl}_2$  and HBr plasma etching of silicon. *J Vac Sci Technol B* 18:820–833
- Weiss M, Binder H, Schwalm R (1995) Modeling and simulation of chemically amplified DUV resist using the effective acid concept. *Microelectron Eng* 27:405–409
- Wong AK, Neureuther AR (1994) Mask topography effects in projection printing of phase-shifting masks. *IEEE Trans on Electron Devices* 41:895–902
- Yang R, Wang WJ (2005) A numerical and experimental study on gap compensation and wavelength selection in UV-lithography of ultra-high aspect ratio SU-8 microstructures. *Sens Actuators B* 110:279–288
- Yang WC, Huang YS, Shew BY et al (2013) Study on diffraction effect and microstructure profile fabricated by one-step backside lithography. *J Micromech Microeng* 23:035004
- Yatziv L, Bartesaghi A, Sapiro G (2006)  $O(N)$  implementation of the fast marching algorithm. *J Comput Phys* 212:393–399
- Yoon YK, Park JH, Allen MG (2006) Multidirectional UV lithography for complex 3-D MEMS structures. *J Microelectromech Syst* 15:1121–1130
- Zanghellini J, Achenbach S, El-Kholi A et al (1998) New development strategies for high aspect ratio microstructures. *Microsyst Technol* 4:94–97
- Zhao HK (2005) A fast sweeping method for eikonal equations. *Math Comp* 74:603–627
- Zhou ZF, Huang QA, Li WH et al (2005) A novel 2-D dynamic cellular automata model for photoresist etching process simulation. *J Micromech Microeng* 15:652–662
- Zhou ZF, Huang QA, Li WH et al (2007a) Improvement of the 2D dynamic CA method for photoresist etching simulation and its application to deep UV lithography simulations of SU-8 photoresists. *J Micromech Microeng* 17:2538–2547
- Zhou ZF, Huang QA, Li WH et al (2007b) The swelling effects during the development processes of deep UV lithography of SU-8 photoresists: theoretical study, simulation and verification. In: *Proceedings of 6th IEEE sensors*, pp 325–328
- Zhou ZF, Huang QA, Li WH et al (2007c) A cellular automaton-based simulator for silicon anisotropic etching processes considering high index planes. *J Micromech Microeng* 17:S38–S49
- Zhou ZF, Huang QA, Li WH et al (2007d) A novel 3D dynamic cellular automata model for photoresist-etching process simulation. *IEEE Trans Comput-Aided Design Integr Circuit Syst* 26:100–114
- Zhou ZF, Huang QA, Li WH (2009) Modeling and simulations of anisotropic etching of silicon in alkaline solutions with experimental verification. *J Electrochem Soc* 156:F29–F37
- Zhou ZF, Huang QA, Li WH et al (2010) Simulations, analysis and characterization of the development profiles for the thick SU-8 UV lithography process. In: *Proceedings of 9th IEEE sensors*, pp 2525–2528
- Zhou ZF, Zhu Z, Huang QA et al (2011) An efficient simulation system for inclined UV lithography processes of thick SU-8 photoresists. *IEEE Trans Semicond Manuf* 24:294–303
- Zhou ZF, Shi LL, Zhang H et al (2014) Large scale three-dimensional simulations for thick SU-8 lithography process based on a full hash fast marching method. *Microelectron Eng*. 123:171–174

- Zhu ZR, Swecker AL, Strojwas AJ (2004) METRO-3D: an efficient three-dimensional wafer inspection simulator for next-generation lithography. *IEEE Trans Semicond Manuf* 17:619–628
- Zhu Z, Zhou ZF, Huang QA et al (2008) Modeling, simulation and experimental verification of inclined UV lithography for SU-8 negative thick photoresists. *J Micromech Microeng* 18:125017
- Zhu Z, Zhou ZF, Huang QA et al (2009) A 3D profile simulator for inclined/multi-directional UV lithography process of negative-tone thick photoresists. In: *Proceedings of 8th IEEE sensors*, pp 57–60
- Zuniga M, Wallraff G, Tomacruz E (1993) Simulation of locally enhanced three-dimensional diffusion in chemically amplified resists. *J Vac Sci Technol B* 11:2862–2866



# Modeling of Electrostatically Actuated Microplates

Libo Zhao, Zhuangde Jiang, Zhikang Li, and Yihe Zhao

## Contents

Introduction .....	100
Finite Element Modeling .....	102
Lumped Electromechanical Model .....	105
Distributed Electromechanical Models of Electrostatically Actuated Circular Microplates ...	109
Problem Formulation .....	109
Reduced-Order Model .....	110
Static Deflection Analysis .....	114
Pull-In Voltage Predication .....	115
Capacitance Calculation .....	116
Validation of Theoretical Analysis .....	117
Distributed Electromechanical Models of Electrostatically Actuated Rectangular Microplates .....	128
Problem Formulation .....	128
Reduced-Order Model .....	129
Static Deflection Analysis .....	131
Pull-In Voltage Prediction .....	132
Capacitance Calculation .....	133
Validation of Theories .....	133
Dynamic Behavior Analysis .....	142
Conclusions and Problems .....	150
References .....	151

L. Zhao · Z. Jiang (✉) · Z. Li · Y. Zhao

State Key Laboratory for Manufacturing Systems Engineering, Collaborative Innovation Center of Suzhou Nano Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China

International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, Xi'an Jiaotong University, Xi'an, Shaanxi, China

e-mail: [zdjiang@mail.xjtu.edu.cn](mailto:zdjiang@mail.xjtu.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies, [https://doi.org/10.1007/978-981-10-5945-2\\_4](https://doi.org/10.1007/978-981-10-5945-2_4)

---

**Abstract**

Electrostatically actuated microplates have been widely used in various microsensors and actuators actuated by electrostatic force. Deep knowledge of the microplates under electrostatic force and other physical quantities is extremely important for the design and optimization of the microsensors and actuators, which can largely reduce the cost and time to develop the proposed devices compared with over and over fabrication and testing in laboratory. This chapter gives a detailed illustration on the modeling methods of electrostatically actuated microplates. Three types of modeling methods are mainly discussed, that is, finite element modeling, lumped electromechanical modeling, and distributed electromechanical modeling. For the finite element modeling method, the electromechanical elements used to model the electrostatic domain and the establishment methods of finite element electromechanical models are given for electrostatically actuated microplates. The lumped electromechanical modeling method models the electrostatically actuated microplate as one-dimensional spring-mass-capacitor system, which can qualitatively analyze the collapse voltage (or pull-in voltage) and the reason why the resonant frequency shifts under electrostatic force. For the distributed electromechanical modeling method, the electromechanical coupling models for circular and rectangular microplates under electrostatic force and hydrostatic pressure are established, and explicit theoretical expressions for collapse voltage, static deflection, and capacitance variation are proposed. In addition, the distributed modeling method also focuses on the dynamic behavior analysis, especially the resonant frequency analysis of the electrostatically actuated microplate.

---

**Keywords**

Circular and rectangular microplates · Electrostatic actuation · Lumped electromechanical modeling · FEM modeling · Distributed electromechanical modeling · Collapse voltage · Static deflection · Capacitance · Resonant frequency · Theoretical expressions · Electrostatic force · Hydrostatic pressure

---

**Introduction**

Owing to the advance in the technology of microelectromechanical systems (MEMS), microsensors and actuators actuated by electrostatic force become the object of intensive investigation for their merits of small size, light weight, rapid response, lower power consumption, batch fabrication, and integrated circuit standard process compatibility. These microsensors and actuators are commonly seen in various structures of pressure sensors, capacitive microphones, micropumps, micromirrors, microswitches, capacitive micromachined ultrasonic transducers (CMUTs), vacuum resonators, and biochemical sensors/actuators. Applications of these electrostatically actuated microdevices spread over biotechnology, image processing, industrial control, nondestructive evaluation, and automotive, chemical, and food industries.

Electrostatically actuated microplates are key components of the aforementioned electrostatic microdevices. Generally, two parallel microplates are simultaneously employed to provide an electrostatic actuation configuration for the microdevice. One of them is a flexible conductor and capable of motion or deformation, while the other one is a stationary rigid conductor and remains fixed under the action of electrostatic force. The gap between two microplates is usually air or a vacuum. When a DC voltage is applied across two microplates, an electrostatic force is produced and forces the flexible microplate to deflect toward the fixed one. The deflection of microplates will induce the capacitance change of parallel-microplate structures. If the applied voltage increases up to a critical value beyond which the mechanical restoring force of microplate can no longer balance the electrostatic force, the flexible microplate will collapse and make contact with the rigid microplate. This phenomenon is called “pull-in,” and the corresponding critical voltage is called the “pull-in voltage.” If an AC voltage excitation is further superimposed on the DC bias voltage, the parallel-plate structure works as an electrostatically actuated resonator, while the DC voltage can change the resonant frequency of the electrostatically actuated microplates. These deflections, capacitances, pull-in voltages, and resonant frequency shifts are key issues for electrostatically actuated microplates and are intensively related to the design and optimization of electrostatic microdevices.

Modeling of electrostatically actuated microplates is greatly necessary for the widespread applications of electrostatic microdevices, which would be more effective to design, tune, and verify the performance of proposed devices than over and over fabrication, testing, and modification in the laboratory. Accurate determination of the deflection is crucial to analyzing the performance of electrostatic device. For example, for CMUTs, this deflection determines the capacitance change and thus the electromechanical coupling coefficient which characterizes the transducer bandwidth and sensitivity (Yaralioglu et al. 2003; Rahman and Chowdhury 2011). For capacitive pressure sensors, the deflection of flexible microplate as a result of both the electrostatic force and the pressure determines the capacitance variation and thus the sensor sensitivity and nonlinearity (Fragiacomo et al. 2010; Zhou et al. 2005). The pull-in voltage is extremely important for the operation of microdevice. For example, the maximum DC bias voltage applied to the CMUTs is dependent on the pull-in voltage (Ahmad and Pratap 2010). The capacitive pressure sensors rely on the pull-in voltage to determine their working modes (normal or touch mode) (Zhou et al. 2005). MEMS switches need to tune the bias voltage across the pull-in voltage back and forth to alternate switch on and off.

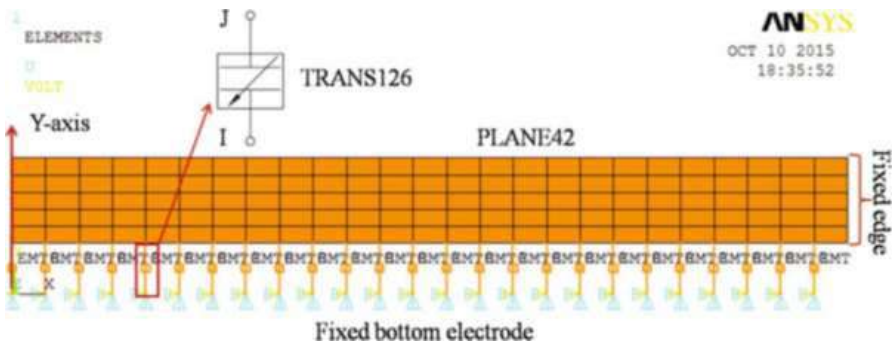
This chapter illustrates the modeling methods of electrostatically actuated microplates, through which their static and dynamic behavior, especially the pull-in voltage, deflection, and resonant frequency shift, are studied. The content of the chapter is arranged as follows: sections “[Finite Element Modeling](#),” “[Lumped Electromechanical Model](#),” “[Distributed Electromechanical Modeling of Electrostatically Actuated Circular Microplates](#),” “[Distributed Electromechanical Modeling of Electrostatically Actuated Rectangular Microplates](#),” “[Dynamic Behavior Analysis](#),” and “[Conclusions and Problems](#).”

## Finite Element Modeling

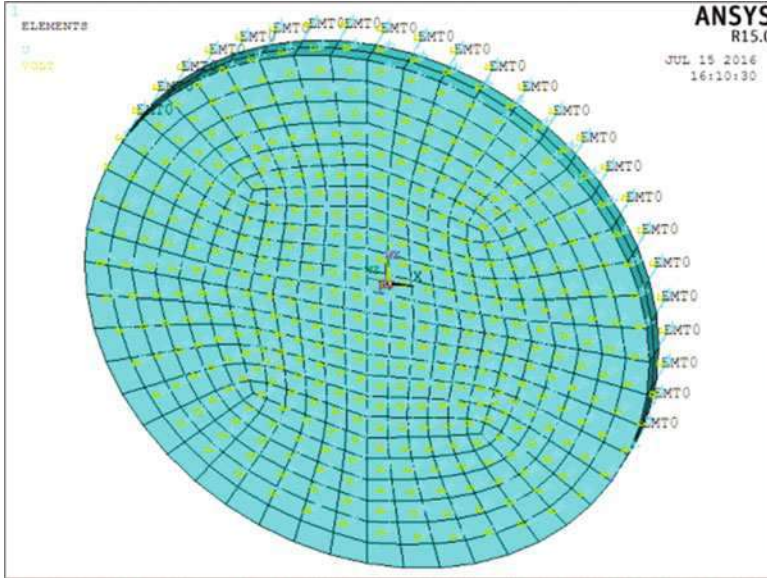
Finite element method (FEM) is widely used to model electrostatically actuated microplates and analyze their static and dynamic behavior. ANSYS and COMSOL are the most used commercially available FEM packages for the finite element modeling. The finite element models for the electrostatically actuated microplates include the electromechanical coupling between the electric and mechanical fields. The key problem in modeling is how to model the electromechanical coupling effects between the top and bottom electrodes. Generally, two methods are used: direct coupling and indirect coupling.

For the direct coupling modeling method, the direct electromechanical coupling elements, TRANS126, are mostly used to model the electrical field of electrostatically actuated microplates. The electromechanical coupling elements can apply electrostatic attraction forces to the nodes to which they are attached. Figure 1 shows an axisymmetric 2D model for an electrostatically actuated circular microplate. As shown in Fig. 1, J and I nodes of TRANS126 element are referred to as its top and bottom nodes, respectively. The top nodes of these elements are attached to the lower surface of the top microplate, and their bottom nodes are simply fixed to represent the bottom electrode. The potentials of bottom nodes of TRANS126 element are set to be zero, and different potentials are applied to the top nodes to simulate the electrostatic force. Generally, 2D plane elements such as PLANE42, PLANE82, and PLANE182 and 3D structural elements such as SOLID45, SOLID185, and SOLID186 are used to model the top microplates. Among them, the 2D elements are more often used to establish an axisymmetric 2D model for the circular microplates, an example of which can be seen in Fig. 1.

For the indirect coupling modeling method, 2D plane electrostatic elements such as PLANE121 and PLANE223 and 3D electrostatic solid elements such as SOLID122 and SOLID123 are used to model the electrostatic domain between the top and bottom electrodes. They are charge-based electric elements and have one degree of freedom, voltage. These electrostatic elements often are jointly used with



**Fig. 1** An axisymmetric 2D model for an electrostatically actuated microplate. Y-axis is the symmetry axis of the 2D model



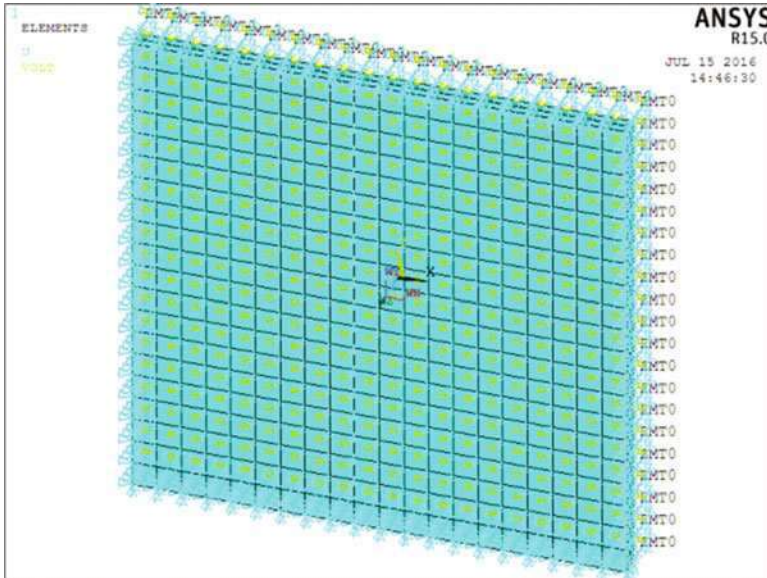
**Fig. 2** 3D finite element model for an electrostatically actuated circular microplate

other command such as CMATRIX to calculate the electrostatic force for the mechanical domain input. The elements used for the microplates are the same as those in the direct coupling modeling method. Some examples for indirect coupling modeling for the microplates can be seen in Ahmad and Pratap (2010) and Chiou et al. (2008).

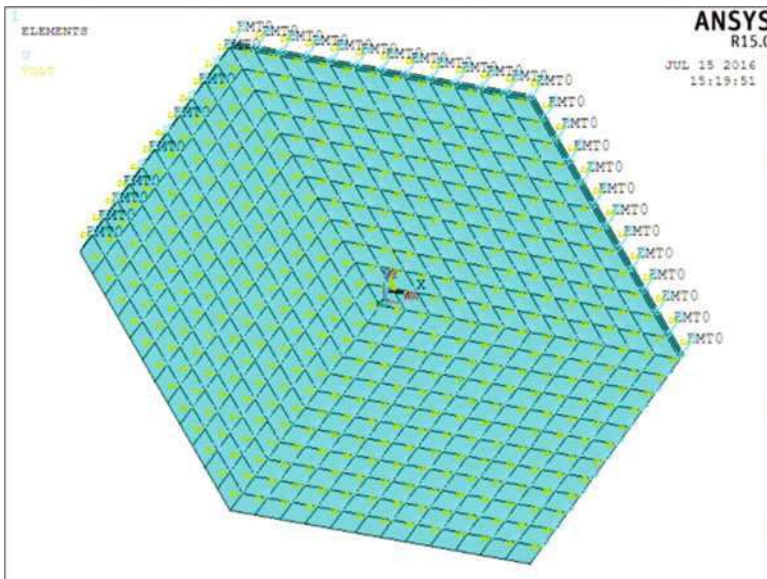
In general, three shapes of microplates are mostly used in electrostatically actuated microdevices, that is, circular, rectangular, and hexagonal microplates. The 3D models of circular, rectangular, and hexagonal electrostatically actuated microplates can be seen in Figs. 2, 3, and 4, respectively. In these models, 3D solid elements are used to model the microplates, and direct electromechanical coupling elements are used to model the electrostatic effects. However, the electrostatically actuated circular microplates are more often modeled using an axisymmetric 2D model as shown in Fig. 1 to reduce computation time, which can give identical results with those obtained by 3D model (Ahmad and Pratap 2010; Chiou et al. 2008).

These finite element models can give static analysis, prestressed modal analysis, harmonic analysis, and transient analysis. The microplate deflection, deformation shape, pull-in voltage, capacitance, displacement-frequency curve, and so on can be obtained from the aforementioned analyses.

Although FEM simulations are adequate for the mechanical behavior analyses of electrostatically actuated microdevices, they are computationally inefficient relative to non-FEA simulations or breakdown near pull-in of electrostatically actuated microplates. As a result, theoretical modeling is gaining attention as a way to make the design and optimization of microdevice more efficient (Nayfeh et al. 2005).



**Fig. 3** 3D finite element model for an electrostatically actuated rectangular microplate



**Fig. 4** 3D finite element model for an electrostatically actuated hexagonal microplate



### Lumped Electromechanical Model

Figure 5 shows the lumped electromechanical model for electrostatically actuated microplates, which can be used to study electrostatically actuated configurations qualitatively. Several simplifications are made in the model establishment: the restoring force of microplate is assumed to be a linear function of its displacement; the electrical fringing fields and microplate curvature under electrostatic forces are neglected. Furthermore, all conductors and contacts are assumed to be perfect. The electrostatic microplate is assumed to operate in a vacuum, which is equivalent to neglecting any loading applied to the membrane. On this foundation, the microplate-based electrostatically actuated configuration is simplified as the lumped model consisting of a spring, a mass, and a parallel capacitor.

The mass is moved under the co-action of electrostatic force and mechanical restoring force. When the static behavior of model is considered, in which the time dependence of two types of forces is neglected, the join force applied to the mass is given as

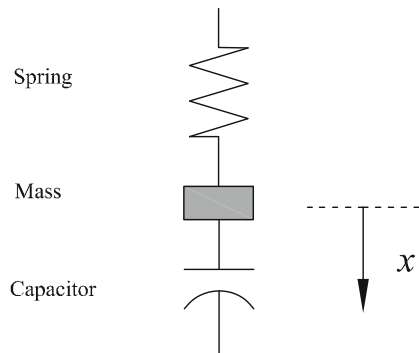
$$F_N = F_s - F_e \tag{1}$$

where the electrostatic force is derived by differentiating the potential energy of the capacitor with respect to the position of the mass (principle of virtual work):

$$F_e = -\frac{d}{dx} \left( \frac{1}{2} C V_{dc}^2 \right) = -\frac{1}{2} V_{dc}^2 \frac{d}{dx} \left( \frac{\epsilon_0 S}{d_0 - x} \right) = \frac{\epsilon_0 S V_{dc}^2}{2(d_0 - x)^2} \tag{2}$$

where  $V_{dc}$  is the bias voltage across the capacitor,  $C$  is capacitance,  $\epsilon_0$  is the electric permittivity,  $S$  is the area of the parallel microplates,  $x$  is the displacement in the direction shown in Fig. 5, and  $d_0$  is the initial electrode distance of the capacitor. The restoring force produced by the spring that is linearly proportional to displacement is given as

**Fig. 5** Lumped electromechanical coupling model of electrostatically actuated microplates



$$F_s = kx \quad (3)$$

where  $k$  is the spring constant.

The mass reaches a stable situation if the electrostatic force induced by the bias voltage is balanced by the mechanical restoring force of a microplate. At the stable situation, the electrostatic force and mechanical restoring forces are equal to each other, that is,

$$kx = \frac{\varepsilon_0 S V_{dc}^2}{2(d_0 - x)^2} \quad (4)$$

Using Eq. (1), the spring constant  $k_s$  of the system at a given bias voltage is given as

$$k_s = \frac{dF_N}{dx} = k - \frac{\varepsilon_0 S V_{dc}^2}{(d_0 - x)^3} \quad (5)$$

Equation (5) shows that the spring constant of the system decreases with the displacement  $x$  and the bias voltage  $V_{dc}$ . When the bias voltage and displacement increase to certain values, the spring constant of the system decreases to zero; then the microplate cannot bear any loads. The mass reaches a critical stable situation between stable and unstable state, at which any small loads can induce the collapse of the microplates. Setting  $k_s = 0$ , the spring constant of the spring can be derived as

$$k = \frac{\varepsilon_0 S V_{dc}^2}{(d_0 - x)^3} \quad (6)$$

In Eq. (6), the bias voltage is equal to the collapse voltage of the system.

The electrostatic force and the mechanical restoring force satisfy the equilibrium given by Eq. (4) at the critical situation. Substituting Eq. (6) into Eq. (4), the critical displacement  $x_{pi}$  of the mass corresponding to the critical stable situation can be derived as

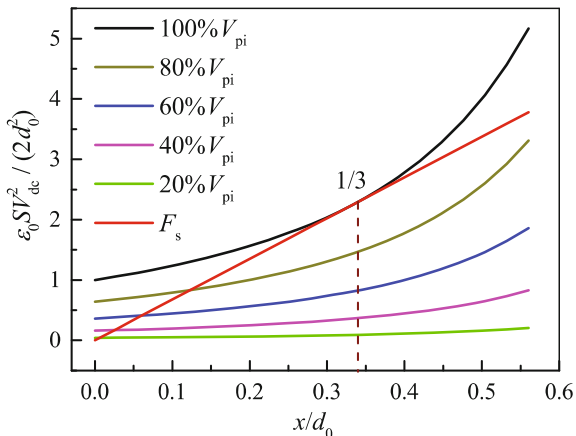
$$x_{pi} = \frac{1}{3} d_0 \quad (7)$$

Substituting Eq. (7) into Eq. (6), the corresponding collapse voltage can be obtained:

$$V_{pi} = \sqrt{\frac{8kd_0^3}{27\varepsilon_0 S}} \quad (8)$$

The spring constant of the spring can also be derived from Eq. (8) as

**Fig. 6** Variations of electrostatic force and mechanical restoring force with the displacement of mass under different bias voltages



$$k = \frac{27\epsilon_0 V_{pi} S}{8k d_0^3} \quad (9)$$

Substituting Eq. (9) into Eq. (3), the mechanical restoring force of the spring can be expressed as

$$F_s = \frac{27\epsilon_0 V_{pi} S}{8k d_0^3} x \quad (10)$$

Using Eqs. (2) and (10) to calculate the variations of electrostatic force and restoring force with displacements under different bias voltages can give an insight into the interaction of electromechanical coupling and collapse phenomenon. Figure 6 shows the comparison of electrostatic and mechanical restoring force variations. As shown in the Fig. 6, the gradient of the electrostatic force increases with the bias voltages, while the gradient of the spring mechanical restoring force is a constant. The curve of electrostatic force is tangent with that of the spring mechanical restoring at a certain point when the applied bias voltage is equal to the collapse voltage. At the tangent point, the electrostatic force is equal to the mechanical restoring force, and their gradients are also equal to each other. The mechanical restoring of the spring just can bear the electrostatic force. When the bias voltage is smaller than the collapse voltage, the curve of electrostatic force has one intersection point with that of mechanical restoring force. At the point of intersection, the gradient of electrostatic force is smaller than that of mechanical restoring force, which means the incremental increase in the mechanical restoring force overcomes the incremental increase in electrostatic force. When the bias voltage is larger than the collapse voltage, the curve of electrostatic force has no intersection point with that of mechanical restoring force. At this situation, the incremental increase in the electrostatic force overcomes the incremental increase in mechanical restoring force, and the collapse phenomenon occurs.

When the small signal AC voltage is superimposed on the bias voltage, the mass vibrates under the coaction of the electrostatic force, which can be expressed as

$$F_s + F_e = F_m \quad (11)$$

where the electrostatic force,  $F_e$ , and mechanical restoring force,  $F_s$ , are the time-dependent variables. Substituting Eqs. (2) and (3) into Eq. (11) and noting all time dependency give

$$m \frac{d^2x(t)}{dt} - \frac{\varepsilon_0 S [V(t)]^2}{2[d_0 - x(t)]^2} + kx(t) = 0 \quad (12)$$

Equation (12) is a nonlinear differential equation of second order, and its solution is not trivial.

In addition to membrane collapse, electrostatic spring softening also is observed experimentally. Such softening arises from the fact that, as the capacitor plate displaces in the  $+x$  direction, a spring force is generated in the  $-x$  direction. However, displacement in the  $+x$  direction under constant  $V_{dc}$  also causes an increase in the electrostatic force in the  $+x$  direction. This increase in electrostatic force can be interpreted as a spring softening. A more mathematical explanation begins by linearizing Eq. (12) with a Taylor expansion about the point  $x(t) = 0$ :

$$m \frac{d^2x(t)}{dt} - \left[ \frac{\varepsilon_0 S V_{dc}^2}{2d_0^2} + \frac{\varepsilon_0 S V_{dc}^2}{d_0^3} x(t) \right] + kx(t) = 0 \quad (13)$$

Equation (13) can be rewritten as

$$m \frac{d^2x(t)}{dt} + \left[ k - \frac{\varepsilon_0 S V_{dc}^2}{d_0^3} \right] x(t) = \frac{\varepsilon_0 S V_{dc}^2}{2d_0^2} \quad (14)$$

Equation (14) can be further written as

$$m \frac{d^2x(t)}{dt} + k_s x(t) = \frac{\varepsilon_0 S V_{dc}^2}{2d_0^2} \quad (15)$$

Thus, we expect to see a drop in the resonance frequency of the system as the bias voltage  $V_{dc}$  is increased.

The collapse voltage and resonant frequency drop can be qualitatively analyzed using the lumped electromechanical model. As the model does not take the membrane deflection into account, distributed electromechanical models are necessary for correct analyses of the collapse voltage, collapse point, microplate deflections, and resonant frequency shifts.

## Distributed Electromechanical Models of Electrostatically Actuated Circular Microplates

### Problem Formulation

Figure 7 shows the schematic of an electrostatically actuated configuration for theoretical analysis. The upper flexible microplate has a radius  $R$  and thickness  $h$  and is clamped at its edge. It is assumed to be thin, homogenous, and isotropic. The lower bottom microplate is parallel fixed below the upper one with an effective electrode distance  $d_0$ . The space between the upper and lower microplates is assumed as air or a vacuum with a dielectric permittivity  $\epsilon_0$ . When subjected to an electrostatic force or a uniform hydrostatic pressure, the upper microplate is deflected toward the fixed bottom microplate. Kirchhoff plate theory is used for the deflection analysis, which can provide accurate results when the ratio of microplate thickness to its diameter is very small ( $<1/20$ ) (Nabian et al. 2008). It should be noted that the plate referred to in the following sections is the upper flexible microplate unless stated otherwise.

For a circular plate with symmetrically distributed loads, the transverse deflection  $w$  is symmetrical. The general equation that governs the symmetrical bending is given elsewhere (Timoshenko and Woinowsky-Krieger 1959). For the clamped circular plate subjected to both axisymmetrical electrostatic force and hydrostatic pressure, the deflection governing equation can be expressed as

$$D \left( \frac{d^4 w}{dr^4} + \frac{2}{r} \frac{d^3 w}{dr^3} - \frac{1}{r^2} \frac{d^2 w}{dr^2} + \frac{1}{r^3} \frac{dw}{dr} \right) = F_e(r) + P \tag{16}$$

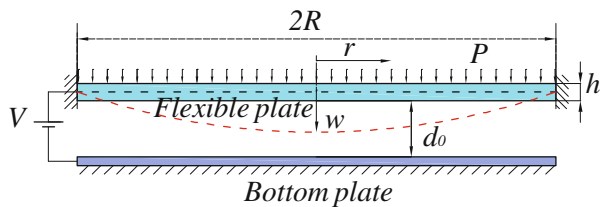
where

$$\nabla^4 w = \frac{d^4 w}{dr^4} + \frac{2}{r} \frac{d^3 w}{dr^3} - \frac{1}{r^2} \frac{d^2 w}{dr^2} + \frac{1}{r^3} \frac{dw}{dr}$$

$0 \leq r \leq R$ ,  $w(r)$  is the transverse deflection,  $D$  is the flexural rigidity,  $F_e(r)$  is the electrostatic force per unit area, and  $P$  is the applied pressure from sources such as air or water. The flexural rigidity  $D$  is defined as

$$D = Eh^3 / [12(1 - \nu^2)] \tag{17}$$

**Fig. 7** Schematic of a clamped circular microplate under nonuniform electrostatic force and uniform hydrostatic pressure



where  $E$  is the Young's modulus and  $\nu$  is the Poisson's ratio. The electrostatic force per unit area  $F_e(r)$  in (7) is given by

$$F_e(r) = \frac{\varepsilon_0 V^2}{2[d_0 - w(r)]^2} \quad (18)$$

where  $V$  is assumed as the applied DC voltage across the upper and lower plates without considering its time dependence (Nabian et al. 2008).

The boundary conditions for the clamped circular plate with axisymmetrical deformation are (Timoshenko and Woinowsky-Krieger 1959)

$$w(R) = 0, \quad dw(R)/dr = 0, \quad dw(0)/dr = 0 \quad (19)$$

## Reduced-Order Model

Due to the nonlinear electrostatic force in the deflection governing equation, it is hard to establish an accurate analytical solution. So many researchers searched for an approximation one (Ahmad and Pratap 2010; Chao et al. 2006; Younis et al. 2003; Vogl and Nayfeh 2005; Zhao et al. 2004; Liao et al. 2009). A common idea in their works was to use the well-known Galerkin method to decompose the deflection governing equation into a coupled set of ordinary equations (the reduced-order model), through which the approximate solution can be obtained.

To this end, an initial assumption was made about the function used to approximate the transverse deflection  $w$ . For a circular plate with axisymmetrical deflection, it can be expressed as

$$w(r) = \sum_{m=1}^N \eta_m \phi_m(r), \quad (20)$$

where  $\eta_m$  are the unknown coefficients to be determined,  $\phi_m(r)$  are the linearly independent coordinate functions (they are also called trial functions) that satisfy all the prescribed boundary conditions (Ventsel and Krauthammer 2001), and  $N$  is the number of the chosen trial functions.

Subsequently, there are two methods to generate a reduced-order model for Eq. (16) according to the idea mentioned above. In the first method (Ahmad and Pratap 2010; Chao et al. 2006; Younis et al. 2003; Liao et al. 2009), the electrostatic force is expanded using a Taylor series around  $w = 0$ . Thus, Eq. (18) can be written as

$$F_e(r) = \frac{\varepsilon_0 V^2}{2d_0^2} \left( 1 + 2\frac{w}{d_0} + 3\frac{w^2}{d_0^2} + 4\frac{w^3}{d_0^3} + 5\frac{w^4}{d_0^4} + \dots \right). \quad (21)$$

Substituting Eqs. (20) and (21) into Eq. (16), multiplying both sides of the resulting equation by  $\phi_n(r)$ , and integrating the outcome from  $r = 0$  to  $R$ , a series of ordinary equations can be obtained as

$$\begin{aligned}
 \int_0^R D\nabla^4 w \times \phi_n dr &= \frac{\varepsilon_0 V^2}{2d_0^2} \left( \int_0^R \phi_n dr + \frac{2}{d_0} \sum_{m=1}^N \eta_m \int_0^R \phi_n \phi_m dr \right. \\
 &+ \frac{3}{d_0^2} \sum_{m,i=1}^N \eta_m \eta_i \int_0^R \phi_n \phi_m \phi_i dr \\
 &+ \frac{4}{d_0^3} \sum_{m,i,j=1}^N \eta_m \eta_i \eta_j \int_0^R \phi_n \phi_m \phi_i \phi_j dr \\
 &+ \left. \frac{5}{d_0^4} \sum_{m,i,j,k=1}^N \eta_m \eta_i \eta_j \eta_k \int_0^R \phi_n \phi_m \phi_i \phi_j \phi_k dr + \dots \right) \\
 &+ P \int_0^R \phi_n dr \\
 &\text{for } n = 1, 2, 3, \dots, N,
 \end{aligned} \tag{22}$$

In the second method (Vogl and Nayfeh 2005; Zhao et al. 2004), Eq. (18) is directly used to establish the reduced-order model. Both sides of Eq. (16) are multiplied by  $(d_0 - w)^2$  to represent the electrostatic force term exactly. Then, substituting Eq. (20) into the result, multiplying both sides of the resulting equation by  $\phi_n(r)$ , and integrating the outcome from  $r = 0$  to  $R$ , another series of ordinary equations can be obtained as

$$\begin{aligned}
 d_0^2 \int_0^R D\nabla^4 w \times \phi_n dr &- 2d_0 \sum_{m=1}^N \eta_m \int_0^R D\nabla^4 w \times \phi_n \phi_m dr \\
 &+ \sum_{m,i=1}^N \eta_m \eta_i \int_0^R D\nabla^4 w \times \phi_n \phi_m \phi_i dr \\
 &= \frac{\varepsilon_0 V^2}{2} \int_0^R \phi_n dr + P d_0^2 \int_0^R \phi_n dr \\
 &- 2P d_0 \sum_{m=1}^N \eta_m \int_0^R \phi_n \phi_m dr \\
 &+ P \sum_{m,i=1}^N \eta_m \eta_i \int_0^R \phi_n \phi_m \phi_i dr \\
 &\text{for } n = 1, 2, 3, \dots, N.
 \end{aligned} \tag{23}$$

The integrals in Eqs. (22) and (23) can be evaluated once the trial functions  $\phi_m(r)$  in Eq. (20) and the number  $N$  are chosen. Then, once all  $\eta_m$  are determined by solving Eq. (22) or Eq. (23), an approximate solution to Eq. (16) can be established by Eq. (20).

However, both methods mentioned above have inherent flaws for establishing explicit expressions which can be easily used to design and operate microdevices based on electrostatic actuation. For the first method, the complexity of Eq. (22) and the accuracy of its solution depend on the order of the electrostatic force expansion. The first and second Taylor series expansions have relatively simple forms; however, they yield large errors with respect to the electrostatic force term shown in Eq. (18) (Liao et al. 2009). Though the higher-order expansions can reduce the error, the degree of the term  $\eta_m$  in Eq. (22) will be three or higher, which will greatly increase the complexity of establishing an explicit solution. For the second method, each of Eq. (23) is a standard cubic equation of  $\eta_m$  which makes it complicated to establish an explicit and simple analytical solution to the equation system, especially when the number  $N$  of trial functions increases. So it is necessary to establish an improved reduced-order model to reduce the degree of  $\eta_m$  and simultaneously obtain a solution with sufficient accuracy for engineering applications.

To this end, we propose a new method to treat the electrostatic force term. In this method, we expand one of two  $(d_0 - w)^{-1}$  terms as a Taylor series around  $w = 0$ . Thus, Eq. (18) can be expressed as

$$F_e(r) = \frac{\varepsilon_0 V^2}{2d_0(d_0 - w)} \left( 1 + \frac{w}{d_0} + \frac{w^2}{d_0^2} + \frac{w^3}{d_0^3} + \frac{w^4}{d_0^4} + \dots \right) \quad (24)$$

Substituting Eq. (24) into Eq. (16), multiplying both sides of the resulting equation by  $(d_0 - w)$ , substituting Eq. (20) into the result, then multiplying both sides of the resulting equation by  $\phi_n(r)$ , and finally integrating the outcome from  $r = 0$  to  $R$ , a reduced-order equation system for Eq. (16) is obtained as (Li et al. 2015a)

$$\begin{aligned} & d_0 \int_0^R D\nabla^4 w \times \phi_n dr - \sum_{m=1}^N \eta_m \int_0^R D\nabla^4 w \times \phi_n \phi_m dr \\ &= \frac{\varepsilon_0 V^2}{2d_0} \left( \int_0^R \phi_n dr + \frac{1}{d_0} \sum_{m=1}^N \eta_m \int_0^R \phi_n \phi_m dr \right. \\ &+ \frac{1}{d_0^2} \sum_{m,i=1}^N \eta_m \eta_i \int_0^R \phi_n \phi_m \phi_i dr \\ &+ \frac{1}{d_0^3} \sum_{m,i,j=1}^N \eta_m \eta_i \eta_j \int_0^R \phi_n \phi_m \phi_i \phi_j dr \\ &+ \left. \frac{1}{d_0^4} \sum_{m,i,j,k=1}^N \eta_m \eta_i \eta_j \eta_k \int_0^R \phi_n \phi_m \phi_i \phi_j \phi_k dr + L \right) \\ &+ P d_0 \int_0^R \phi_n dr - P \sum_{m=1}^N \eta_m \int_0^R \phi_n \phi_m dr \end{aligned} \quad (25)$$

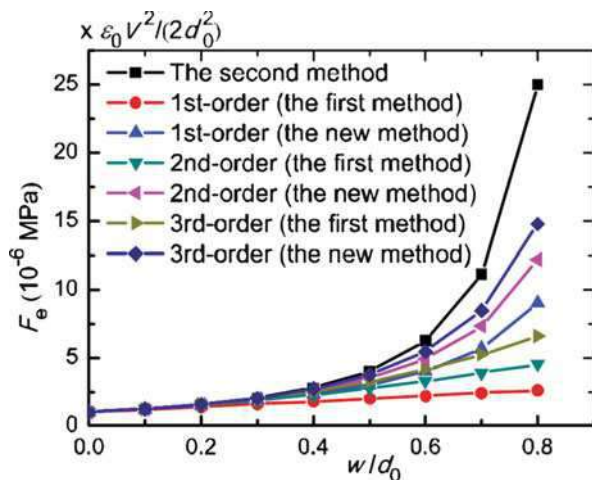
for  $n = 1, 2, 3, LN$ .



Similarly, Eq. (25) can be solved once the trial functions  $\phi_m(r)$  and the number  $N$  in Eq. (20) are determined. An approximate solution can be established by Eq. (20) after all  $\eta_m$  are calculated.

Equation (25) shows following advantages in establishing an approximate solution, especially an explicit expression. Firstly, for the same order Taylor series expansion of the electrostatic force term, the new method shows a higher accuracy than the first method. As can be seen in Fig. 8, a simple comparison of electrostatic forces was performed. The first-, second-, and third-order electrostatic force expansions in Eq. (24) are much closer to the electrostatic force calculated by Eq. (18) than the same order expansions in Eq. (21). Figure 8 also shows that the first-order electrostatic force expansion using the new method exhibits a better approximation than the second-order electrostatic force expansion using the first method. That is because Eq. (24) has a similar term  $(d_0-w)^{-1}$  to Eq. (18) and thus can better capture the variation characteristic of Eq. (18) than Eq. (21). This advantage suggests that a lower-order Taylor series expansion in Eq. (25) can reach the accuracy of a higher-order one in Eq. (22). Equation (25) can reduce the degree of term  $\eta_m$  without any reduction in the accuracy, which can greatly simplify the calculation process. Secondly, according to the order of the electrostatic force expansion used in Eq. (25), the highest degree of  $\eta_m$  in Eq. (25) can be set as low as two, which is less than the degree of  $\eta_m$  in Eq. (23) (three). This advantage leads to an explicit and simple solution. In addition, although the second method is more accurate, it results in much more complicated calculation than the first method (Younis et al. 2003). So the new method acts as a compromise between accuracy and complexity for the models given by the first and second methods.

**Fig. 8** Variation of the electrostatic force with respect to the ratio of deflection  $w$  to distance  $d_0$ . The electrostatic forces were calculated using Eq. (21) of the first method, Eq. (18) of the second method, and Eq. (24) of the new method



## Static Deflection Analysis

For a clamped circular plate with axisymmetric transverse deflection, the trial functions can be chosen as (Timoshenko and Woinowsky-Krieger 1959; Ventsel and Krauthammer 2001)

$$\phi_m = (R^2 - r^2)^{m+1}, \text{ for } m = 1, 2, 3 \dots N \quad (26)$$

These trial functions satisfy all the boundary conditions shown in Eq. (19). As this study aimed to establish an explicit expression to predict the deflection and pull-in voltage of the electrostatically actuated microplate, the number  $N = 1$  is used. Substituting Eq. (26) into Eq. (20), it can be rewritten as

$$w(r) = \eta_1 \phi_1 = \eta_1 (R^2 - r^2)^2 \quad (27)$$

This function is often used and is sufficient to approximate the axisymmetric deflection of circular plates (Timoshenko and Woinowsky-Krieger 1959; Ventsel and Krauthammer 2001). Substituting Eqs. (26) and (27) into Eq. (25) and using the first-order electrostatic force expansion, Eq. (25) can be simplified as

$$\begin{aligned} d_0 \int_0^R D \nabla^4 w \times (R^2 - r^2)^2 dr - \eta_1 \int_0^R D \nabla^4 w \times (R^2 - r^2)^4 dr \\ = Q \left( \int_0^R (R^2 - r^2)^2 dr + \frac{\eta_1}{d_0} \int_0^R (R^2 - r^2)^4 dr \right) \\ + d_0 P \int_0^R (R^2 - r^2)^2 dr - \eta_1 P \int_0^R (R^2 - r^2)^4 dr \end{aligned} \quad (28)$$

where  $Q = \varepsilon_0 V^2 / 2d_0$ . Evaluating the integrals and then solving the resulting equation, two possible solutions to the constant  $\eta_1$  are given by

$$\eta_{1a,b} = (84Dd_0^2 + d_0R^4P - QR^4 \pm B) / (128Dd_0R^4) \quad (29)$$

where

$$B = \sqrt{(84Dd_0^2 - R^4Q - d_0R^4P)^2 - 336Dd_0^2R^4Q} \quad (30)$$

As can be derived from Eq. (27), the term  $\eta_1$  actually determines the maximum deflection of the microplate, which increases with the electrostatic force and hydrostatic pressure (Nabian et al. 2008). So the correct value is considered to be

$$\eta_{1b} = (84Dd_0^2 + d_0R^4P - QR^4 - B) / (128Dd_0R^4) \quad (31)$$

Thus, replacing  $\eta_1$  in Eq. (27) with  $\eta_{1b}$ , the deflection  $w_{V,P}$  of the clamped circular microplate under both electrostatic force and uniform hydrostatic pressure is obtained as

$$w_{V,P}(r) = \frac{84Dd_0^2 + d_0R^4P - QR^4 - B}{128Dd_0R^4} \times (R^2 - r^2)^2 \quad (32)$$

When the electrostatic force equals zero, Eq. (32) can be simplified to

$$w_P(r) = P(R^2 - r^2)^2 / (64D) \quad (33)$$

Equation (33) is the well-known expression for the deflection of the plate under uniform hydrostatic pressure only (Timoshenko and Woinowsky-Krieger 1959). If only the effect of the electrostatic force on the deflection is considered, Eq. (32) can be simplified to

$$w_V(r) = \frac{84Dd_0^2 - QR^4 - B_1}{128Dd_0R^4} \times (R^2 - r^2)^2 \quad (34)$$

where the term  $w_V$  in Eq. (34) denotes the deflection of the microplate subjected only to the electrostatic force and  $B_1$  is given by

$$B_1 = \sqrt{7056D^2d_0^4 - 504Dd_0^2R^4Q + R^8Q^2} \quad (35)$$

## Pull-In Voltage Predication

The solution for the pull-in voltage of the clamped upper microplate can be obtained from Eq. (29). It can be shown that there is a voltage at which the value of  $\eta_{1a}$  equals that of  $\eta_{1b}$ . This voltage is the critical point where the pull-in phenomenon occurs (Ahmad and Pratap 2010; Vogl and Nayfeh 2005; Pelesko and Bernstein 2002). Thus, an explicit expression for the pull-in voltage of the electrostatically actuated plate under uniform hydrostatic pressure can be obtained using a similar method to that in (Ahmad and Pratap 2010). Setting Eq. (30) equal to zero, four possible solutions for the pull-in voltage can be calculated (not given in this paper). As the pull-in voltage is positive, and increases with the electrode distance  $d_0$ , yet reduces with an increased hydrostatic pressure (Ahmad and Pratap 2010; Nabian et al. 2008; Bozkurt et al. 1999), its correct expression is

$$V = \frac{d_0}{\sqrt{\epsilon_0 R^2}} \sqrt{504Dd_0 + 2R^4P - 8\sqrt{3528D^2d_0^2 + 42Dd_0R^4P}} \quad (36)$$

This equation provides a convenient method to predict the pull-in voltage of the electrostatically actuated plate under uniform hydrostatic pressure. If there is no hydrostatic pressure applied to the plate, the pull-in voltage is

$$V_{pi} = \frac{5.369d_0}{R^2} \sqrt{\frac{Dd_0}{\epsilon_0}} \quad (37)$$

where  $V_{pi}$  represents the pull-in voltage of the electrostatically actuated plate without other loads or initial stress.

## Capacitance Calculation

Determination of the capacitance change is crucial to the performance analysis of electrostatically actuated microdevices such as CMUTs and capacitive pressure sensors (Rahman and Chowdhury 2011; Zhou et al. 2005). Based on the above static deflection analysis, the capacitance variation as a result of the plate deformation can be calculated using a surface integral over the plate area. The results are given by

$$C = \int_0^{2\pi} \int_0^R \epsilon_0 r / [d_0 - w(r)] dr d\theta \quad (38)$$

Substituting Eq. (32) into Eq. (38) and calculating the integral, the capacitance of the parallel-plate configuration subjected to electrostatic force as well as hydrostatic pressure is given by

$$C_{V,P} = \frac{\pi\epsilon_0}{2\sqrt{d_0\eta_{1b}}} \ln \frac{\sqrt{d_0} + \sqrt{\eta_{1b}R^2}}{\sqrt{d_0} - \sqrt{\eta_{1b}R^2}} \quad (39)$$

If only the effect of the electrostatic force on the capacitance is considered, the corresponding expression is

$$C_V = \frac{\pi\epsilon_0}{2\sqrt{d_0H_1}} \ln \frac{\sqrt{d_0} + \sqrt{H_1R^2}}{\sqrt{d_0} - \sqrt{H_1R^2}} \quad (40)$$

where

$$H_1 = (84Dd_0^2 - QR^4 - B_1) / (128Dd_0R^4) \quad (41)$$

The term  $C_V$  represents the capacitance variation induced solely by the electrostatic force. When both applied voltage and hydrostatic pressure equal zero, the capacitance of the parallel-plate configuration is the limit of Eq. (39). It is

$$C_{V=0, P=0} = (\epsilon_0 \pi R^2) / d_0. \quad (42)$$

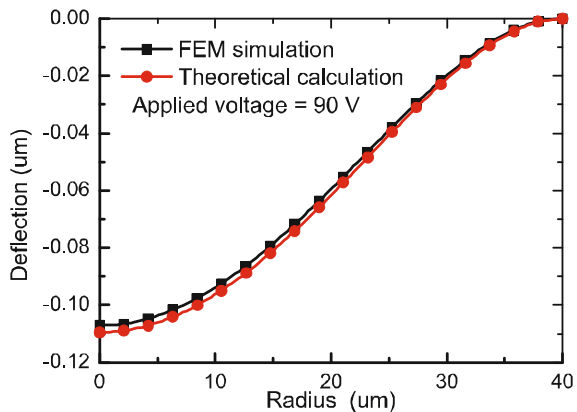
This equation is the most used capacitance approximation to an undeformed parallel-plate configuration (Fragiacomo et al. 2010).

## Validation of Theoretical Analysis

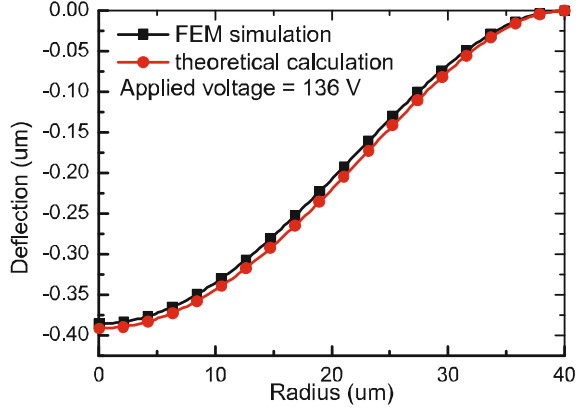
In order to validate the theoretical analyses obtained above, their results were compared with those from simulations and previous literature. The simulations were carried out using a commercially available FEM package (ANSYS12.0, ANSYS Inc., Canonsburg, PA). Due to the axial symmetry of the circular plates and applied loads shown in Fig. 7, a 2D axisymmetrical model was constructed for a reduced computation time. The upper microplate was constructed using plane structural elements (PLANE182), and its rim was constrained. The vacuum and the electrostatic effect between the upper and bottom microplates were modeled using electromechanical coupling elements (TRANS126). The electromechanical coupling elements can apply electrostatic attraction forces to the nodes to which they are attached. The top nodes of these elements were attached to the upper microplate, and their bottom nodes were simply constrained representing the bottom microplate. Based on this model, the ANSYS Multi-Field Solver was employed to determine the static deflection of upper flexible microplate under different voltages and uniform pressures. The simulated results are shown in Figs. 9, 10, 11, 12, 13, 14, 15, and 16. Detailed material and geometry parameters used for simulations are shown in Tables 1 and 2. Table 2 was only used for Fig. 13, and Table 1 was used for the others. The material parameters in Table 2 are as same as those in Table 1.

Figures 9 and 10 show a comparison of the deflection curves of the microplate obtained by the FEM simulation and Eq. (34). Figure 9 shows the deformation profile of the plate under an applied voltage of 90 V with a maximum displacement of 0.11  $\mu\text{m}$ , which is less than 30% of the thickness of the plate (shown in Table 1).

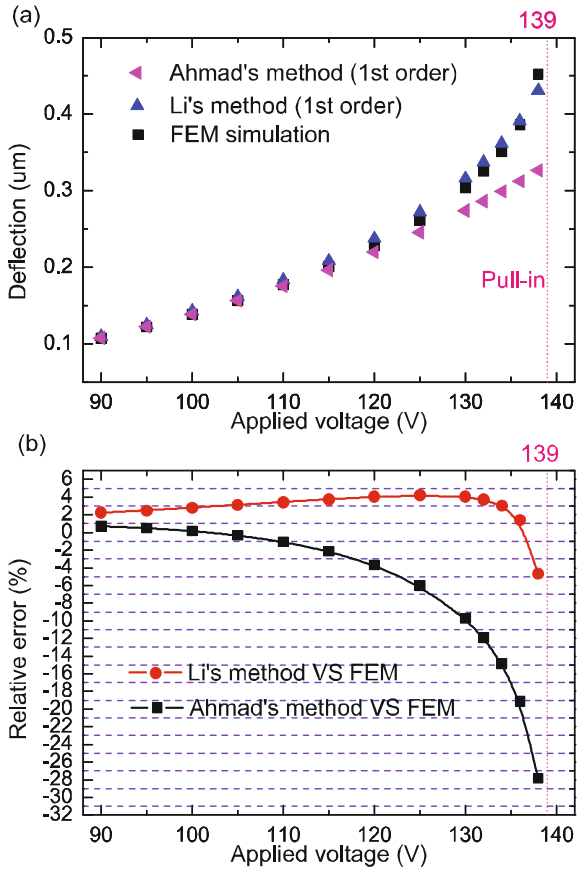
**Fig. 9** Deflection profile of the microplate subjected only to a voltage of 90 V



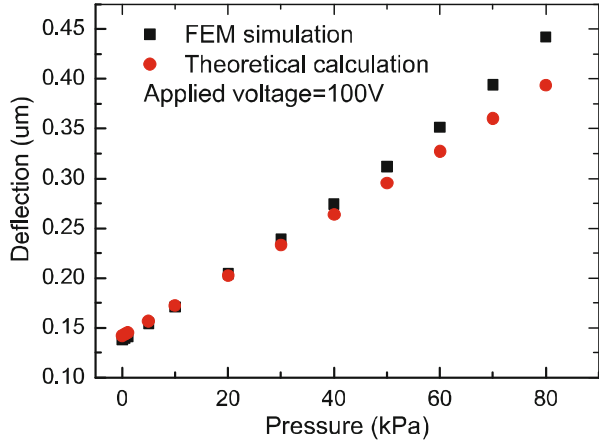
**Fig. 10** Deflection profile of the microplate subjected only to a voltage of 136 V



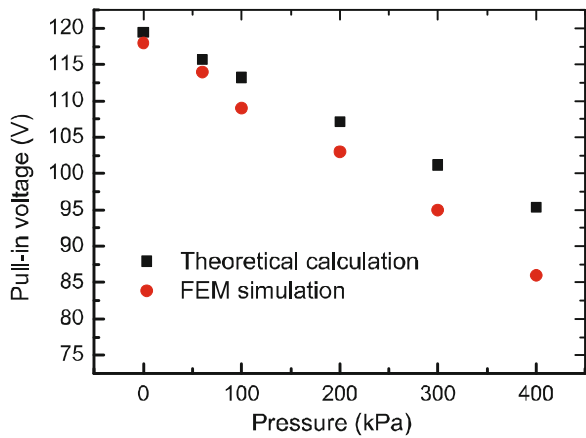
**Fig. 11** The displacements of microplate center point subjected to voltages varying from 90 to 138 V; (a) the displacements obtained by FEM simulations, our and Ahmad's analytical methods; (b) the errors of the results from our and Ahmad's analytical methods with respect to those from the FEM simulations



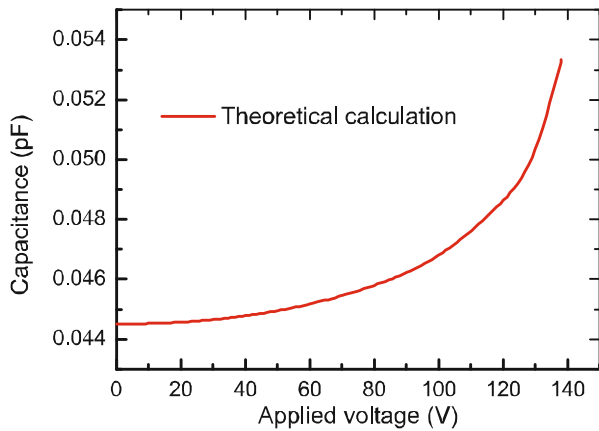
**Fig. 12** The displacements of the microplate center point under different uniform pressures with a fixed voltage of 100 V obtained by FEM simulations and Eq. (32)



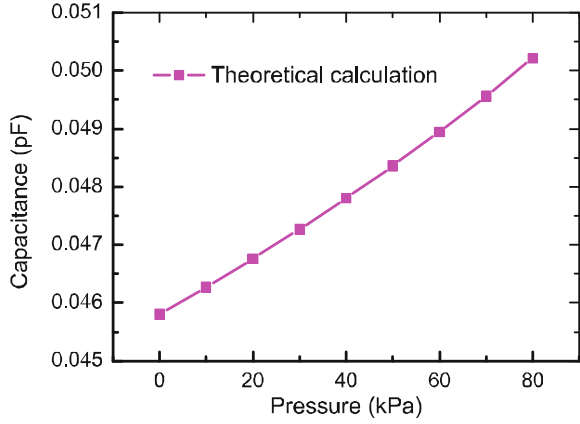
**Fig. 13** Pull-in voltages of the microplate under different uniform hydrostatic pressures



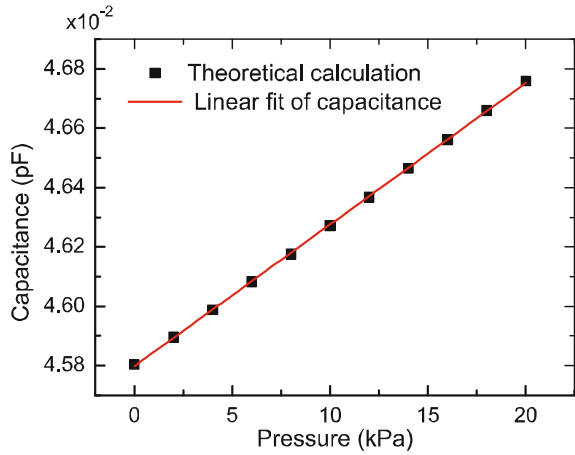
**Fig. 14** Capacitance variation with respect to the applied voltage



**Fig. 15** Capacitance variation with respect to uniform hydrostatic pressure within the range from 0 to 80 kPa with a fixed voltage of 80 V



**Fig. 16** Capacitance variation with respect to uniform hydrostatic pressure within the range from 0 to 20 kPa with a fixed bias voltage of 80 V



**Table 1** The first group of material and geometry parameters for simulations (μMKS units)

Parameter	Value
Radius $R$	40 μm
Thickness $h$	1 μm
Poisson's ratio $\nu$	0.3
Separation distance $d_0$	1 μm
Young's modulus $E$	$1.69 \times 10^5$ MPa
Dielectric constant $\epsilon_0$	$8.854 \times 10^{-6}$ pF/μm

Figure 10 shows the deformation profile under an applied voltage of 136 V with a maximum displacement of 0.39 μm, which is larger than 30% of the thickness (Shahiri-Tabarestani et al. 2012). As can be seen from the two figures, the analytical deflection curves agree well with the simulated results for both small and



**Table 2** The second group of parameters for simulations ( $\mu\text{MKSV}$  units)

Parameter	Value
Radius $R$	25 $\mu\text{m}$
Thickness $h$	1.2 $\mu\text{m}$
Separation distance $d_0$	0.4 $\mu\text{m}$

relatively large deflections. The shape function given by Eq. (27) provides a good approximation to the deformation profile of the plate under the electrostatic force. It can be seen in Figs. 9 and 10 that a better deformation approximation may be achieved if the displacement of the center point (the maximum displacement) of the plate is determined accurately. Furthermore, as the shape function is sufficient to approximate the deformation profile of the clamped circular plate under uniform hydrostatic pressure (Timoshenko and Woinowsky-Krieger 1959; Ventsel and Krauthammer 2001), it can be derived that Eq. (32) using the same shape function can closely approximate the shape of deformed microplate subjected to both electrostatic force and uniform pressure, if the displacement of microplate center point is predicted exactly.

Figure 11 shows a comparison of the displacements of the microplate center point (the maximum displacements) under different voltages obtained by FEM simulations, Eq. (34) in our study and Eq. (29) in Ahmad’s research (Ahmad and Pratap 2010). Figure 11a shows the displacements as a function of the voltages within a range of 90 to 138 V. It can be seen that the results from our and Ahmad’s analytical methods almost overlap with those from the FEM simulations when the applied voltage is below 120 V. When the applied voltage is above 120 V, the displacements predicted by Ahmad’s method begin to deviate appreciably from the simulated results. However, our analytical values still show small difference from the simulated results even when the applied voltage approaches the pull-in voltage, 139 V (determined using FEM simulations). Figure 11b shows the errors of the results predicted by our and Ahmad’s analytical methods compared with those obtained by the FEM simulations. The error from Ahmad’s method is smaller than 5% when the voltage is below 120 V and increases to about 28% when the voltage rises up to 138 V. However, the error from our method is smaller than 5% within the whole voltage range (from 90 to 138 V). In contrast, our solution in Eq. (29) shows a better approximation to the deflection of the plate than Eq. (29) in (Ahmad and Pratap 2010). Furthermore, a comparison of the maximum displacements predicted by Eq. (34) in our study and Eq. (34) in Ahmad’s research using the second-order term of electrostatic force expansion was also carried out (Ahmad and Pratap 2010). The maximum error of the results from Ahmad’s method with respect to those from the FEM simulations is 9.5% within the voltage range from 90 to 138 V. However, as mentioned above, the maximum error from our analytical method is 4.7%. Equation (34) in our study also shows a better prediction for the deflection than the nonlinear solution in Eq. (34) of Ahmad’s (Ahmad and Pratap 2010). Therefore, our analytical method can predict the deflection with a better accuracy within the voltage range from zero to the one approaching the pull-in voltage.

Figure 12 presents the maximum displacements of the clamped circular microplate under different uniform hydrostatic pressures with a fixed applied voltage of 100 V. As shown in the Fig. 12, the results calculated using Eq. (32) are consistent with those obtained by FEM simulations when the applied pressure is below 50 kPa. The corresponding relative error with the simulated results is less than 5.2%. When the pressure increases, the relative error rises accordingly and reaches up to about 11% at a pressure of 80 kPa. Additionally, it can be observed that the analytical displacement is smaller than that from the FEM simulations. The reason for this phenomenon lies in the fact that, when the deflection of the plate under a fixed bias voltage increases with the hydrostatic pressure, a corresponding electrostatic force is produced which will further lead to an increase in the deflection. As a result, the hydrostatic pressure applied to the plate under the fixed voltage can induce a larger deflection than that caused by the pressure alone. Similar to the electrical spring softening in Ladabaum et al. (1998), the additional electrostatic force caused by the pressure can be interpreted as a pressure-induced electrical spring softening (PIESS). As Eq. (32) does not take the effect of the PIESS into account, the displacements predicted by it are smaller than those simulated values. However, as mentioned above, Eq. (32) can provide an acceptable prediction for the deflection of the electrically actuated microplate under the pressure within a low range.

In order to further validate the availability of Eq. (32), a comparison of the center-point deflections calculated using Eq. (32) with the values reported by Nabian et al. (2008) is shown in Table 3. The deflections were calculated under different voltages with a constant uniform pressure of 200 kPa, for which the corresponding material and geometry parameters are shown in Table 4. As can be seen from Table 3, the analytical results from Eq. (32) agree well with those reported by Nabian. This further demonstrated the validity of Eq. (32) in predicting the microplate deflections under both electrostatic force and uniform hydrostatic pressure.

**Table 3** Center point deflections from our analytical calculations and Nabian's simulations (Nabian et al. 2008)

Bias voltage (V)	Our solutions ( $\mu\text{m}$ )	Nabian's results ( $\mu\text{m}$ )
0	0.099	0.102
100	0.123	0.129
200	0.212	0.226

**Table 4** Parameters used by Nabian (2008), Osterberg (1995) ( $\mu\text{MKSV}$  units)

Parameter	Value
Radius $R$	250 $\mu\text{m}$
Thickness $h$	20 $\mu\text{m}$
Separation distance $d_0$	1 $\mu\text{m}$
Poisson's ratio $\nu$	0.3
Young's modulus $E$	$1.69 \times 10^5$ MPa
Dielectric constant $\epsilon_0$	$8.854 \times 10^{-6}$ pF/ $\mu\text{m}$

Having demonstrated the availabilities of Eqs. (32) and (34) for static deflection analysis, we studied the performance of Eqs. (36) and (37) for the pull-in voltage prediction in the following section. Equation (37) has a same form as (35) given by Ahmad (Ahmad and Pratap 2010) and with the equation derived by setting the residual stress in Eq. (42) reported by Liao (2009) to zero. For a clamped circular plate with the material and geometry parameters shown in Table 4, we used our, Ahmad’s, and Liao’s equations to calculate the pull-in voltage  $V_{pi}$ . The results are shown in Table 5 along with the pull-in voltage from CoSolve-EM simulations reported by Osterberg (1995). The pull-in voltage predicted by Eq. (37) in our study deviates 0.7% from the full 3D simulation value produced by CoSolve-EM, which is much smaller than errors from Ahmad’s and Liao’s method. Therefore, Eq. (37) in our study shows a higher accuracy in predicting the pull-in voltage than Ahmad’s and Liao’s methods.

Furthermore, a comparison of the pull-in voltages of the electrostatically actuated microplate under different uniform pressures obtained through Eq. (36) and FEM simulations is shown in Fig. 13. The FEM simulation is time-consuming because the analysis is nonlinear and needs a large number of elements for the large-size structures. Therefore, the small-size parameters shown in Table 2 were used to simulate the pull-in phenomenon in order to reduce the computation time. Additionally, the FEM simulation shows poor simulation performance at the pull-in point of the electrostatic plate; however, it can accurately determine the static deflection when the applied voltage is below the pull-in voltage. A voltage making the maximum deflection of the plate equal to 50% of the air gap was used as the pull-in voltage in our FEM simulations, which is very close to the actual pull-in position being 53% of the air gap reported by Vogl et al. (Chao et al. 2006; Vogl and Nayfeh 2005). Though the pull-in voltage obtained from our simulations is a little smaller than its actual value, it is sufficient to validate theoretical predictions. Results obtained from FEM simulations and Eq. (36) in Fig. 13 show the pull-in voltage decreases with the uniform hydrostatic pressure. The error of the calculated values with respect to those from the FEM simulations is below 4% within the pressure range from 0 to 200 kPa. When the pressure increases, the error rises accordingly up to 10.9% at the pressure of 400 kPa. As the simulated pull-in voltage was underestimated, the error of our prediction should be smaller than those mentioned above. The reason for the relatively large error in the pressure range from 200 to 400 kPa is that the PIESS effect is not considered in Eq. (36). However, even with this relatively large error for high pressure, Eq. (36) can well determine the pull-in

**Table 5** Comparison of pull-in voltage with the results from Ahmad (Ahmad and Pratap 2010); Liao (2009); Osterberg (1995)

$V_{pi}$ (V) by Ahmad	$V_{pi}$ (V) by Liao	$V_{pi}$ (V) by us	$V_{pi}$ (V) by Osterberg	$\Delta\%$ between Ahmad and Osterberg	$\Delta\%$ between Liao and Osterberg	$\Delta\%$ between us and Osterberg
326	325	321	319	2.2	1.9	0.7

voltage of the electrostatic microplate under a uniform hydrostatic pressure in the low range.

Figure 14 shows the capacitance variation of the parallel-plate configuration as a function of the applied voltage, which is achieved using Eq. (40). As shown in Fig. 14, the capacitance increases slowly with the voltage below 80 V and rises fast when the voltage becomes larger, especially for the applied voltage close to the pull-in value, 139 V. This trend for the capacitance variation is similar to that in (Ahmad and Pratap 2010). The accuracy of the capacitance variation depends on the static deflection analysis of microplate under the applied voltage. An improved deflection analysis may contribute to a relatively accurate capacitance prediction.

Figure 15 shows the capacitance of the parallel-plate configuration under different uniform hydrostatic pressures with a fixed bias voltage of 80 V, in which the value of the capacitance is obtained using Eq. (39). The capacitance increases almost linearly with the applied pressure in the range from 0 to 80 kPa. This linear relationship may be more obvious when the pressure changes in a lower range, such as in the pressure range from 0 to 20 kPa shown in Fig. 16. So Eq. (39) is helpful for the pressure-induced capacitance variation analysis of the electrostatic actuation structure, especially for designing and optimizing those capacitive pressure sensors.

The theoretical results mentioned above have high accuracy when compared with those from FEM simulations and literature. These demonstrate that the expressions, derived from the model based on thin plate theory and first-order electrostatic force expansion, are accurate enough for the mechanical behavior analysis of the microplate under electrostatic force and uniform hydrostatic pressure. However, there are several points in our analytical model that should be stated for better understanding and using it.

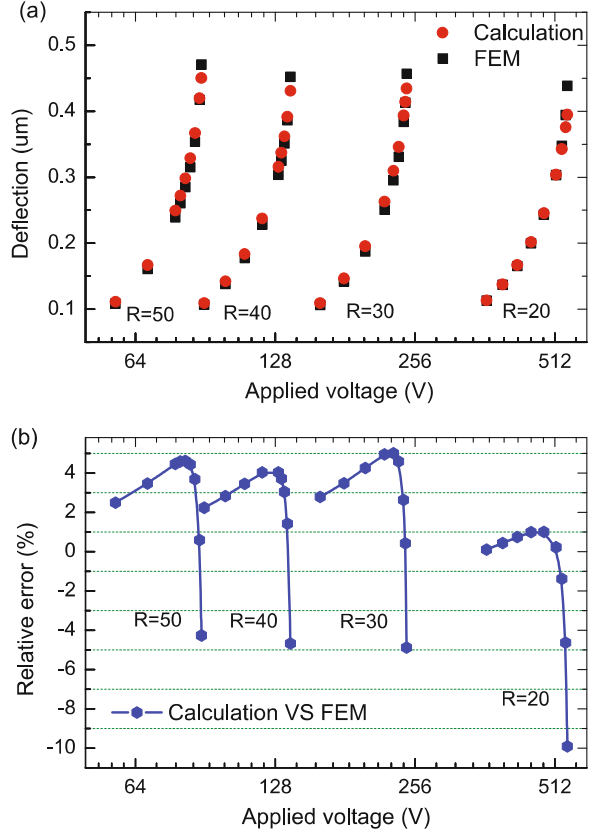
Firstly, the thin plate theory shows poor accuracy when the deflection is larger than 30% of the thickness because it does not consider the stretch of the middle plane of the plate (Timoshenko and Woinowsky-Krieger 1959). However, the theoretical analysis based on the thin plate theory in our study shows a small difference even for the deflection larger than 30% of the plate thickness, as shown in Fig. 11. In addition, the expression for the pull-in voltage prediction has a smaller error than those in literature, as shown in Table 5. The reasons for these are as follows. The pull-in position of the electrostatically actuated configuration in our study occurs at about 54% of the electrode distance  $d_0$ , which very approximates to 53% of the electrode distance given in (Vogl and Nayfeh 2005; Liao et al. 2009). For the situations where  $d_0 \leq h$ , the maximum deflection range in our study is expanded to 54% of the plate thickness, which is not a very large expansion with respect to the relatively accurate range (30% of the plate thickness) for the thin plate theory (Timoshenko and Woinowsky-Krieger 1959; Ventsel and Krauthammer 2001). Therefore, although the thin plate theory will overestimate the deflection in comparison with that taking the stretch of middle plane of the plate into account (especially for the deflection larger than 30% of the plate thickness (Timoshenko and Woinowsky-Krieger 1959)), it does not generate a large deflection difference. Furthermore, the first-order electrostatic force expansion is smaller than, but has a relatively accurate approximation to, the actual value of electrostatic force for  $w/d_0 \leq 54\%$ , as shown in Fig. 8.

Therefore, the first-order electrostatic force expansion proposed in our study does not generate a large difference either. As a result, the reduced-order model based on thin plate theory and first-order electrostatic force expansion can contribute to analytical solutions with a high accuracy.

Secondly, although the second-order electrostatic force expansion has a better approximation to Eq. (18) than the first-order term, the reduced-order model using the second-order term, similar to Eq. (28), does not contribute to more accurate solutions than the model Eq. (28) using the first-order term. As mentioned before, the reduced-order model in Eq. (28) based on the thin plate theory and the first-order electrostatic force expansion can obtain relatively accurate solutions. Although the second-order electrostatic force expansion is larger and more accurate than the first-order term, the solution to the model including the second-order electrostatic force term largely overestimates the deflection which has a larger difference with the results from FEM simulations and other literature than the solution to Eq. (28). Actually, as shown in Fig. 8, the first-order electrostatic force expansion even has a better approximation to the electrostatic force than the second-order term in Eq. (21). Therefore, it has high approximation accuracy and can contribute to solutions with higher accuracy than those in literature, which is sufficient for the static mechanical behavior analysis of the electrostatically actuated microplates. The second-order electrostatic force expansion or higher-order terms given in Eq. (24) may be used for the model taking the stretch of the middle plane of the plate into account to obtain more accurate solutions.

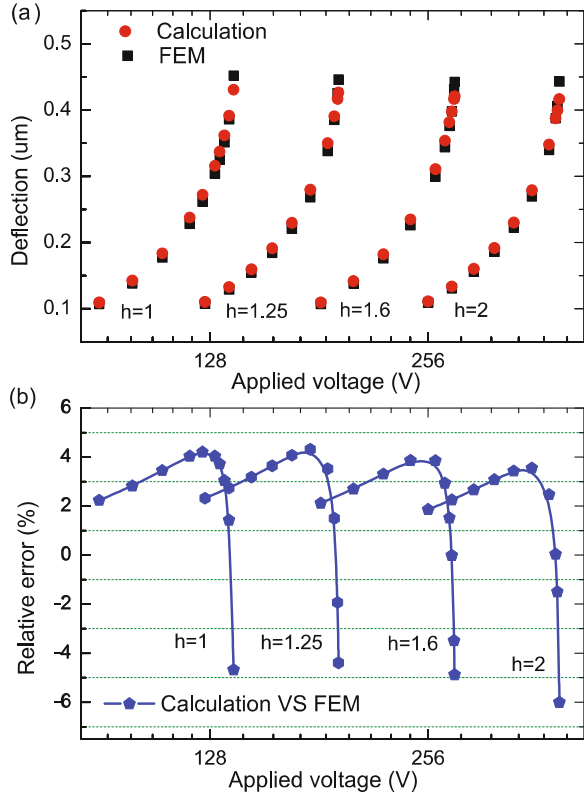
Finally, the proposed reduced-order model is advantageous for establishing simple and explicit expressions for the mechanical behaviors of electrostatically actuated microplates. The model using the first-order electrostatic force expansion can contribute to expressions with high accuracy, which are general and robust enough for different device dimensions. To show the general performance of the model, a parametric study was carried out. The results are shown in Figs. 17 and 18, for which the basic parameters are as same as those in Table 1. Figure 17a shows the maximum deflections of the microplates with different radii  $R$  (i.e., different ratios of plate radius to thickness  $R/h$ ). As shown in Fig. 17b, the errors of the calculated values with respect to those from the FEM simulations are less than 5.0% within the deflection range from very small deformations to near pull-in positions for  $R/h = 50, 40, \text{ and } 30$ . However, the maximum error becomes as large as 9.9% for  $R/h = 20$ . Figure 18a shows the maximum deflections of the microplates with different thicknesses  $h$  (i.e., different ratios of plate radius to thickness  $R/h$  and different ratios of thickness to electrode distance  $h/d_0$ ). As shown in Fig. 18b, the errors between the calculated and simulated values are less than 5.0% within the large deflection range for  $R/h = 40, 32, \text{ and } 25$ . When the ratio  $R/h = 20$ , the maximum error is 6.0% which just has a very little increase in comparison with other cases. However, this error is much smaller than that shown in Fig. 17 for the same ratio  $R/h = 20$ . It is because the thickness is larger than the electrode distance as being  $h/d_0 = 2$  in Fig. 18; yet the thickness is equal to the electrode distance as being  $h/d_0 = 1$  in Fig. 17. As the pull-in position occurs at 54% of the electrode distance  $d_0$ , for the same  $d_0$ , the larger ratio of thickness to electrode distance  $h/d_0$  means the smaller ratio of the deflection at

**Fig. 17** Comparison of deflections of the microplate center point with different radii  $R$ , fixed thickness  $h$ , and electrode distance  $d_0$ ; (a) the deflections from small deformation to near pull-in positions obtained from FEM simulations and theoretical calculation; (b) the errors of the simulated results with those from FEM simulations



pull-in position to thickness  $w_{pi}/h$ , which will result in deflection analysis with a higher accuracy (Ventsel and Krauthammer 2001; Shahiri-Tabarestani et al. 2012). Additionally, for the cases shown in Fig. 17, the relative errors between the calculated and simulated pull-in voltages change from 0.2% to 2.5% when the radius to thickness ratios vary from 50 to 20. However, for the cases shown in Fig. 18, the relative errors of the pull-in voltages change from 0.7% to 1.7% when the radius to thickness ratios vary from 40 to 20 (i.e., the thickness to electrode distance ratios vary from 1 to 2). It can be found that the pull-in voltage with  $h/d_0 = 2$  in Fig. 18 has a higher accuracy than that with  $h/d_0 = 1$  in Fig. 17 for the same ratio  $R/h = 20$ . This suggests that a larger thickness to electrode distance ratio also leads to higher prediction accuracy for the pull-in voltage for the same radius to thickness ratio. However, even for the harsh case where  $R/h = 20$  and  $h/d_0 = 1$  shown in Fig. 17, the error of the predicted pull-in voltage is comparable to that given in (Vogl and Nayfeh 2005) using a reduced order with five modes. In conclusion, the case with  $R/h = 20$  and  $h/d_0 = 1$  may be a limit of device dimensions which the theoretical analysis is applicable to. The theoretical analysis has high accuracy when the ratio of radius to thickness varies from 50 to 20 (i.e., the diameter to thickness ratio varies from 100 to 40),

**Fig. 18** Comparison of deflections of the microplate center point with different thicknesses  $h$ , fixed radius  $R$ , and electrode distance  $d_0$ ; (a) the deflections from small deformations to near pull-in positions obtained from FEM simulations and theoretical calculation; (b) the errors of simulated results with those from FEM simulations



which covers a very large part of the dimension range of thin plates (Ventsel and Krauthammer 2001).

In all, as the stretch of middle plane of the flexible microplate is not considered in our theoretical model, the derived analytical expressions are applicable to the situations in which the ratio of microplate diameter to its thickness varies from 100 to 40, and the microplate deflection is smaller than its thickness, and the electrode distance is smaller than or equal to the thickness (these cases have been widely used in the electrostatically actuated configurations (Yaralioglu et al. 2003; Ahmad and Pratap 2010; Vogl and Nayfeh 2005; Liao et al. 2009)). Based on these assumptions, Eq. (34) can predict the static deflection of microplate under only an electrostatic force even if the deflection is close to the pull-in position. Equation (37) can provide a relatively accurate pull-in voltage prediction in comparison with the analytical methods in literature. Equations (32) and (36) are more suitable to analyzing static deflection and pull-in voltage of electrostatically actuated microplates subjected to hydrostatic pressure within a low range. Equation (40) can be used to analyze the capacitance variation and electromechanical coefficient of electrostatic microdevices such as CMUTs (Yaralioglu et al. 2003). Equation (39) suggests that the capacitance of the parallel-plate configuration varies linearly with

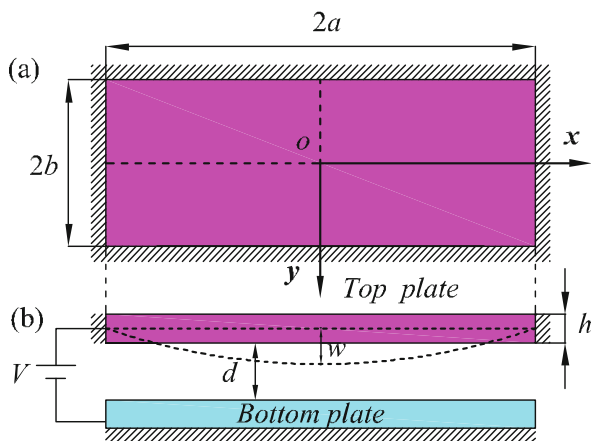
the pressure within a low range. These theoretical analyses are capable of widespread application to microelectromechanical system (MEMS) as a growing number of electrostatic microdevices require a high sensitivity, low energy consumption, and low pull-in voltage, for which the electrostatic microdevices are often required to have a thin plate thickness and a small electrode separation distance. Additionally, the methods proposed can be further expanded to analyze other mechanical behaviors of the electrostatic microdevices such as dynamic pull-in voltages, resonant frequencies, and large deflections if the strain of the middle plane of the microplate is taken into account.

## Distributed Electromechanical Models of Electrostatically Actuated Rectangular Microplates

### Problem Formulation

A schematic of an electrostatically actuated configuration composed of two parallel rectangular microplates is considered in this theoretical analysis. Its schematic is shown in Fig. 19. The top rectangular microplate with its four edges clamped has a length  $2a$ , width  $2b$ , and thickness  $h$ . The microplate is assumed to be a thin, linear, and isotropic microplate. It is parallel separated with the fully fixed bottom microplate with an effective electrode distance  $d_0$ . The space between the two microplates is assumed as a vacuum or air with a dielectric constant  $\epsilon_0$ . The top and bottom microplates are electrically conductive and work as the top and bottom electrodes, respectively. When a bias voltage is applied across the two microplates, the top microplate is deflected toward the bottom one by the resulting electrostatic force. The deflection of the top microplate is assumed to be small compared to its thickness  $h$ . The classical thin microplate theory is used for the deflection analysis, which is adequate when the thickness-to-length ratio ( $h/2b$ ) is relatively small (i.e.,  $(h/2b) < 1/20$ ) (Nabian et al. 2008). It should be noted the plate referred to in the following sections is the top microplate unless stated otherwise.

**Fig. 19** Schematic of a parallel rectangular microplate configuration under electrostatic actuation; (a) top view of the configuration; (b) cross-sectional view of the configuration





The general equation governing the bending of an isotropic rectangular microplate is given in (Timoshenko and Woinowsky-Krieger 1959). For the transverse deflection of the microplate under a distributed electrostatic force, the partial differential equation can be written as

$$D \left( \frac{\partial^4 w}{\partial x^4} + 2 \frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} \right) = F_e(x, y) \quad (43)$$

where  $w = w(x, y)$  is the transverse deflection of the microplate,  $x$  and  $y$  are the coordinates originated at the center point of the microplate, and  $F_e(x, y)$  is the electrostatic force per unit area acting on the microplate and can be given by

$$F_e(x, y) = \frac{\varepsilon_0 V^2}{2[d_0 - w(x, y)]^2} \quad (44)$$

where  $V$  is the bias voltage applied to the microplate.

For the rectangular microplate with its four edges clamped as shown in Fig. 19, the boundary conditions are

$$\begin{aligned} w(x, y) = 0 \text{ and } \frac{dw(x, y)}{dx} = 0 \text{ at } x = \pm a, \forall y \\ w(x, y) = 0 \text{ and } \frac{dw(x, y)}{dy} = 0 \text{ at } y = \pm b, \forall x \end{aligned} \quad (45)$$

When  $2a = 2b$ , the rectangular microplate becomes a square one. The boundary conditions in Eq. (46) are also applicable to the clamped square microplate (Chao et al. 2006).

## Reduced-Order Model

For Eq. (43), it is hard to establish an accurate analytical solution owing to the nonlinearity of electrostatic force term. So Galerkin method is employed to decompose the deflection governing equation into a coupled set of ordinary algebraic equations (a reduced-order model). Then the mechanical behavior of the electrostatic actuated microplate can be simulated by solving these ordinary equation systems. This technique and its application to nonlinear system have been discussed in (Nayfeh and Mook 1979; Nayfeh 2000).

The function used to approximate the transverse deflection  $w(x, y)$  is assumed as

$$w(x, y) = \sum_{i, j=0}^N k_{ij} \phi_{ij}(x, y), \quad (46)$$

where  $k_{ij}$  is the unknown coefficient to be determined,  $\phi_{ij}(x, y)$  is the linearly independent coordinate functions (they are also called trial functions) that satisfy

all the prescribed boundary conditions (Ventsel and Krauthammer 2001), and  $N$  is the largest value of the variables  $i$  and  $j$ . When  $N$  approaches infinity, the approximation in Eq. (46) becomes exact if the chosen trial functions form a complete set. For the rectangular microplate, the cosine-like functions are chosen as the trial functions to approximate its deformed shape. The expression for the function is given by (Kerrou and Hobar 2006)

$$\phi_{ij} = \cos^2\left(\frac{(2i+1)\pi x}{2a}\right) \cos^2\left(\frac{(2j+1)\pi y}{2b}\right). \quad (47)$$

These trial functions satisfy all the boundary conditions given in Eq. (45), the lowest order ( $i = 0, j = 0$ ) of which has been widely used to simulate the deflection of silicon diaphragms under hydrostatic pressure (Elgamel 1995, 1999).

Having determined the treatment method of the nonlinear electrical force and the trial functions, the process of establishing reduced-order model can be started by substituting Eq. (46) into Eq. (43) and treating the electrostatic force using the same method as Eq. (24). Then, multiplying both sides of the resulting equation with the term  $(d_0 - w)$  and functions  $\phi_{lm}(x, y)$  and finally integrating the outcome over the area of the microplate, the ordinary algebraic equation system for Eq. (43) is obtained (Li et al. 2015b):

$$\begin{aligned} d_0 \iint \nabla^4 w \times \phi_{lm} dx dy - \sum_{i,j=0}^N k_{ij} \iint \nabla^4 w \times \phi_{ij} \phi_{lm} dx dy &= Q_1 \left( \iint \phi_{lm} dx dy \right. \\ &+ \frac{1}{d_0} \sum_{i,j=0}^N k_{ij} \iint \phi_{ij} \phi_{lm} dx dy \\ &+ \frac{1}{d_0^2} \sum_{i,j,p,q=0}^N k_{ij} k_{pq} \iint \phi_{ij} \phi_{pq} \phi_{lm} dx dy \\ &\left. + \frac{1}{d_0^3} \sum_{i,j,p,q,g,h=0}^N k_{ij} k_{pq} k_{gh} \iint \phi_{ij} \phi_{pq} \phi_{gh} \phi_{lm} dx dy + \dots \right), \end{aligned} \quad (48)$$

$$i, j, p, q, g, h, l, m = 0, 1, 2, 3 \dots N,$$

where

$$\begin{aligned} \nabla^4 w &= \frac{\partial^4 w}{\partial x^4} + 2 \frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} \\ &= \sum_{i,j=0}^N k_{ij} \left[ \frac{\partial^4 \phi_{ij}(x, y)}{\partial x^4} + 2 \frac{\partial^4 \phi_{ij}(x, y)}{\partial x^2 \partial y^2} + \frac{\partial^4 \phi_{ij}(x, y)}{\partial y^4} \right], \end{aligned} \quad (49)$$

and

$$Q_1 = \frac{\epsilon_0 V^2}{2d_0 D} \tag{50}$$

The integrals in Eq. (48) can be evaluated once the trial functions of Eq. (47) are substituted into Eq. (48) and the number  $N$  is chosen. Then, once all  $k_{ij}$  are determined by solving Eq. (48), an approximate solution to Eq. (43) can be established by Eq. (46).

### Static Deflection Analysis

Owing to the closeness between the shape given by the trial function  $\phi_{i=0, j=0}$  and that of a practical deformed microplate (Elgamel 1995, 1999), the lowest order ( $N = 0$ ) of the trial function in Eq. (48) is used. Additionally, as the complexity of both solving process and final solution increases with the order of the Taylor’s expansion of the electrostatic force, the first-order electrostatic force expansion is used to establish explicit expressions for the static deflection. These simplifications can not only contribute to a simple theoretical expression for the relationship between the deflection and the structure parameters but also are sufficient for predicting the behavior of microplate under electrostatic actuation.

Substituting Eq. (47) into Eq. (46) and setting  $N = 0$ , the deflection function in Eq. (46) can be rewritten as

$$\begin{aligned} w(x, y) &= k_{00} \phi_{00}(x, y) \\ &= k_{00} \cos^2\left(\frac{\pi x}{2a}\right) \cos^2\left(\frac{\pi y}{2b}\right), \end{aligned} \tag{51}$$

where  $k_{00}$  in fact gives the largest deflection of the microplate (the deflection of the center point). Then substituting Eq. (51) into Eq. (48) and using the first-order electrostatic force expansion, Eq. (48) can be simplified as

$$\begin{aligned} d_0 \int_{-b}^b \int_{-a}^a \nabla^4 w \times \phi_{00} dx dy - k_{00} \int_{-b}^b \int_{-a}^a \nabla^4 w \times \phi_{00}^2 dx dy \\ = Q_1 \left( \int_{-b}^b \int_{-a}^a \phi_{00} dx dy + \frac{k_{00}}{d_0} \int_{-b}^b \int_{-a}^a \phi_{00}^2 dx dy \right). \end{aligned} \tag{52}$$

Integrating the integrals and then solving the resulting equation, two possible solutions to the constant  $k_{00}$  are given by

$$k_{00} = \frac{(3a^4 + 2a^2b^2 + 3b^4)d_0^2\pi^4 - 9a^4b^4Q_1 \pm M}{(5a^4 + 4a^2b^2 + 5b^4)d_0\pi^4}, \tag{53}$$

where

$$M = \sqrt{(3a^4 + 2a^2b^2 + 3b^4)^2 d_0^4 \pi^8 - 2a^4 b^4 (107a^4 + 82a^2b^2 + 107b^4) d_0^2 \pi^4 Q_1 + 81a^8 b^8 Q_1^2}. \quad (54)$$

Actually, only one of the two possible solutions is consistent with the physical situation in which the deflection of the microplate increases with the applied voltage. So the correct expression for  $k_{00}$  is

$$k_{00} = \frac{(3a^4 + 2a^2b^2 + 3b^4) d_0^2 \pi^4 - 9a^4 b^4 Q_1 - M}{(5a^4 + 4a^2b^2 + 5b^4) d_0 \pi^4}. \quad (55)$$

Substituting Eq. (55) into Eq. (51), the complete expression for the transverse deflection of an electrostatically actuated rectangular microplate is obtained:

$$w(x, y) = \frac{(3a^4 + 2a^2b^2 + 3b^4) d_0^2 \pi^4 - 9a^4 b^4 Q_1 - M}{(5a^4 + 4a^2b^2 + 5b^4) d_0 \pi^4} \cos^2\left(\frac{\pi x}{2a}\right) \cos^2\left(\frac{\pi y}{2b}\right). \quad (56)$$

When  $a = b$ , the rectangular plate becomes a square one, which can also be understood as a special case of the rectangular microplate. Equation (56) can be simplified as

$$w_{sq}(x, y) = \frac{8d_0^2 \pi^4 - 9a^4 Q_1 - M_{sq}}{14d_0 \pi^4} \cos^2\left(\frac{\pi x}{2a}\right) \cos^2\left(\frac{\pi y}{2a}\right), \quad (57)$$

where  $w_{sq}(x, y)$  represents the transverse deflection of a clamped square microplate under electrostatic force;  $M_{sq}$  is given by

$$M_{sq} = \sqrt{64d_0^4 \pi^8 - 592a^4 d_0^2 \pi^4 Q_1 + 81a^8 Q_1^2}. \quad (58)$$

Equations (56) and (57) give the static deflections of the clamped rectangular and square microplates under electrostatic force. Furthermore, the two equations provide approaches to analyze the pull-in voltage and the capacitance changes of electrostatically actuated microplates.

## Pull-In Voltage Prediction

Pull-in voltages of electrostatically actuated microplates can be determined from the two possible solutions in Eq. (53). It can be derived that there is a certain voltage at which the two solutions equal each other. This voltage is a critical point where the pull-in phenomenon occurs (Ahmad and Pratap 2010; Vogl and Nayfeh 2005; Pelesko and Bernstein 2002). The pull-in voltage can be obtained using the similar

method in (Ahmad and Pratap 2010). Setting Eq. (54) equal to zero and solving the resulting equation, the expression for the pull-in voltage is obtained:

$$V_{pi} = \frac{1.549d_0^{3/2}D^{1/2}}{a^2b^2\sqrt{\epsilon_0}} \left[ 107a^4 + 82a^2b^2 + 107b^4 - 4\sqrt{2(335a^8 + 518a^6b^2 + 870a^4b^4 + 518a^2b^6 + 335b^8)} \right]^{\frac{1}{2}}. \quad (59)$$

For the pull-in voltage of a clamped square microplate (i.e.,  $a = b$ ), this expression can be simplified as

$$V_{pi(sq)} = \frac{4.618d_0^{3/2}D^{1/2}}{a^2\sqrt{\epsilon_0}}. \quad (60)$$

This expression provides a simple method to predict the pull-in voltage of the clamped square microplate under electrostatic force.

## Capacitance Calculation

The capacitance of an electrostatically actuated configuration changes with the deformation of the upper flexible microplate. It can be evaluated by integrating over the area of the deformed microplate. The expression of capacitance is given by (Rahman and Chowdhury 2010)

$$C = \iint_A \frac{\epsilon_0 dx dy}{(d_0 - w(x, y))}, \quad (61)$$

where  $A$  denotes the integration area. As Eqs. (56) and (57) are cosine-like functions, the integration in Eq. (61) cannot be solved analytically (Elgamel 1999). Therefore, Gauss-Legendre numerical integration is used to evaluate Eq. (61). When the deflection functions in Eq. (56) or Eq. (57) are determined, the capacitance between the fixed bottom microplate and the deformed upper microplate can be calculated.

## Validation of Theories

In order to validate the theoretical analyses obtained above, their results were compared with those from simulations and previous literature. The simulations were carried out using a commercially available FEM package (ANSYS12.0, ANSYS Inc., Canonsburg, PA). A3-D electromechanical coupling model was constructed for the electrostatically actuated configuration shown in Fig. 19. The top microplate was constructed using 3D structural elements (SOLID185), and its rim was constrained. The vacuum and the electrostatic effect between the upper flexible and bottom

microplates were modeled using electromechanical coupling elements (TRANS126) which were generated using EMTGEN command. The ANSYS Multi-Field Solver was employed to determine the static deflection of the flexible microplate under different voltages. Detailed material and geometry parameters used for simulations are shown in Table 6.

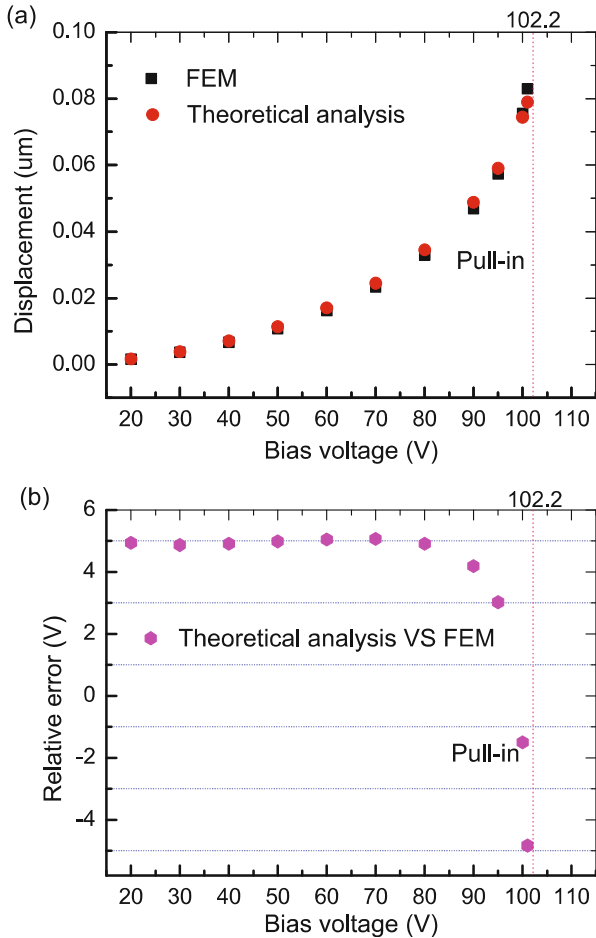
Figure 20 shows a comparison of the center displacements of a clamped rectangular microplate under different voltages. The dotted line in Fig. 20 corresponds to the pull-in voltage of the microplate. Figure 20a shows the center displacements of the microplate under a bias voltage range from 20 to 101 V which nearly approaches the pull-in voltage, 102.2 V. It can be seen that the calculated values agree with the simulated results over a large deflection range, from the undeformed state to that approaching the pull-in position. The relative errors of the calculated values with those simulated ones are shown in Fig. 20b. The relative errors are less than 5% within the large deflection range. It is demonstrated that Eq. (58) has a high accuracy in predicting the center deflection of a rectangular microplate.

Figure 21 shows the displacements of the rectangular microplate along the axis directions under a bias voltage of 90 V. Owing to the symmetry of the microplate deformation under a bias voltage, Fig. 20 only shows the microplate deflection along positive axis directions. Figure 21a shows a comparison of the displacements along the  $x$ -axis direction calculated using Eq. (56) with those obtained by FEM simulations. Figure 21b shows a comparison between the displacements along the  $y$ -axis direction calculated using Eq. (56) and those obtained by FEM simulations. As can be seen from the Fig. 21a, b, both theoretical results are consistent with the simulated values. It is also shown that the calculated displacements along the  $y$ -axis direction have a better agreement with the simulated values than those along the  $x$ -axis direction. This may result from different symmetries of the rectangular plate about the  $x$ -axis and the  $y$ -axis and the nonlinearity of the electrostatic force. Owing to the different lengths of rectangular microplate in the  $x$ -axis and the  $y$ -axis directions, the symmetries of electrostatic force distribution and deformation about the  $y$ -axis are not consistent with those about the  $x$ -axis. So the difference between simulated and calculated deflections along the  $y$ -axis is not consistent with that along the  $x$ -axis. When the length to width ratio approaches one, the symmetries along the  $x$ -axis and the  $y$ -axis approach are consistent and the differences along the  $x$ -axis and the  $y$ -axis approach are similar. The reason why the difference along the  $y$ -axis is smaller than

**Table 6** The material and geometry parameters for simulations ( $\mu\text{MKSV}$  units)

Parameters	Value
Length $2a$	40 $\mu\text{m}$
Width $2b$	20 $\mu\text{m}$
Thickness $h$	1 $\mu\text{m}$
Poisson's ratio $\nu$	0.28
Separation distance $d$	0.2 $\mu\text{m}$
Young's modulus $E$	$1.30 \times 10^5$ MPa
Dielectric constant $\epsilon_0$	$8.854 \times 10^{-6}$ pF/ $\mu\text{m}$

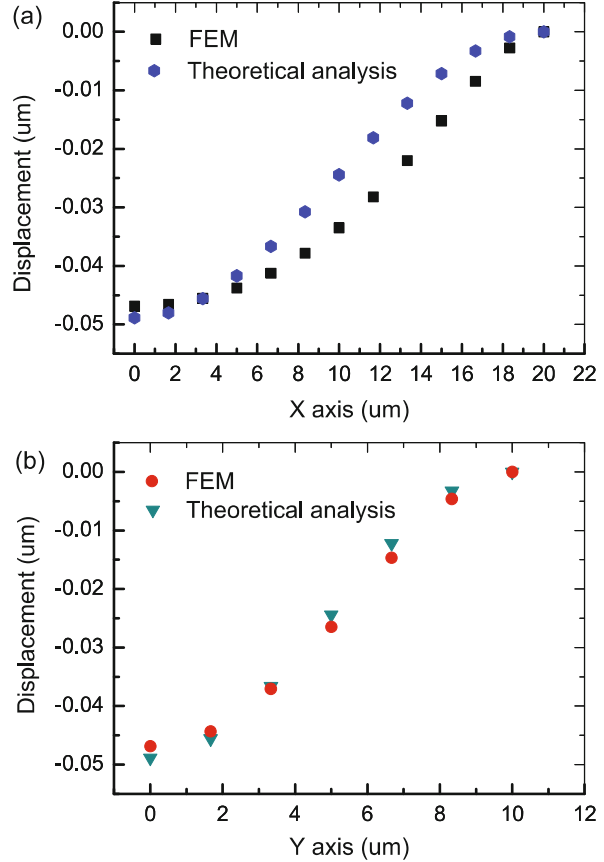
**Fig. 20** Displacements of the center point of a rectangular microplate under different bias voltages; (a) comparison between the displacements obtained by FEM simulations and those calculated using Eq. (56); (b) relative error between the calculated and simulated displacements



that along the  $x$ -axis depends on the used trial function itself. The shape function used in Eq. (56) shows better performance to simulate the deformation of rectangular microplate along its width direction. However, Eq. (56) is enough for predicting the deformation of rectangular microplate under electrostatic force. For better simulating the deformation along the  $x$ -axis, a mostly used method (Kerrou and Hobar 2006; Elgamel 1995) is to make a partial modification to the deflection function in Eq. (56). For instance, a modified function for Eq. (56) was given in Appendix C, which better approaches the deformation of rectangular microplate along the  $x$ -axis.

Figure 22 shows center displacements of a square microplate under a voltage range from 20 to 66 V. The dotted line along vertical direction in Fig. 22 denotes the pull-in voltage of the microplate as being 67.2 V. The edge length of square microplate is 30  $\mu\text{m}$ , that is,  $2a = 2b = 30 \mu\text{m}$ . The other geometry and material parameters for the microplate are the same as those in Table 6. Figure 22a shows a comparison of the calculated displacements using Eq. (57) with those obtained by

**Fig. 21** Comparison of displacements of the rectangular microplate along the axis directions under a bias voltage of 90 V; (a) displacements along the  $x$ -axis direction; (b) displacements along the  $y$ -axis direction



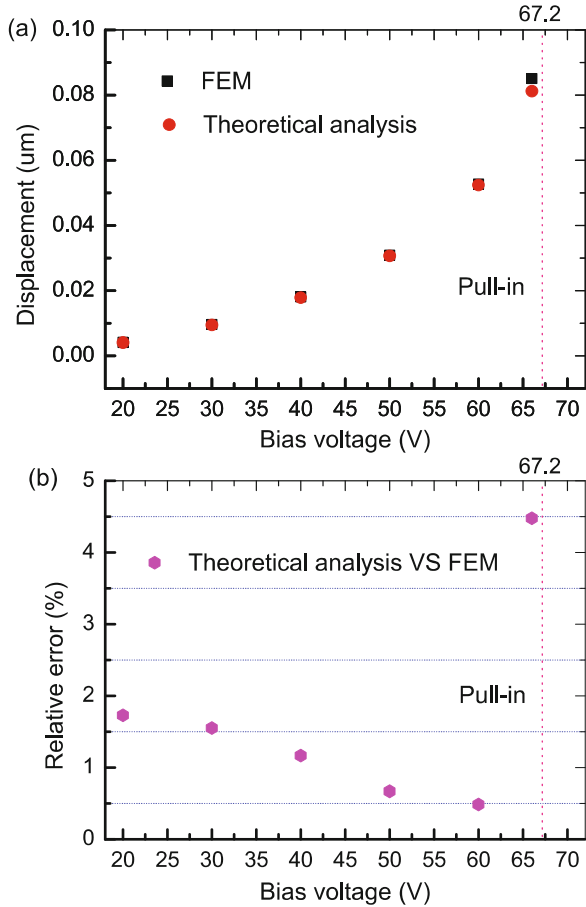
FEM simulations. There is an agreement between them. As shown in Fig. 22b, the difference of theoretical results with respect to those from FEM simulations is less than 4.8% even if the deflection approaches the pull-in position of microplate. These results demonstrate that the theoretical model (or Eq. (57)) has robustness, being able to predict the deflections up to the pull-in position.

A comparison of the displacements of the square microplate along the axis directions calculated using Eq. (57) with those obtained by FEM simulations is shown in Fig. 23. Figure 23a, b shows the displacements of square microplate along the  $x$ -axis and the  $y$ -axis directions, respectively. Both theoretical results in Fig. 23a, b agree with those simulated values. These indicate that Eq. (57) can well predict the deformation of a clamped square microplate.

Furthermore, we also validated the theoretical model using the experimental data in other literature. Figure 24 shows a comparison of the center displacements of a square microplate calculated using Eq. (57) with the experimental results given by Francais and Dufour (1999). It can be observed that there is an agreement between them. Figure 25 shows that the theoretical results using our reduced-order model are



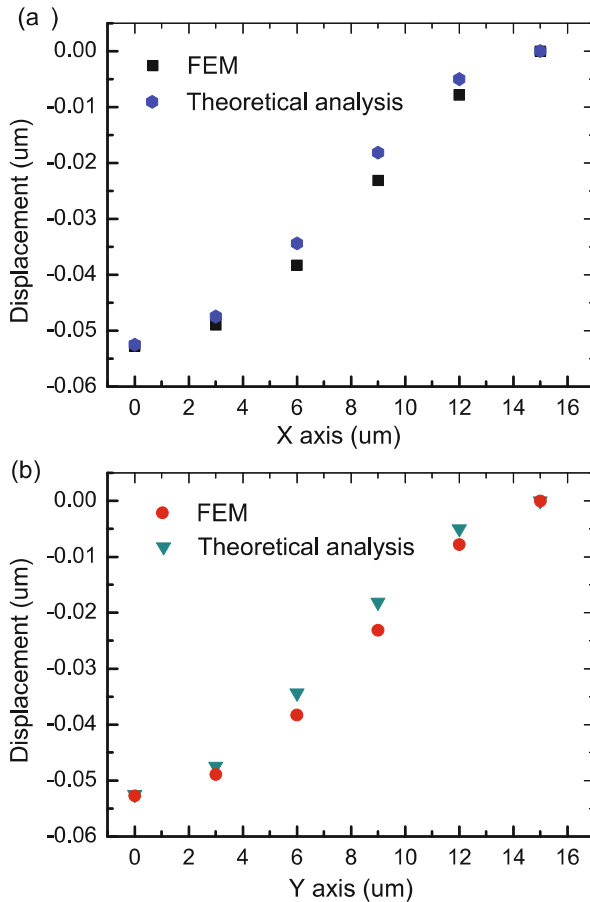
**Fig. 22** Displacements of the center point of a square microplate under different bias voltages; (a) comparison between the displacements obtained by FEM simulations and those calculated using Eq. (59); (b) relative errors of the calculated displacements with those from FEM simulations



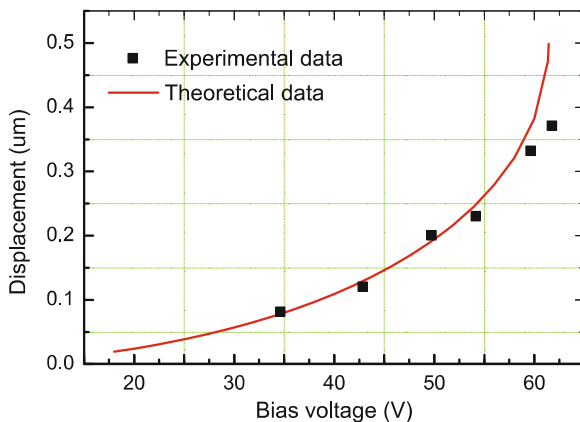
consistent with those calculated values using step-by-step linearized method given by Talebian (Talebian et al. 2010). These comparisons further demonstrate the robustness of our theoretical model in predicting the deflection of the microplate even if the deflection reaches up to the pull-in position. The geometry and material parameters for Figs. 24 and 25 are shown in Table 7.

Figure 26 shows the results of another case study of a larger microplate with edge length of 500 µm, that is,  $2a = 2b = 500 \mu\text{m}$ . The other geometry and material parameters are the same as given in Table 7. The deflections were obtained under the voltages changing from 4 V to 15.2 V which very approaches the pull-in voltage, 15.7 V. As shown in Fig. 26, the calculated deflections show good agreement with the results by nonlinear FEM simulations over the large deflection range, from the undeformed state to that approaching the pull-in position. As the cases of smaller microplates shown in Figs. 20 and 22, the maximum difference of the theoretical results with respect to those by the FEM simulations can also be controlled within

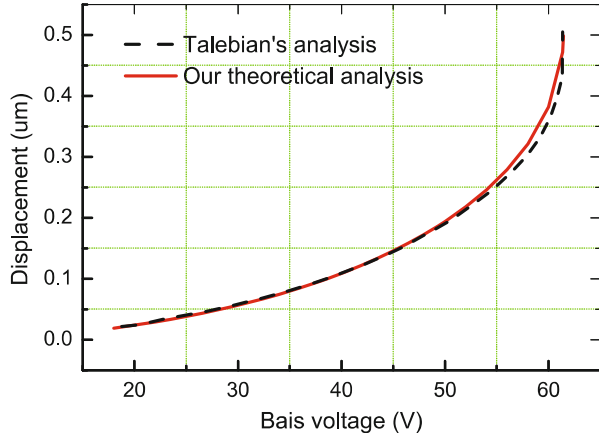
**Fig. 23** Comparison of the displacements of the square microplate along the axis directions under a bias voltage of 60 V; (a) displacements along the  $x$ -axis direction; (b) displacements along the  $y$ -axis direction



**Fig. 24** Comparison of center displacements of the center point of a square microplate calculated using Eq. (57) with the experimental results of Francois and Dufour (1999)



**Fig. 25** Comparison of center displacements of a square microplate calculated using Eq. (57) with the theoretical results of Talebian (2010)



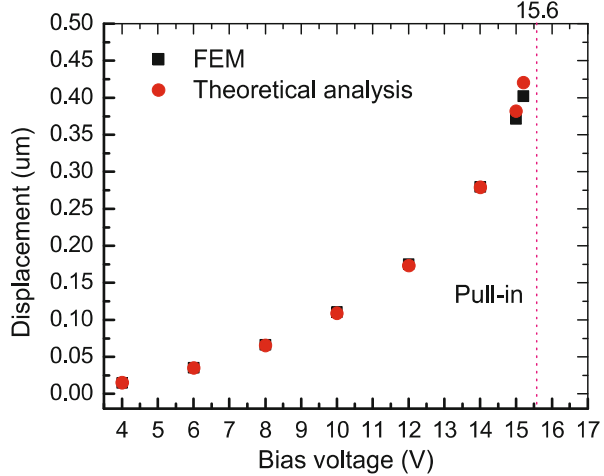
**Table 7** The material and geometry parameters used by Francais (Francais and Dufour 1999) and Talebian (2010)

Parameters	Value
Length $2a$	250 $\mu\text{m}$
Width $2b$	250 $\mu\text{m}$
Thickness $h$	3 $\mu\text{m}$
Poisson's ratio $\nu$	0.06
Separation distance $d$	1 $\mu\text{m}$
Young's modulus $E$	$1.69 \times 10^5$ MPa
Dielectric constant $\epsilon_0$	$8.854 \times 10^{-6}$ pF/ $\mu\text{m}$

the range of 5%. These results further demonstrated the availability of our theoretical model in predicting the static deflection of the electrostatically actuated microplates.

Having demonstrated the availability of the theoretical model for predicting the deflection of the microplate under different voltages, we sequentially validated its ability to predict the pull-in voltage. Table 8 shows a comparison of the calculated pull-in voltage of a square microplate with that from FEM simulations, for which the parameters are the same as that used for Fig. 22. The difference between the calculated and simulated results is as low as 0.45%. Table 9 shows a comparison of the pull-in voltage calculated using Eq. (60) with the result given by Talebian's work is as low as 0.49%. Figure 27 shows a comparison of the pull-in voltages calculated using Eq. (59) with those obtained by FEM simulations under different ratios of electrode separation distance  $d_0$  to plate thickness  $h$ . The parameters in Table 6 were used for the comparison, in which only the separation distance  $d_0$  was changed to form the different ratios. The calculated results are consistent with those from the FEM simulations. The relative errors between them are less than 1.0% for different ratios of  $d_0$  to  $h$ . These results demonstrate that the theoretical analyses can

**Fig. 26** A comparison of deflections obtained by theoretical calculation and FEM simulations for a large microplate with edge length of 500  $\mu\text{m}$



**Table 8** Comparison of the pull-in voltage calculated using Eq. (60) with that from FEM simulations

Theoretical analysis	FEM	Relative error
66.9 V	67.2 V	0.45%

**Table 9** Comparison of the pull-in voltage calculated using Eq. (60) with that from Talebian (2010)

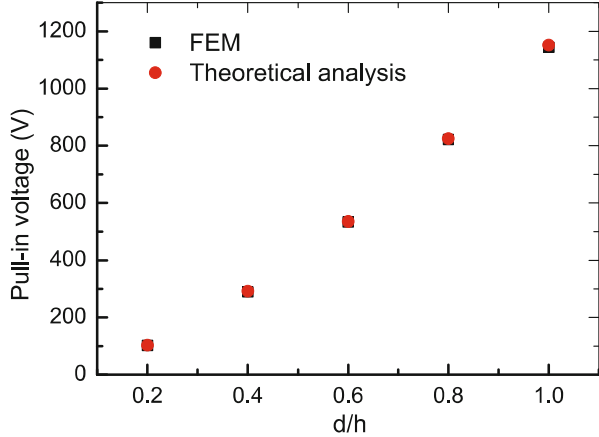
Theoretical analysis	Talebian’s result	Relative error
61.4 V	61.7 V	0.49%

predict the pull-in voltages of clamped rectangular and square microplates with a very high accuracy.

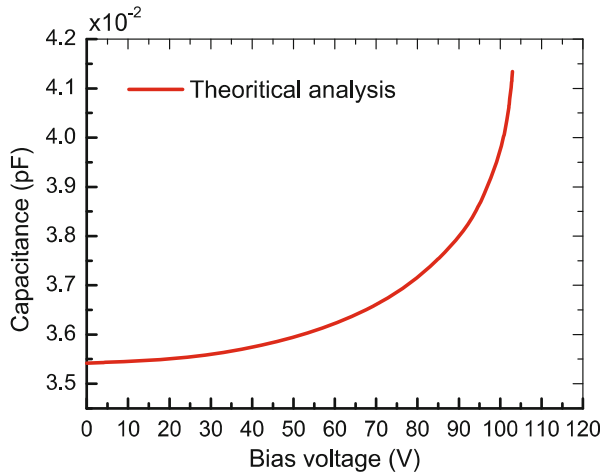
Figure 28 shows capacitance of a rectangular microplate with parameters given in Table 6 as a function of bias voltage. The theoretical value of the capacitance was calculated combining Eq. (61) and eight nodes Gauss-Legendre numerical integration. As can be seen from Fig. 28, the capacitance changes slow below about 80 V and changes fast above that, especially when the bias voltage approaches the pull-in voltage, 102.2 V. This trend of capacitance variation is consistent with that given by Ahmad and Pratap (2010). The capacitance of square microplates can also be calculated using Eq. (61) and Gauss-Legendre numerical integration. The accuracy of the quasi-analytical capacitance depends on both the deflection functions (Eqs. (58) and (59)) and accuracy of the numerical integration.

The aforementioned results are based on the lowest mode of the trial functions (i.e.,  $N = 0$ ). To further validate performance of the established reduced-order model in predicting the behavior of a microplate under electrostatic force, higher modes were also studied. Figure 29 shows comparison of the deflection curves of the

**Fig. 27** Comparison of pull-in voltages of the rectangular microplate with different ratios of electrode separation distance  $d_0$  to plate thickness  $h$



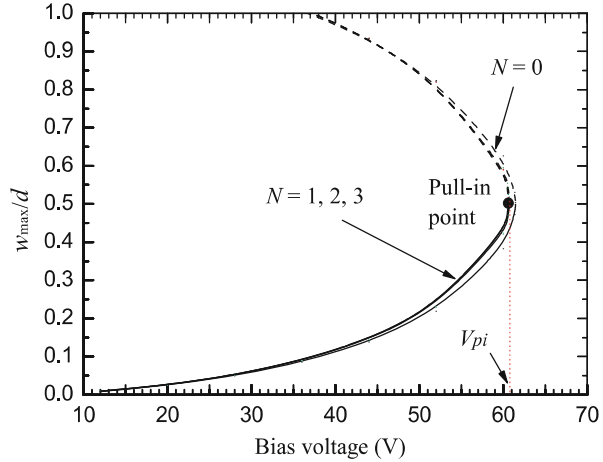
**Fig. 28** Capacitance variation of a rectangular microplate with bias voltages



microplate center point of a microplate for various numbers of modes  $N$ . The parameters used for the microplate are given in Table 7. As shown in Fig. 29, when  $N = 1, 2,$  and 3, the corresponding three deflection curves overlap each other. The deflection curves converge fast when the mode number  $N$  is more than zero. The modes with  $N = 1$  can be sufficient to model the static deflection of microplate over its stable physical range. However, the deflection curve for  $N = 0$  very approaches that for  $N = 1$ . The relative error of pull-in voltages between  $N = 0$  and  $N = 1$  is as low as 1.15%. So it is accurate enough to use the lowest-mode-based theoretical expression to predict the mechanical behavior of electrostatically actuated microplate. They are more efficient and convenient for design and optimization of electrostatically actuated microdevices than those higher-mode functions based on numerical solutions.

As the stretching effects of the microplate middle plane, which is discussed in detail in (Younis and Nayfeh 2007; Younis 2011), are not considered in our

**Fig. 29** Deflections of microplate center versus applied voltages with various  $N$



theoretical model, the analytical solutions mentioned above are applicable to the cases where stretching effects are negligible (or the deflection of the plate is smaller than its thickness) and where the electrode distance  $d_0$  is equal to or smaller than the microplate thickness  $h$  (these cases have been widely used in the electrostatically actuated configurations (Yaralioglu et al. 2003; Ahmad and Pratap 2010; Nabian et al. 2008; Vogl and Nayfeh 2005; Liao et al. 2009)). Based on these assumptions, Eqs. (56) and (57) can accurately predict the static deflection of rectangular and square microplates under electrostatic force even if the deflection is very close to the pull-in position. Equations (61) and (60) can accurately determine pull-in voltages of electrostatically actuated rectangular and square microplates, respectively. Equation (61) can be used to analyze capacitance variation and electromechanical coupling coefficient of electrostatic microdevices such as CMUTs (Yaralioglu et al. 2003). The theoretical model and these expressions are capable of widespread application to MEMS as a growing number of electrostatic microdevices require a high sensitivity, low energy consumption, and low pull-in voltage, for which the electrostatic microdevices are often required to have a small electrode separation distance.

## Dynamic Behavior Analysis

The vibration equation of the flexible microplate under a distributed electrostatic force can be written as

$$D\nabla w^4 + \rho h \frac{\partial^2 w}{\partial t^2} = F_e(r, t), \quad (62)$$

where  $\rho$  is the density of the plate and  $F_e(r, t)$  is time-dependent variable given as

$$F_e(r, t) = \frac{\varepsilon_0 V^2}{2[d_0 - w(r, t)]^2}, \quad (63)$$

In order to analyze the resonant frequency of the microplate under electrostatic force, the corresponding eigenvalue equation should be derived. We start the derivation by introducing the electrical excitation and deflection functions.

The electric excitation applied to the microplate is composed of a DC bias voltage and an AC voltage, which can be given by

$$V(t) = V_{dc} + V_{ac}(t), \quad (64)$$

where  $V(t)$  is the total electric excitation applied to the microplate which changes with time variable  $t$ ,  $V_{dc}$  is the DC bias voltage, and  $V_{ac}(t)$  is the AC voltage which is assumed as a harmonic signal. Generally, the amplitude of the AC voltage is much smaller than the value of the bias voltage.

As a result of the electric excitation, the microplate deflection can be correspondingly decomposed into an initial static deflection component caused by the DC bias voltage and a small vibration around the initial static deflection under the small AC voltage. Therefore, the deflection of the microplate can be written as

$$w(x, y, t) = w_{dc}(x, y) + w_{ac}(x, y, t), \quad (65)$$

where  $w_{dc}$  is the initial static deflection of the microplate under the DC bias voltage and  $w_{ac}(x, y, t)$  is the small harmonic vibration function around the initial static deflection.

Substituting Eqs. (63), (64), and (65) into (62) and simplifying the resulting Eq. (62) can be rewritten as

$$D\nabla^4 w_{dc} + D\nabla^4 w_{ac} + \rho h \frac{\partial^2 w_{ac}}{\partial t^2} = \frac{\varepsilon_0 [V_{dc} + V_{ac}]^2}{2\{d_0 - [w_{dc} + w_{ac}]\}^2}. \quad (66)$$

Expanding the electrostatic force term in Eq. (66) at the static equilibrium position ( $w_{dc}$ ,  $V_{dc}$ ) using the Taylor series yields

$$\begin{aligned} D\nabla^4 w_{dc} + D\nabla^4 w_{ac} + \rho h \frac{\partial^2 w_{ac}}{\partial t^2} &= \frac{\varepsilon_0 V_{dc}^2}{2(d_0 - w_{dc})^2} + \frac{\varepsilon_0 V_{dc}^2}{(d_0 - w_{dc})^3} w_{ac} \\ &+ \frac{\varepsilon_0 V_{dc}}{(d_0 - w_{dc})^2} V_{ac} + \frac{3\varepsilon_0 V_{dc}^2}{2(d_0 - w_{dc})^4} w_{ac}^2 + \frac{2\varepsilon_0 V_{dc}}{(d_0 - w_{dc})^3} V_{ac} w_{ac} \\ &+ \frac{\varepsilon_0}{2(d_0 - w_{dc})^2} V_{ac}^2 \dots \end{aligned} \quad (67)$$

Equation (67) can be decoupled into an equation of static deflection of the microplate under the bias voltage and an equation of motion to describe the dynamic behavior around the initial static deflection. The equation for the static deflection is

$$D \left( \frac{\partial^4 w_{dc}}{\partial x^4} + 2 \frac{\partial^4 w_{dc}}{\partial x^2 \partial y^2} + \frac{\partial^4 w_{dc}}{\partial y^4} \right) = \frac{\epsilon_0 V_{dc}^2}{2(d_0 - w_{dc})^2}, \quad (68)$$

and the equation for the small vibration is

$$\begin{aligned} D \nabla^4 w_{ac} + \rho h \frac{\partial^2 w_{ac}}{\partial t^2} &= \frac{\epsilon_0 V_{dc}^2}{(d_0 - w_{dc})^3} w_{ac} + \frac{\epsilon_0 V_{dc}}{(d_0 - w_{dc})^2} V_{ac} + \frac{3\epsilon_0 V_{dc}^2}{2(d_0 - w_{dc})^4} w_{ac}^2 \\ &+ \frac{2\epsilon_0 V_{dc}}{(d_0 - w_{dc})^3} V_{ac} w_{ac} + \frac{\epsilon_0}{2(d_0 - w_{dc})^2} V_{ac}^2 \dots \end{aligned} \quad (69)$$

Generally, the amplitude of the applied AC voltage is much smaller than the value of the DC bias voltage; thus, a small-amplitude motion is excited around the initial deflection. This means  $V_{ac} \ll V_{dc}$  and  $w_{ac} \ll w_{dc}$ . Therefore, using the first-order Taylor expansion of the electric force, Eq. (69) can be further simplified as

$$D \nabla^4 w_{ac} + \rho h \frac{\partial^2 w_{ac}}{\partial t^2} = \frac{\epsilon_0 V_{dc}^2}{(d_0 - w_{dc})^3} w_{ac} + \frac{\epsilon_0 V_{dc}}{(d_0 - w_{dc})^2} V_{ac}. \quad (70)$$

It can be seen from Eq. (70) that the two terms on the right-hand side are related to the electrostatic force. The term with  $w_{ac}$  represents the variation of the electrostatic force induced by the small vibration around the initial deflection, which is a special intrinsic condition for the microplate under the DC bias voltage. However, the term with  $V_{ac}$  is actually an electric excitation which should be considered as an external load. Eliminating the term with the AC voltage,  $V_{ac}$ , in Eq. (70), the eigenvalue equation for the resonant frequency (natural frequency) of the microplate under the DC voltage is derived as

$$D \nabla^4 w_{ac} + \rho h \frac{\partial^2 w_{ac}}{\partial t^2} = \frac{\epsilon_0 V_{dc}^2}{(d_0 - w_{dc})^3} w_{ac}. \quad (71)$$

This eigenvalue equation has a different form from the one without any initial loads. An additional term appears representing the effect of the DC bias voltage on the resonant frequency. Equation (71) indicates that the resonant frequency (or natural frequency) of the electrostatically actuated microplate is not only a function of the common geometry and material properties of the microplate but also is a function of the applied DC voltage.

In order to solve Eq. (71), the small vibration function was assumed to be (Li et al. 2013)

$$w_{ac} = \eta(r) \sin(\omega t + \varphi) \quad (72)$$

where  $\eta(r)$  denotes the vibration mode function of the clamped plate and satisfies the boundary conditions as  $w_0(r)$ ,  $\omega$  is the fundamental angular frequency of the



microplate with initial deflection  $w_0(r)$ ,  $t$  is the time variable, and  $\varphi$  is the initial phase. Substituting Eqs. (72) into Eq. (71), the small variation of electrostatic force can be written as

$$\Delta Q(r, t) = \frac{\varepsilon_0 V_{dc}^2}{[d_0 - w_{dc}(r)]^3} \eta(r) \sin(\omega t + \varphi) \quad (73)$$

and Eq. (71) can be simplified as

$$D\nabla^4 \eta(r) - \omega^2 \rho h \eta(r) = \Delta Q(r) \quad (74)$$

where

$$\Delta Q(r) = \frac{\varepsilon_0 V^2}{[d_0 - w_0(r)]^3} \eta(r). \quad (75)$$

It is difficult to solve Eq. (74) because of its right nonlinear term  $\Delta Q(r)$ , especially for the nonlinear coefficient  $\varepsilon_0 V_{dc}^2 / [d_0 - w_0(r)]^3$ . Here, an approximate method was proposed to give an equivalent deflection  $w_{\text{eff}}$  instead of  $w_0(r)$  and turn the nonlinear coefficient into  $\varepsilon_0 V_{dc}^2 / [d_0 - w_{\text{eff}}]^3$ . An approximate vibration mode function  $\eta_1(r)$  was given instead of  $\eta(r)$  in Eq. (75). It can be written as

$$\eta_1(r) = C_1 (R^2 - r^2)^2 \quad (76)$$

where  $C_1$  is a constant. This approximate function also satisfies all the boundary conditions as  $w_{dc}(r)$ . Then the small variation of electrostatic force  $\Delta Q(r)$  in Eq. (75) can be rewritten as

$$\Delta Q(r) \approx \frac{\varepsilon_0 V^2}{[d_0 - w_{dc}(r)]^3} \eta_1(r). \quad (77)$$

and the equivalent electrostatic force variation can be written as

$$\Delta Q_{\text{eff}}(r) = \frac{\varepsilon_0 V^2}{(d_0 - w_{\text{eff}})^3} \eta_1(r). \quad (78)$$

Then let the work done by  $\Delta Q(r)$  in Eq. (77) on the plate within the small deviation  $\eta_1(r)$  be equal to that of  $\Delta Q_{\text{eff}}(r)$ . The corresponding expression is

$$\int_0^R \int_0^{2\pi} \Delta Q_{\text{eff}}(r) \eta_1(r) r dr d\theta = \int_0^R \int_0^{2\pi} \Delta Q(r) \eta_1(r) r dr d\theta. \quad (79)$$

Substitute Eqs. (76), (77), and (78) into Eq. (79) to solve this integration, the equivalent deflection  $w_{\text{eff}}$  was obtained:

$$w_{\text{eff}} = d_0 - \sqrt[3]{\frac{R^{10}}{10N_1}} \quad (80)$$

where

$$N = \frac{5CR^6 - 3d_0R^2}{16C^2(d_0 - CR^4)^2} + \frac{3}{32\sqrt{d_0}C^5} \ln \frac{\sqrt{d_0} + \sqrt{CR^2}}{\sqrt{d_0} - \sqrt{CR^2}}, \quad (81)$$

and

$$C = \frac{84Dd_0^2 + d_0R^4P - QR^4 - B}{128Dd_0R^4}. \quad (82)$$

Then substituting Eq. (80) into Eq. (74) and replacing the term  $w_0(r)$ , a simplified equation can be obtained:

$$\nabla^4 \eta(r) - \beta^4 \eta(r) = 0 \quad (83)$$

where

$$\beta^4 = \omega^2 \frac{\rho h}{D} + \frac{10\varepsilon_0 V_{\text{dc}}^2 N}{DR^{10}}. \quad (84)$$

It is obvious that Eq. (83) has the same form as the undeformed plate and an approximate solution can be obtained using the same method (Soedel 2004). Due to the fundamental (the first-order) resonant frequency studied, the value of  $\beta R$  is chosen to be 3.196, that is,  $\beta = 3.196/R$  (Soedel 2004). Substituting  $\beta$  into Eq. (84), the fundament angular frequency was obtained:

$$\omega = \sqrt{\frac{104.334D}{\rho h R^4} - \frac{10\varepsilon_0 V_{\text{dc}}^2 N}{\rho h R^{10}}}. \quad (85)$$

And the resonant frequency  $f$  can be expressed as

$$\begin{aligned} f &= f_0 \sqrt{1 - \frac{\varepsilon_0 V_{\text{dc}}^2 N}{10.4334DR^6}} \\ &= f_0 \sqrt{1 - \frac{\varepsilon_0 V_{\text{dc}}^2}{10.4334DR^6} \left[ \frac{5CR^6 - 3d_0R^2}{16C^2(d_0 - CR^4)^2} + \frac{3}{32\sqrt{d_0}C^5} \ln \frac{\sqrt{d_0} + \sqrt{CR^2}}{\sqrt{d_0} - \sqrt{CR^2}} \right]} \end{aligned} \quad (86)$$

where

$$f_0 = \frac{10.21}{2\pi R^2} \sqrt{\frac{D}{\rho h}}. \quad (87)$$

When used for low or ultra-low pressure sensing, Eq. (85) can be linearized using Maclaurin's theorem. It can be written as

$$f = f_0(a - bP) \quad (88)$$

where  $a$  and  $b$  are functions of the applied voltage and their expressions are

$$a = \sqrt{1 - 0.0958 \times \frac{\varepsilon_0 V_{\text{bias}}^2}{DR^6} \left[ \frac{5CR^6 - 3d_0R^2}{16C^2(d_0 - CR^4)^2} + \frac{3}{32\sqrt{d_0}C^5} \ln \frac{\sqrt{d_0} + \sqrt{CR^2}}{\sqrt{d_0} - \sqrt{CR^2}} \right]} \quad (89)$$

and

$$b = 0.0479 \times \frac{\varepsilon_0 d V_{\text{bias}}^2 (84Dd_0^2 - QR^4 + B)}{aDC^2B(84Dd_0^2 - QR^4 - B)} \left[ \frac{(5CR^4 - d_0)CR^4}{16(d_0 - CR^4)^3} + \frac{(15d_0 - 23CR^4)}{32(d_0 - CR^4)^2} - \frac{15}{64\sqrt{d_0}CR^2} \ln \frac{\sqrt{d_0} + \sqrt{CR^2}}{\sqrt{d_0} - \sqrt{CR^2}} \right] \quad (90)$$

where

$$C_1 = \frac{84Dd_0^2 - QR^4 - B_1}{128Dd_0R^4} \quad (91)$$

and

$$B_1 = \sqrt{7056D^2d_0^4 - 504Dd_0^2R^4Q + R^8Q^2} \quad (92)$$

The pressure sensitivity can be defined as (Kagawa 1974)

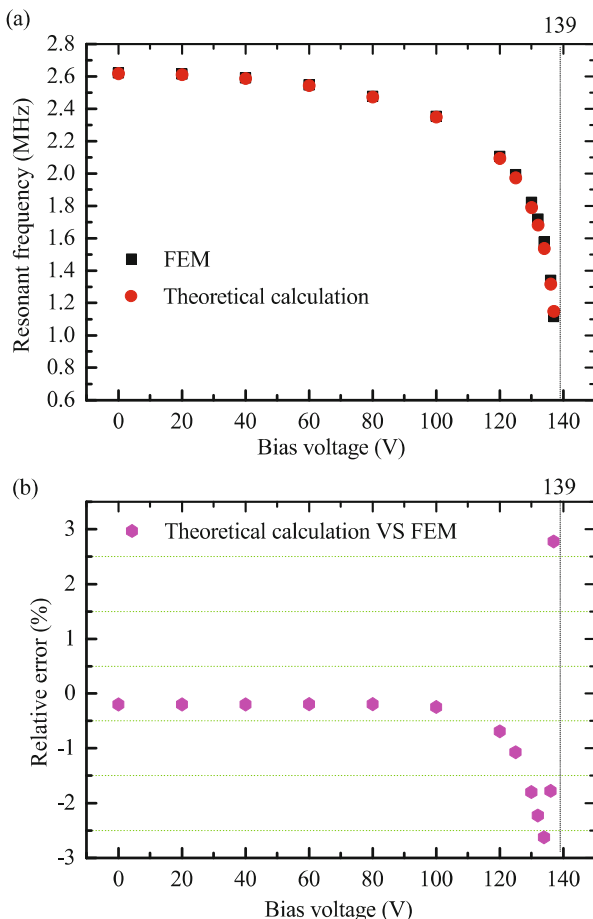
$$S_P = \frac{1}{af_0} \frac{\partial f}{\partial P} = -\frac{b}{a} \quad (\text{ppm/Pa}). \quad (93)$$

In order to validate the analytical solutions obtained above, FEM simulations were carried out using a commercially available FEM package. The detailed material and geometry parameters used for the FEM simulations are shown in Table 1.

A comparison of the results of analytical calculations and FEM simulations under different voltages with a fixed zero pressure is presented in Fig. 30. The calculated resonant frequencies by Eq. (86) are compared with the simulated ones in Fig. 30a. The resonant frequencies obtained by the analytical calculation agree well with those obtained by the FEM simulation. Both results show that the resonant frequency decreases with the voltage, which is same as (Liao et al. 2009). Figure 30b shows the relative errors of the calculated data with respect to the simulated data in Fig. 30a, b. It can be seen that the relative errors are very small within almost all working range

**Fig. 30** A comparison of the resonant frequencies obtained by theoretical calculation and FEM simulations.

(a) Calculated and simulated results. (b) Error analysis between theoretical calculation and FEM simulations

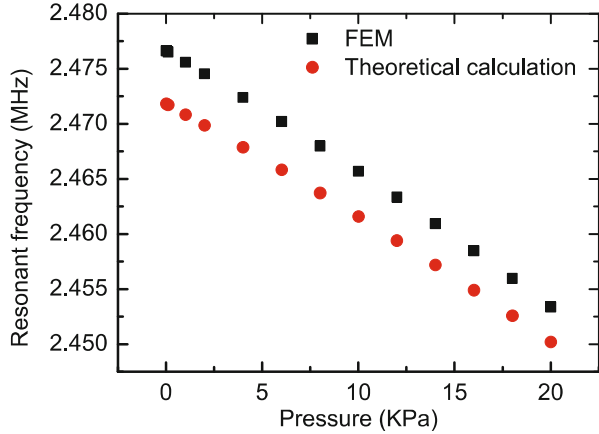


of the applied bias voltage (from zero to 137 V, 96% of the collapse voltage) and the maximum relative error is smaller than 3%. So Eq. (86) can predict the effect of the electrostatic force on the resonant frequency of a microplate more accurately.

Figure 31 shows the effect of uniform hydrostatic pressure on the resonant frequency of microplate under an applied voltage of 80 V. As shown in Fig. 31, both analytical and numerical results show that the resonant frequency of the microplate decreases with the pressure under the bias voltage. Good agreement between the analytical and numerical results is observed, and the maximum error is just 0.19% within the pressure range of 0 to 20 kPa. These results indicate that Eq. (86) can well analyze the resonant frequency changes of the microplates under both electrostatic force and hydrostatic pressure.

Figure 32 shows a comparison of the pressure-induced relative frequency shifts obtained by Eq. (88) and FEM simulations. It can be seen that the frequency shifts vary linearly with the pressure in the pressure range from 0 to 10 kPa. The resonant frequencies have a linear relationship with the pressure change.

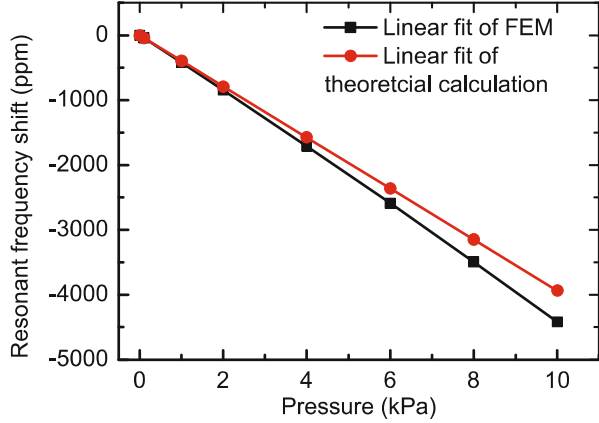
**Fig. 31** The effect of uniform hydrostatic pressure on the resonant frequency with the applied voltage of 20 V. The red circular and the black square dots denote the data obtained by Eq. (86) and FEM simulations, respectively



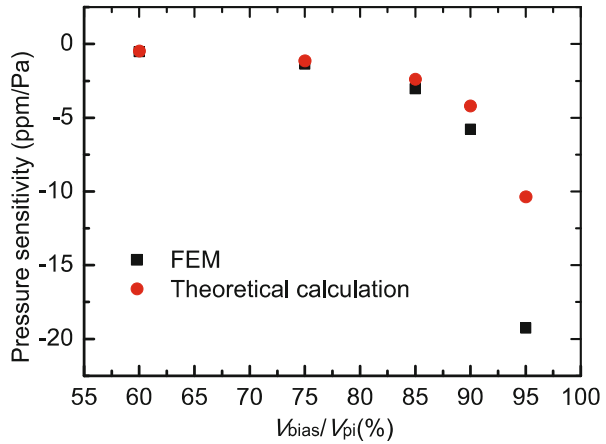
The pressure sensitivities from the theoretical calculation and FEM simulation are  $-0.52$  ppm/Pa and  $-0.64$  ppm/Pa, respectively. The relative error is less than 5%. These results demonstrate that Eq. (88) can well predict the resonant frequency changes under different uniform pressures.

Though the results and discussions above demonstrate that the derived expressions can well predict the resonant frequencies of microplate under different voltages and uniform pressures, a further insight into these expressions is quite necessary for their applications. Firstly, the derived expressions are valid only when the maximum displacement of flexible microplate under two types of loads simultaneously satisfies the following two conditions. Owing to the assumptions of small deflections and the use of the classical Kirchhoff thin plate theory, the maximum displacement should be smaller than the microplate thickness (Osterberg 1995). To avoid the occurrence of pull-in phenomenon of the flexible microplate (Vogl and Nayfeh 2005; Li et al. 2015a; Kerrou and Hobar 2006), the maximum displacement should also be less than about half of the separation distance  $d_0$  between the upper flexible and lower plates. So, the two limits should be comprehensively considered when these expressions are used. Secondly, as shown in Figs. 31 and 32, the resonant frequency holds a linear relationship with the low or ultra-low pressure. It suggests a potential application of these expressions in the design of electrostatically actuated pressure sensors. Additionally, it can be derived from Eq. (93) that the corresponding pressure sensitivity depends on the applied voltage. The pressure sensitivity can be adjusted by the applied voltage. This phenomenon is clearly shown in Fig. 33. Both results of analytical calculations and FEM simulations show that the sensitivity is changed by the applied voltages within the same pressure range of 0 to 10 kPa. The pressure sensitivity increases with the applied voltage. Finally, it should be mentioned that the strain of the microplate middle plane was neglected in our analysis due to small deflection assumptions. Though the derived approximate solutions were well validated by FEM simulations, a more accurate analysis taking into account the effect of the middle plane strain of the middle plane might be meaningful.

**Fig. 32** The relative frequency shift vs the uniform hydrostatic pressure under the applied voltage of 80 V. The black square and the red circular dots denote the data obtained by FEM simulations and Eq. (88), respectively



**Fig. 33** The pressure sensitivities under different voltages within the same pressure range of 0 to 60 Pa. The black square and the red circular dots denote the data obtained by the FEM simulation and Eq. (93), respectively



## Conclusions and Problems

This chapter illustrates the finite element modeling, lumped electromechanical modeling, and distributed electromechanical modeling methods.

The use of finite elements for microplates and electrostatic domain and establishment of the finite element electromechanical model are discussed in the finite element modeling method. Examples of 2D axisymmetrical model and 3D models for electrostatically actuated microplates are given.

A lumped model for the electrostatically actuated microplates is given, based on which the collapse position is calculated out as one-third of the air gap. The collapse voltage and resonant frequency of the microplates are qualitatively analyzed. The simplified vibration equation shows the resonant frequency of microplate decreases with applied DC voltage.

For the distributed modeling, a method to approximate electrostatic force is proposed, and the electromechanical coupling models for electrostatically actuated circular and rectangular microplates under electrostatic force and hydrostatic pressure are established. Based on these models, theoretical expressions for collapse voltage, static deflection, and capacitance variation are proposed. These expressions illustrate the effects of structure parameters, applied electrostatic force and pressure on the collapse voltages, static deflections, and capacitances of the electrostatically actuated circular and rectangular microplates. Finite element method (FEM) simulations are carried out to validate these theoretical analyses. The applicable dimension ranges of the theoretical analyses are given by parametric studies. The results indicate that these theories are more applicable to electrostatically actuated microplates with the membrane thickness-to-dimension ratio less than  $1/20$  and membrane thickness-to-electrode distance ratio larger than 1; the errors of deflection and collapse voltage with respect to FEM simulations are less than 5% and 2.5%, respectively, within the whole bias voltage variation range (from 0% to 98% of the collapse voltage).

Energy equivalent method is proposed to accomplish the microplate deflection equivalence, and expressions of electrostatically actuated microplates under only electrostatic force and the coaction of electrostatic force and hydrostatic pressure are established, which reveal the variation regulation of the resonant frequency under only electrostatic force and both electrostatic force and hydrostatic pressure. The expression accuracies are well validated by FEM simulations.

As the stretch of electrostatically actuated microplates is not considered in the distributed electromechanical models, the theories are applicable to the cases that the microplate deflection is smaller than its thickness. Additionally, the resonant frequency of electrostatically actuated rectangular microplates still needs further study since closed-form expression is not established. Other typical distributed electromechanical modeling methods such as step-by-step linearized method and Galerkin method-based reduced-order modeling methods for electrostatically actuated microplates can be seen in Ahmad and Pratap (2010), Nabian et al. (2008), Chao et al. (2006), Younis et al. (2003), Vogl and Nayfeh (2005), Zhao et al. (2004), Liao et al. (2009), and Talebian et al. (2010).

---

## References

- Ahmad B, Pratap R (2010) Elasto-electrostatic analysis of circular microplates used in capacitive micromachined ultrasonic transducers[J]. *IEEE Sensors J* 10(11):1767–1773
- Bozkurt A, Ladabaum I, Atalar A et al (1999) Theory and analysis of electrode size optimization for capacitive microfabricated ultrasonic transducers[J]. *IEEE Trans Ultrason Ferroelectr Freq Control* 46(6):1364–1374
- Chao PCP, Chiu CW, Tsai CY (2006) A novel method to predict the pull-in voltage in a closed form for micro-plates actuated by a distributed electrostatic force[J]. *J Micromech Microeng* 16(5):986
- Chiou DY, Chen MY, Chang MW et al (2008) Finite element modeling, characterization, and optimization design for the polymer-typed capacitive micro-arrayed ultrasonic transducer[J]. *Microsyst Technol* 14(6):787–797

- Elgamel HE (1995) Closed-form expressions for the relationships between stress, diaphragm deflection, and resistance change with pressure in silicon piezoresistive pressure sensors[J]. *Sensors Actuators A Phys* 50(1–2):17–22
- Elgamel HEA (1999) A simple and efficient technique for the simulation of capacitive pressure transducers[J]. *Sensors Actuators A Phys* 77(3):183–186
- Fragiacomo G, Ansbæk T, Pedersen T et al (2010) Analysis of small deflection touch mode behavior in capacitive pressure sensors[J]. *Sensors Actuators A Phys* 161(1):114–119
- Francais O, Dufour I (1999) Normalized abacus for the global behavior of diaphragms: pneumatic, electrostatic, piezoelectric or electromagnetic actuation[J]. *J Model Simul Microsyst* 2(1):149–160
- Kagawa Y (1974) Vibrator sensors for atmospheric pressure, temperature, and humidity measurements[J]. *J Acoust Soc Am* 56(5):1644–1649
- Kerrou F, Hobar F (2006) A novel numerical approach for the modeling of the square shaped silicon membrane[J]. *Semicond Phys Quantum Electron Optoelectron* 9(4):52–57
- Ladabaum I, Jin X, Soh HT et al (1998) Surface micromachined capacitive ultrasonic transducers [J]. *IEEE Trans Ultrason Ferroelectr Freq Control* 45(3):678–690
- Li ZK, Zhao LB, Ye ZY et al (2013) Resonant frequency analysis on an electrostatically actuated microplate under uniform hydrostatic pressure[J]. *J Phys D Appl Phys* 46(19):195108
- Li ZK, Zhao LZ, Jiang ZD et al (2015a) An improved method for the mechanical behavior analysis of electrostatically actuated microplates under uniform hydrostatic pressure[J]. *J Microelectromech Syst* 24(2):474–485
- Li ZK, Zhao LB, Jiang ZD et al (2015b) Mechanical behavior analysis on electrostatically actuated rectangular microplates[J]. *J Micromech Microeng* 25(3):035007
- Liao LD, Chao PCP, Huang CW et al (2009) DC dynamic and static pull-in predictions and analysis for electrostatically actuated clamped circular micro-plates based on a continuous model[J]. *J Micromech Microeng* 20(2):025013
- Nabian A, Rezazadeh G, Haddad-derafshi M et al (2008) Mechanical behavior of a circular micro plate subjected to uniform hydrostatic and non-uniform electrostatic pressure[J]. *Microsyst Technol* 14(2):235–240
- Nayfeh AH (2000) *Nonlinear interactions: analytical, computational and experimental methods*[M]. New York, NY, Wiley
- Nayfeh AH, Mook DT (1979) *Nonlinear oscillations*[M]. New York, Wiley
- Nayfeh AH, Younis MI, Abdel-Rahman EM (2005) Reduced-order models for MEMS applications [J]. *Nonlinear Dyn* 41(1):211–236
- Osterberg PM. (1995) Electrostatically actuated microelectromechanical test structures for material property measurement[D]. Massachusetts Institute of Technology
- Pelesko JA, Bernstein DH (2002) *Modeling Memes and Nems*[M]. Boca Raton (FL): Chapman & Hall, CRC Press
- Rahman M, Chowdhury S (2010) A new deflection shape function for square membrane CMUT design[C]. In: *IEEE international symposium on circuits and systems(ISCAS)*, Paris, 30 May–02 June, pp 2019–2022
- Rahman MM, Chowdhury S (2011) Square diaphragm CMUT capacitance calculation using a new deflection shape function[J]. *J Sensors*, 2011
- Shahiri-Tabarestani M, Ganji BA, Sabbaghi-Nadooshan R (2012) Design and simulation of high sensitive capacitive pressure sensor with slotted diaphragm[C]. In: *IEEE international conference on biomedical engineering (ICoBE)*, Penang, 27–28 Feb, pp 484–489
- Soedel W (2004) *Vibrations of shells and plates*[M]. Marcel Dekker, New York, CRC Press
- Talebani S, Rezazadeh G, Fathalilou M et al (2010) Effect of temperature on pull-in voltage and natural frequency of an electrostatically actuated microplate[J]. *Mechatronics* 20(6):666–673
- Timoshenko S, Woinowsky-Krieger S (1959) *Theory of plates and shells*[M]. New York, London, McGraw-Hill Book Company, Inc.
- Ventsel E, Krauthammer T (2001) *Thin plates and shells: theory: analysis, and applications*[M]. New York : Marcel Dekker, CRC Press



- Vogl GW, Nayfeh AH (2005) A reduced-order model for electrically actuated clamped circular plates[J]. *J Micromech Microeng* 15(4):684
- Yaralioglu GG, Ergun AS, Bayram B et al (2003) Calculation and measurement of electromechanical coupling coefficient of capacitive micromachined ultrasonic transducers[J]. *IEEE Trans Ultrason Ferroelectr Freq Control* 50(4):449–456
- Younis MI (2011) MEMS linear and nonlinear statics and dynamics[M]. Boston, MA, Springer Science & Business Media
- Younis MI, Nayfeh AH (2007) Simulation of squeeze-film damping of microplates actuated by large electrostatic load[J]. *J Comput Nonlinear Dyn* 2(3):232–241
- Younis MI, Abdel-Rahman EM, Nayfeh A (2003) A reduced-order model for electrically actuated microbeam-based MEMS[J]. *J Microelectromech Syst* 12(5):672–680
- Zhao X, Abdel-Rahman EM, Nayfeh AH (2004) A reduced-order model for electrically actuated microplates[J]. *J Micromech Microeng* 14(7):900
- Zhou MX, Huang QA, Qin M (2005) Modeling, design and fabrication of a triple-layered capacitive pressure sensor[J]. *Sensors Actuators A Phys* 117(1):71–81



# Behavior Modeling and Simulation of an Inertial Sensor

Honglong Chang, Jianbing Xie, and Ying Liu

## Contents

Introduction .....	156
Parametric Model Order Reduction (PMOR) Method Based on Implicit Moment Matching .....	157
Parametric Model Order Reduction for a MEMS Capacitive Accelerometer .....	157
Modeling and Simulation for a MEMS Vibratory Gyroscope .....	165
Trajectory Piecewise-Linear (TPWL) Method .....	180
Thermoelectric Governing Differential Equations .....	180
Trajectory Piecewise-Linear Method .....	181
Compact Model Extraction and Result Analysis .....	183
Conclusions .....	191
References .....	191

## Abstract

Two macromodeling techniques for inertial sensors are discussed in this chapter. The first one is a parametric model order reduction (PMOR) method based on the implicit moment matching to accommodate the parameter variation. The second one is the trajectory piecewise-linear (TPWL) method which is developed for dealing with the strong nonlinearity. For each technique, its effectiveness is demonstrated by the applications to read devices characterization.

## Keywords

MEMS · Sensor · Simulation · Model order reduction

H. Chang (✉) · J. Xie  
MOE Key Laboratory of Micro/Nano Systems for Aerospace, Northwestern Polytechnical University, Xi'an, China  
e-mail: [changhl@nwpu.edu.cn](mailto:changhl@nwpu.edu.cn)

Y. Liu  
School of Arts and Sciences, Shaanxi University of Science and Technology, Xi'an, China

## Introduction

MEMS inertial sensors including gyroscopes and accelerometers are used for angular and linear motion measurement, which have been the two killer applications in MEMS field. The modeling and simulation for complex MEMS inertial sensors has been the major motive for developing efficient computer-aided design tools. Thermal, electrical, and environmental actions should be accounted for in a fully coupled multiphysics modeling of the sensors (Chang et al. 2010). Usually, MEMS inertial sensors are described by a set of complex motion equations. Direct solution of these equations based on fully meshed models is computationally intensive and hard to perform (Clark and Pister 2007). Modeling and simulation based on low order behavioral models is more widely accepted. There mainly exist two kinds of low order models: numerical macromodel and parametrized lumped-element behavioral model.

The modeling and simulation based on parametrized lumped-element models (Clark and Pister 2007; Mukherjee et al. 2000; Tilmans 1996; Veijola et al. 1995; Zhou and Agogino 2002) is an efficient method for MEMS inertial sensors with regular geometry. In this approach, the inertial sensor is divided into a few components like beams, proof masses, comb-finger capacitors with simple or regular geometrical shapes, then their behavioral models are deduced according to classical physical theory such as matrix structural analysis, structural dynamics (Clark and Pister 2007; Vandemeer 1998; Coventor Inc 2008). The distinguished advantage of the methodology is that the models are parameterized thus are reusable by modifying relevant parameters, which is especially important for rapid system-level synthesis and optimization. However, the lumped element behavioral models are hard to achieve when MEMS structures have high geometric complexity (Chen et al. 2005). For the components with eccentric geometrical shape, the behavior model will be very hard to achieve analytically. In other words, the method has poor extensibility and does not support well the innovation design for some novel MEMS or components. Some significant improvements (Clark and Pister 2007; Coventor Inc 2008; Jing et al. 2002; Zhang et al. 2005) have been achieved for enriching the lumped-element model library. Many models of complex components or clusters such as perforated beam, perforated comb drive, serpentine beam, eccentric shaped beam or mass have been developed in order to expand the supporting scope of the lumped-element model library. Coventor (2008) almost developed the most abundant model library including some nonideal components such as arc beam, curve comb drive, beam actuator, L-shaped beam. This greatly facilitates the MEMS design and expands the supporting scope of the methodology. However, the limited extension of the model library cannot meet the requirement of the limitless creativity of designers. Multifarious MEMS devices may need some different models which do not exist in the model library and are usually unable to be forecasted.

The modeling and simulation (Zhang et al. 2008; Yang and Yen 2004; Reitz et al. 2003; Schrag and Wachutka 2002, 2004; Bedyk et al. 2008; Nayfeh et al. 2005; Chatterjee and Aluru 2005) based on numerical macromodels is another popular system-level modeling and simulation method. The numerical macromodels are

usually extracted from a series of FEA runs and can be directly inserted into circuit or system-level simulators for complete MEMS simulation. The numerical macro-modeling strategy is the base of the methodology, and many macromodeling techniques have been developed during the past two decades.

In this chapter, two distinctive macromodeling techniques will be discussed. The first one is a parametric model order reduction (PMOR) method based on the implicit moment matching to accommodate the parameter variation (Chang et al. 2009, 2010). For instance, in the reduction process of gyroscope system, not only the material density, Young's modulus, Rayleigh damping coefficient, angle velocity, but also the coefficient of thermal expansion and the change in temperature are preserved. The second one is the recently proposed trajectory piecewise-linear (TPWL) method (Liu et al. 2014a, b). This kind of method is constructed for dealing with the strong nonlinearity and avoiding the difficulties rising from the well-known proper orthogonal decomposition (POD) method.

---

## **Parametric Model Order Reduction (PMOR) Method Based on Implicit Moment Matching**

In order to simulate and validate the performance of MEMS devices in a short time, so as to ameliorate the design scheme, these extracted macromodels need to be parameterized. But the traditional model order reduction (MOR) methods are incapable of dealing with the parametric macromodels.

Recently, some parametric MOR (PMOR) methods have been proposed to preserve the parameter during the MOR process. In this section, a PMOR method based on implicit moment matching is introduced by using two examples including a MEMS capacitive accelerometer and a MEMS Vibratory Gyroscope.

### **Parametric Model Order Reduction for a MEMS Capacitive Accelerometer**

Squeeze film damping is very important for oscillating MEMS devices since it mainly determines their dynamical behaviors. In order to consider the damping effects into system-level behavior simulation, researchers have presented many analytic damping models which can be easily inserted into system-level simulators. However, these analytic models are restricted to simple regular geometries. Unfortunately, the structures of most MEMS devices are complex and irregular. In addition, perforations are widely used to control gas damping or reduce the release etching time. The squeeze film damping in perforated surfaces with arbitrarily-shaped geometries is not easy for modeling using analytic method, while it is usually modeled, simulated, and analyzed accurately through the finite element method (FEM). However, the numerical FEM analysis process is very time-consuming when the geometries of perforated devices are complex due to the high dimensionality of the finite element model. Furthermore, the models are

hard to be directly cosimulated with the interface circuit in the system-level simulators.

The modified Reynolds equation (Yang and Yu 2002) accounting for perforation effect in the squeeze film damping is

$$\frac{h_0}{P_0} \frac{\partial p}{\partial t} = \frac{h_0^3}{12\mu} \left( \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} \right) + \frac{p}{Z} + f \frac{\partial d}{\partial t} \quad (1)$$

where  $h_0$  is the gap separation;  $\mu$  is the dynamic viscosity;  $p$  is the pressure change;  $P_0$  is the ambient pressure;  $d$  is the variation of the displacement;  $f$  is the coefficient about shape of the displacement;  $Z$  is the flow impedance of the hole which is a function of the radius  $r$  and length  $th$  of the hole and the viscosity  $\mu$ . If considering the slip flow boundary conditions, an effective viscosity  $\mu_{\text{eff}}$  is introduced as  $\mu_{\text{eff}} = \mu / (1 + 9.638K_n^{1.159})$ ,  $K_n = P_{\text{ref}} mfp / P_0 h_0$ , where  $K_n$  is Knudsen number,  $mfp$  is mean free path length and the  $P_{\text{ref}}$  is reference pressure for the mean free path. If the moving structure is rigid and moves in a translational motion, then  $f = 1$ . If in a torsional motion, then  $f = x$  and  $d = \varphi$ , where  $\varphi$  is the tilting angle. Two cases of squeeze film damping of perforated plates when they are sinusoidally oscillating in normal translational motions or torsional motions are shown in Fig. 1.

Equation (1) can be discretized in space by FEM to a state space form as

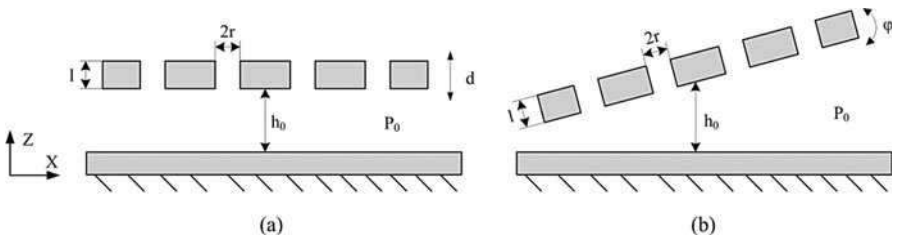
$$\frac{h_0}{p_0} \mathbf{E} \frac{d\mathbf{p}}{dt} = \frac{h_0^3}{12\mu} \mathbf{A}_D \mathbf{p} + \frac{1}{Z} \mathbf{A}_Z \mathbf{p} + \mathbf{B} \mathbf{u}(t) \quad \mathbf{y} = \mathbf{L}^T \mathbf{p} \quad (2)$$

where  $\mathbf{p}$  is the unknown pressure vector.  $\mathbf{E}$ ,  $\mathbf{A}_D$ ,  $\mathbf{A}_Z$  are the system matrix;  $\mathbf{B}$  is the input matrix, and  $\mathbf{L}$  is the output matrix. The matrix  $\mathbf{A}_Z$  contains the perforations location. In fact, after the discretization,  $\mathbf{B} = \mathbf{L} = \mathbf{E}f$ .

Furthermore, Eq. (2) can be rewritten as

$$k_1 \mathbf{E} \frac{d\mathbf{p}}{dt} = (\mathbf{A}_D + k_2 \mathbf{A}_Z) \mathbf{p} + k_3 \mathbf{B} \mathbf{u}(t) \quad \mathbf{y} = \mathbf{L}^T \mathbf{p} \quad (3)$$

where  $k_1 = 12\mu / (h_0^2 P_0)$ ,  $k_2 = 12\mu / (h_0^3 Z)$ ,  $k_3 = 12\mu / (h_0^3)$ . If we take the conventional MOR method to reduce order of the model described by Eq. (3), then the parameter  $k_2$  must be chosen as a constant number and kept unchanged, and the



**Fig. 1** Squeeze film damping of perforated plates (a) translational motion (b) torsional motion

system matrix  $\mathbf{A}_D$  and  $\mathbf{A}_Z$  can be merged into one matrix to compute projection matrix spanning the Krylov subspace. However, the parameter  $k_2$  cannot be preserved. To preserve the parameter, PMOR method must be applied. At the same time, the parameter  $k_1$  and  $k_2$  can also be preserved.

The PMOR method is based on moment matching via Krylov subspace projection to generate a low-dimensional orthogonal matrix  $\mathbf{V}$  which does not depend on parameters in system (3). After the Laplace transformation, system (3) becomes

$$k_1 s \mathbf{E} \mathbf{p}(s) = (\mathbf{A}_D + k_2 \mathbf{A}_Z) \mathbf{p}(s) + k_3 \mathbf{B} \mathbf{U}(s) \quad \mathbf{Y}(s) = \mathbf{L}^T \mathbf{p}(s) \quad (4)$$

Then, the transfer function for Eq. (3) can be written as

$$\mathbf{H}(s) = \frac{\mathbf{Y}(s)}{\mathbf{U}(s)} = k_3 \mathbf{L}^T (k_1 s \mathbf{E} - \mathbf{A}_D - k_2 \mathbf{A}_Z)^{-1} \mathbf{B} \quad (5)$$

Firstly, expanding the transfer function in Taylor series about  $s = 0$ , Eq. (5) can be expressed as

$$\begin{aligned} \mathbf{H}(s) &= -k_3 \mathbf{L}^T \left[ \mathbf{I} - k_1 (\mathbf{A}_D + k_2 \mathbf{A}_Z)^{-1} \mathbf{E} s \right]^{-1} (\mathbf{A}_D + k_2 \mathbf{A}_Z)^{-1} \mathbf{B} \\ &= -k_3 \mathbf{L}^T \sum_{j=0}^{\infty} \left[ k_1 (\mathbf{A}_D + k_2 \mathbf{A}_Z)^{-1} \mathbf{E} \right]^j (\mathbf{A}_D + k_2 \mathbf{A}_Z)^{-1} \mathbf{B} s^j \\ &= -k_3 k_1^j \mathbf{L}^T \sum_{j=0}^{\infty} \tilde{\mathbf{M}}_j s^j \end{aligned} \quad (6)$$

where  $\mathbf{M}_j = -k_3 k_1^j \mathbf{L}^T \tilde{\mathbf{M}}_j$ ,  $j = 0, 1, \dots$ , are the Taylor series coefficients of  $\mathbf{H}(s)$ , known as the moments of  $\mathbf{H}(s)$ .  $k_3 k_1^j$ ,  $j = 0, 1, \dots$ , are the coefficients before each matrix  $\tilde{\mathbf{M}}_j$ . Hence, the subspace spanned by the columns of  $\tilde{\mathbf{M}}_j$  does not depend on the parameter  $k_3$  and  $k_1$ . In the same way, in order to preserve the parameter  $k_2$ , each  $\tilde{\mathbf{M}}_j$  can be further expanded in Taylor series about  $k_2 = 0$ .

$$\begin{aligned} \tilde{\mathbf{M}}_j &= \sum_{i_j=0}^{\infty} \sum_{i_{j-1}=0}^{\infty} \cdots \sum_{i_1=0}^{\infty} \sum_{i_0=0}^{\infty} \left[ (-\mathbf{A}_D^{-1} \mathbf{A}_Z)^{i_j} \mathbf{A}_D^{-1} \mathbf{E} (-\mathbf{A}_D^{-1} \mathbf{A}_Z)^{i_{j-1}} \mathbf{A}_D^{-1} \mathbf{E} \cdots \right. \\ &\quad \left. (-\mathbf{A}_D^{-1} \mathbf{A}_Z)^{i_1} \mathbf{A}_D^{-1} \mathbf{E} (-\mathbf{A}_D^{-1} \mathbf{A}_Z)^{i_0} \mathbf{A}_D^{-1} \mathbf{B} \right] k_2^{i_j} k_2^{i_{j-1}} \cdots k_2^{i_1} k_2^{i_0} \end{aligned} \quad (7)$$

It can be seen from Eq. (7) that  $k_2^m$ ,  $m = 1, 2, \dots, j$  are the coefficients before each matrix

$$(-\mathbf{A}_D^{-1} \mathbf{A}_Z)^{i_j} \mathbf{A}_D^{-1} \mathbf{E} (-\mathbf{A}_D^{-1} \mathbf{A}_Z)^{i_{j-1}} \mathbf{A}_D^{-1} \mathbf{E} \cdots (-\mathbf{A}_D^{-1} \mathbf{A}_Z)^{i_1} \mathbf{A}_D^{-1} \mathbf{E} (-\mathbf{A}_D^{-1} \mathbf{A}_Z)^{i_0} \mathbf{A}_D^{-1} \mathbf{B} \quad (8)$$

where  $i_j, i_{j-1}, \dots, i_1, i_0 = 0, 1, \dots, \infty$ . Therefore,  $\tilde{\mathbf{M}}_j$  actually lies in subspace spanned by the columns of the matrices in Eq. (8) that are independent

of the parameter  $k_2$ . Then the projection matrix  $\mathbf{V}$  can be computed with the PMOR method to use the matrices (8). Given  $\mathbf{V}$ , low-order model can be generated by projecting Eq. (3) as follows:

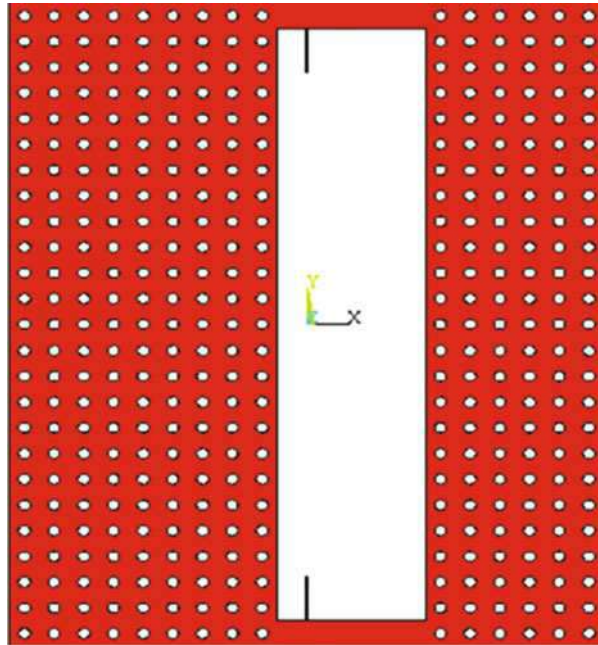
$$k_1 \mathbf{V}^T \mathbf{E} \mathbf{V} \frac{d\mathbf{p}}{dt} = (\mathbf{V}^T \mathbf{A}_D \mathbf{V} + k_2 \mathbf{V}^T \mathbf{A}_Z \mathbf{V}) \mathbf{p} + k_3 \mathbf{V}^T \mathbf{B} u(t) \quad \mathbf{y} = \mathbf{L}^T \mathbf{V} \mathbf{p} \quad (9)$$

Finally, reduced-order model as shown in Eq. (9), which still captures the dynamics of the original model with small error, can be obtained with the parameter  $k_1, k_2, k_3$ .

A MEMS capacitive accelerometer (Fig. 2) is chosen to verify the low-order parametric damping model. The perforated proof-mass is supported by two torsional beams. Because the two masses of the accelerometer are dissymmetrical, it will rotate once there is an acceleration input along the z axis. Then the acceleration can be sensed differentially using the capacitive electrodes below the masses.

The accelerometer's parameters were set as follows for simulation. The plate thickness is  $80 \mu\text{m}$  and the gas gap is  $3 \mu\text{m}$ . The total surface area of the plate is  $2500 \times 2000 \mu\text{m}^2$ . The area of bigger mass is  $2500 \times 900 \mu\text{m}^2$  and the smaller one is  $2500 \times 600 \mu\text{m}^2$ . The total number of perforations is  $25 \times 15$ . The distance between every two hole-centers is  $60 \mu\text{m}$ , and the radius of each hole is  $20 \mu\text{m}$ . The accelerometer's resonance frequency is 1027 Hz according to simulation. When the accelerometer's closed-loop interface circuit is applied, its measurement scale is  $\pm 22 \text{ g}$ , and its maximal tilting angle is  $3\text{e-}5 \text{ rad}$  which is very small. Therefore, the

**Fig. 2** Schematic of the torsional capacitive accelerometer



torsional squeezed film damping of the accelerometer can be approximately solved by the linearized Reynolds equation.

In simulations, the finite element model is built by ANSYS, system matrices are extracted by MOR4ANSYS, and reduced order model is simulated by MATLAB. The order of the original damping model is 10,261 while the order of reduced parametric macromodel is only 60. The comparisons of the simulation results based on the models with different values are shown in Table 1. The relative error for harmonic response was estimated as follows:

$$\text{error} = | \mathbf{A} - \hat{\mathbf{A}} | / | \mathbf{A} | \tag{10}$$

where  $\mathbf{A}$  is the amplitude of harmonic solution by the original model and  $\hat{\mathbf{A}}$  is the amplitude of harmonic solution by the reduced order model.

The relative error for transient response was estimated as follows (Rudnyi et al. 2005):

$$\text{error} = \left[ \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right]^{1/2} / \left[ \sum_{i=1}^n y_i^2 \right] \tag{11}$$

where  $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$  is the transient solution by the original model and  $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)^T$  is the transient solution by the reduced order model.

It can be seen from Table 1 that the five parameters of squeeze film damping influence the errors between the reduced and original damping model in a different

**Table 1** Simulation results for reduced macromodel as compared with the original model

Parameter values					Original harmonic solution		Error (%) in harmonic	Error (%) in transient
$P_0(\text{MPa})$	$\mu(\text{kg}/\mu\text{m}/\text{s})$	$mfp(\mu\text{m})$	$h_0(\mu\text{m})$	$th(\mu\text{m})$	Amplitude ( $\mu\text{N}\mu\text{m}$ )	Phase (rad/s)		
0.1	2.0e-11	0.1253	3	80	1.2513e + 3	-0.0012	0.0159	1.2603e-6
	3.21e-11	0.1097			2.0635e + 3	-0.0019	0.0163	7.8363e-7
	1.83e-11	0.064			1.2749e + 3	-0.0012	0.0175	1.3620e-6
<b>0.1</b>	<b>1.83e-11</b>	<b>0.064</b>	<b>5</b>	<b>80</b>	<b>323.6001</b>	<b>-4.839e-4</b>	<b>0.0034</b>	<b>1.0316e-6</b>
					579.5127	-7.069e-4	0.0078	1.3257e-6
					1.2749e + 3	-0.0012	0.0175	1.3620e-6
					3.9576e + 3	-0.0025	0.0401	1.0045e-6
					<b>2.6403e + 4</b>	<b>-0.0082</b>	<b>0.1142</b>	<b>4.2912e-7</b>
<b>0.1</b>	<b>1.83e-11</b>	<b>0.064</b>	<b>1</b>	40	2.6380e + 4	-0.0082	0.1644	6.1794e-7
				20	2.6368e + 4	-0.0082	0.2432	9.1401e-7
				<b>10</b>	<b>2.6363e + 4</b>	<b>-0.0082</b>	<b>0.3781</b>	<b>1.4210e-6</b>
0.01	1.83e-11	0.064	3	80	576.0111	-0.0053	0.0065	1.1172e-6
0.001					81.3536	-0.0070	5.4010e-4	6.5827e-7
0.0001					7.7945	-0.0065	6.3991e-5	8.1418e-7
0.00001					0.6556	-0.0054	1.4679e-5	1.1105e-5
<b>0.00001</b>	<b>1.83e-11</b>	<b>0.064</b>	<b>1</b>	40	2.4124	-0.0073	0.0048	1.9679e-4
				<b>10</b>	<b>2.3067</b>	<b>-0.0071</b>	<b>0.0086</b>	<b>3.7001e-4</b>



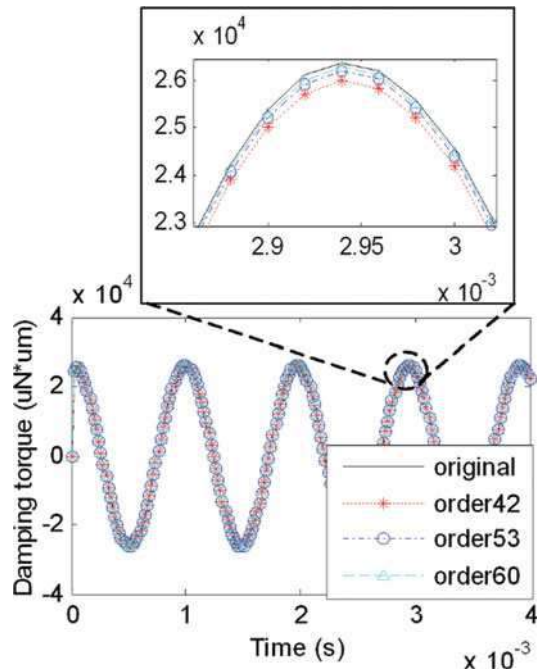
way. The parameters  $\mu$  and  $mfp$  are dependent on the gas property and their variation does not cause much error. In fact, the errors are mainly affected by the parameters  $P_0$ ,  $h_0$ , and  $th$ . And it can be found from Table 1 that the errors in the harmonic analysis is proportional to the value of parameter  $P_0$  and inverse to the parameters  $h_0$  and  $th$ . The main reason lies in that parameter  $k_2$  in parametric MOR process. The parameter  $k_2$  is mainly determined by the gap separation  $h_0$  and the flow impedance  $Z$ . When values of the parameter  $h_0$  and  $Z$  decrease, the parameter  $k_2$  will increase, and this will cause the error to augment.

In order to intuitively show the influences of these parameters, the four different cases from Table 1 are selected to draw the plot of the transient responses as shown in Figs. 3, 4, 5 and 6, respectively. The transient responses of the parametric damping macromodel with different order 42, 53, and 60 are also shown in the figure for comparison. We can find that the higher order of macromodel the better simulation accuracy can be obtained, although the differences are very small.

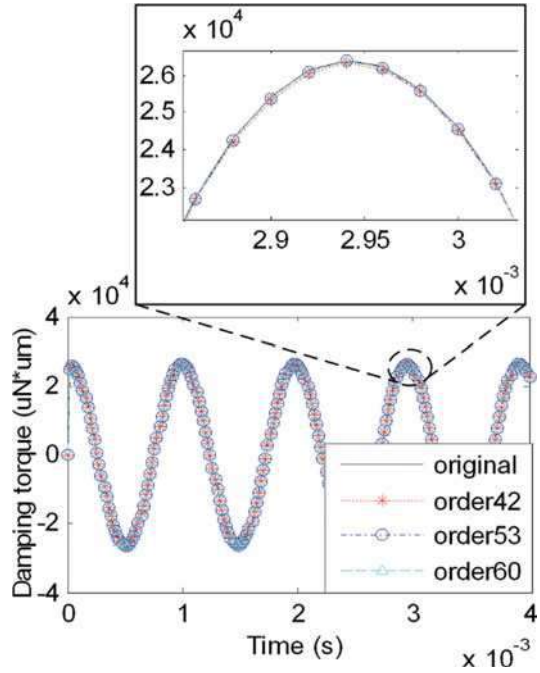
Figure 7 shows the error behavior of the harmonic response over a wide range of  $th$ , when  $P_0 = 0.1$ ,  $\mu = 1.83e-11$ ,  $mfp = 0.064$ ,  $h_0 = 1$ . In fact, the scale of the flow impedance reaches  $1e-14$  when  $th$  changes from 10 to 120, but the maximum error of the model with 60 dimensions is only about 0.3%. The computation time for transient analysis are shown in Table 2. It is obvious that the low order damping model dramatically decreases the simulation time.

The presented PMOR method successfully extracted the parametric macromodels of squeeze film damping with perforations. The five parameters such as the flow

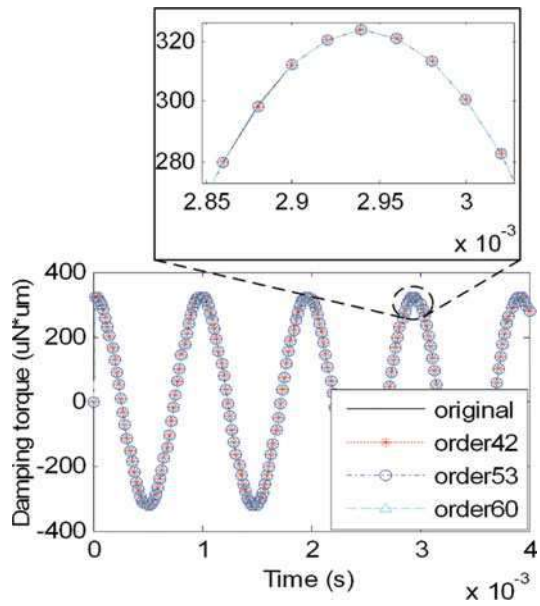
**Fig. 3** Transient response of the original model and the reduced model of order 42, 53, 60 for the sinusoidal input function ( $P_0 = 0.1$ ,  $\mu = 1.83e-11$ ,  $mfp = 0.064$ ,  $h_0 = 1$ ,  $th = 10$ )



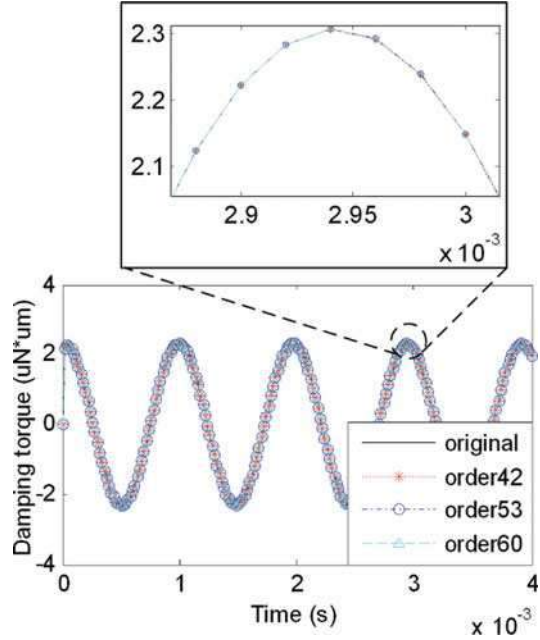
**Fig. 4** Transient response of the original model and the reduced model of order 42, 53, 60 for the sinusoidal input function ( $P_0 = 0.1$ ,  $\mu = 1.83e-11$ ,  $mfp = 0.064$ ,  $h_0 = 1$ ,  $th = 80$ )



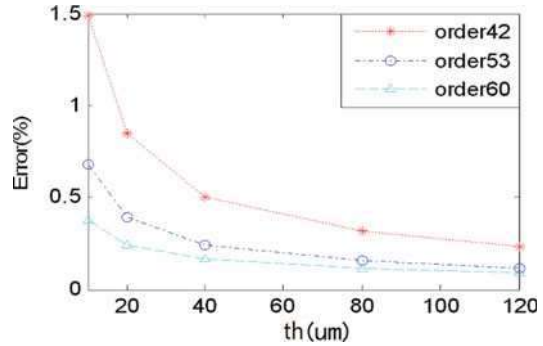
**Fig. 5** Transient response of the original model and the reduced model of order 42, 53, 60 for the sinusoidal input function ( $P_0 = 0.1$ ,  $\mu = 1.83e-11$ ,  $mfp = 0.064$ ,  $h_0 = 5$ ,  $th = 80$ )



**Fig. 6** Transient response of the original model and the reduced model of order 42, 53, 60 for the sinusoidal input function ( $P_0 = 0.00001$ ,  $\mu = 1.83e-11$ ,  $mfp = 0.064$ ,  $h_0 = 1$ ,  $th = 10$ )



**Fig. 7** Error with different values of thickness



**Table 2** Comparison of computation time for transient analysis between original model and reduced-order model

	Computation time of original model	Computation time of reduced order models		
		Order 42	Order 53	Order 60
Transient simulation	185.5 s	15.7 s	16.1 s	16.3 s
Speedup	—	12	11	11

impedance of perforations, the viscosity, the ambient pressure, the plate thickness, and the gap separation are all preserved in the MOR process and packaged into one macromodel. The parametric macromodel could be reused conveniently like a common parametric component. From the simulation results, we can see that the

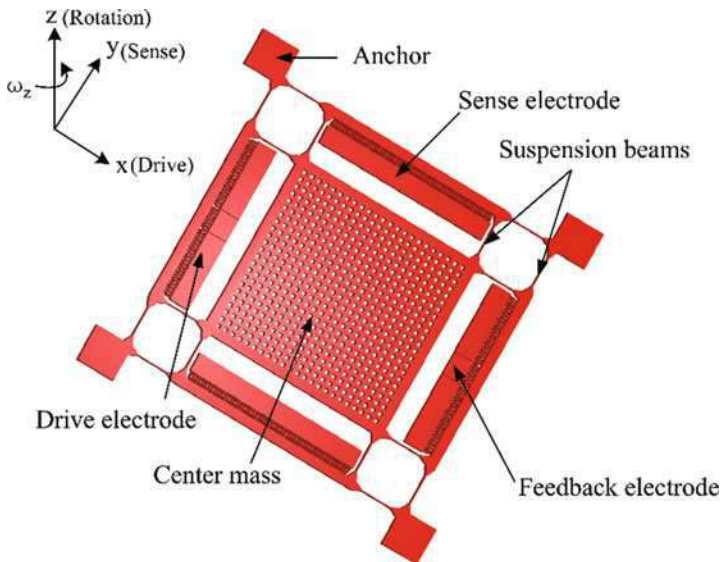
parametric macromodel could produce solutions as accurate as that computed directly from the original large model in a much shorter time. Additionally, the macromodel realized the accurate cosimulation damping effects with the interface circuit.

## Modeling and Simulation for a MEMS Vibratory Gyroscope

In this part, the parametric model order reduction approach is applied to generate a parametric low order model for a gyroscope. The model can directly connect with the external circuit to predict performance of the entire gyroscope system. In the reduction process, not only the material density, Young's modulus, Rayleigh damping coefficient, angle velocity but also the coefficient of thermal expansion and the change in temperature are preserved. So, the structural performance impacted by environmental thermal fluctuations can be simulated in a fast and accurately way as compared with the FEM (Painter and Shkel 2001, 2003; Liu et al. 2008). In addition, we find that the expansion in the angular velocity direction is unnecessary. The simulation results show that the reduced order model can accurately and efficiently approximate the original gyroscope device when the parameters change at certain range. Then the cosimulations with interface circuit demonstrate the feasibility of performing simulations of a sophisticated model under closed-loop control.

### Introduction of the MEMS Gyroscope

Figure 8 shows a typical z-axis MEMS gyroscope with symmetrical structure. The gyroscope is driven in the x-direction by an electrostatic force generated by a



**Fig. 8** Schematic structure of the microgyroscope

**Table 3** Parameters of the gyroscope

Parameters	Values
Young's modulus	$1.3 \times 10^{11}$ Pa
Poisson's ratio	0.22
Coefficient of thermal expansion	$2.6 \times 10^{-3}/^\circ\text{C}$
Temperature coefficient of Young's modulus	$7.5 \times 10^{-5}/^\circ\text{C}$
Density	$2.33 \times 10^3$ kg/m <sup>3</sup>
Suspension beam length	$500 \times 10^{-6}$ m
Suspension beam width	$10 \times 10^{-6}$ m
Suspension beam thickness	$65 \times 10^{-6}$ m
Center mass length	$2000 \times 10^{-6}$ m
Center mass width	$2000 \times 10^{-6}$ m
Finger width	$5 \times 10^{-6}$ m
Finger length	$60 \times 10^{-6}$ m
Finger gap	$3 \times 10^{-6}$ m
Finger overlap	$35 \times 10^{-6}$ m
Comb-to-substrate gap	$30 \times 10^{-6}$ m

DC-biased AC voltage applied across the comb actuators. The Coriolis force produced by an angular velocity around the z-axis causes the proof-mass to oscillate in the y-direction. The output signals are then detected from the capacitance change of comb-fingers in the sense direction. Table 3 shows the parameters of the gyroscope.

### Vibration Equation of the Microgyroscope

Considering damping, temperature, electrostatic, Coriolis, and centrifugal effects simultaneously, the vibration equation with multidegree of freedom for the microgyroscope is as follows:

$$\begin{aligned} \rho \mathbf{M} \ddot{\mathbf{x}}(t) + (\mathbf{D} + \omega \rho \mathbf{G}) \dot{\mathbf{x}}(t) + (\mathbf{EK} + ETC_h \Delta T \mathbf{S}) \mathbf{x}(t) \\ = \mathbf{F}_e + \mathbf{F}_t + \mathbf{F}_r \mathbf{y}(t) = \mathbf{L}^T \mathbf{x}(t) \end{aligned} \quad (12)$$

where  $t$  is the time variable,  $\mathbf{x}(t)$  is the vector of the state variables.  $\mathbf{M}$ ,  $\mathbf{D}$ ,  $\mathbf{G}$ ,  $\mathbf{K}$ , and  $\mathbf{S}$  are the mass, damping, Coriolis, stiffness, and temperature stress stiffening matrices, respectively.  $\mathbf{F}_e$ ,  $\mathbf{F}_t$ , and  $\mathbf{F}_r$  are electrostatic force vector, thermal force vector, and centrifugal force vector.  $\mathbf{L}$  is the output matrix and  $\mathbf{y}(t)$  is the output vector. The parameters  $\rho$ ,  $\omega$ ,  $E$ ,  $TCh$ , and  $\Delta T$  are the material density, angle velocity around z-axis, Young's modulus, the coefficient of thermal expansion and the change in temperature. The damping matrix  $\mathbf{D}$  is often modeled as Rayleigh damping, which is linear combination of mass and stiffness matrices,  $\mathbf{D} = \alpha \mathbf{M} + \beta \mathbf{K}$ . The matrix  $\mathbf{S}$  is obtained by nonlinear static analysis with prestress effect.

In the following, the damping, electrostatic, temperature, and centrifugal effects will be discussed separately.

### Damping Effect

Usually, the MEMS damping effects can be divided into two types: slide film damping and squeeze film damping (Senturia 2001). A slide film damping occurs when two parallel plates are in relatively lateral motion; whereas a squeeze film damping occurs when two parallel plates are in vertical motion.

In the gyroscope, the plates are actuated laterally by the electrostatic comb-drive. Because of the Coriolis force, the plates will oscillate in the y-direction (parallel to the substrate). Thus, the slide film damping will be involved in the gyroscope’s dynamic model. Figure 9 shows some of the energy dissipative processes in the lateral moving gyroscope. Typical applications are damping between the fingers of comb drives and horizontally-moving seismic masses.

Couette flow assumes a constant velocity gradient across the fluid gap. Tangentially moving surfaces at low frequencies produce a nearly constant velocity gradient in the fluid, so Couette flow is applicable to low frequencies. The damping coefficient is (Veijola and Turowski 2001)

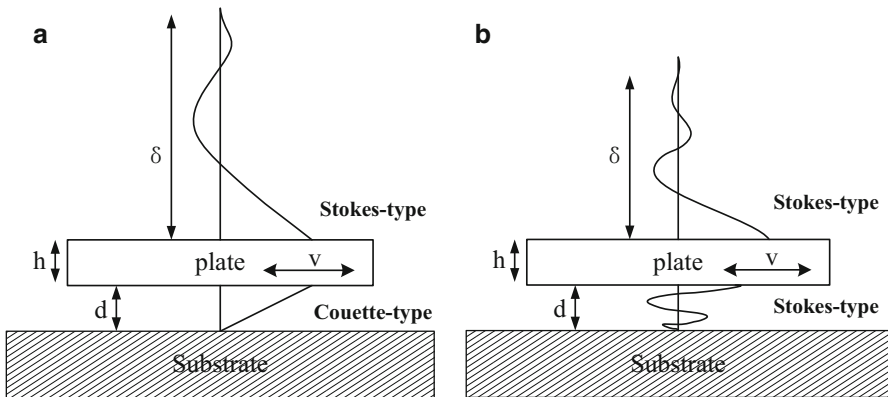
$$c_d = \frac{\mu A}{d} \tag{13}$$

where  $\mu$  is the viscosity of the fluid,  $A$  is the surface area, and  $d$  is gap height.

Stokes flow assumes that the velocity gradient is not constant across the fluid gap. Tangentially moving surfaces at high frequencies do not produce a constant velocity gradient, so the Stokes flow is applicable to high frequencies. The transition from Couette to Stokes flow occurs near the cut-off frequency  $f_c$ .

$$f_c = \frac{\mu}{2\pi\rho d^2} \tag{14}$$

where  $\rho$  is the fluid density.



**Fig. 9** Dissipative processes in the lateral moving microgyroscope (a) at low frequency (b) at high frequency

The damping effects come from both the ambient fluid on the top of the moving plate and the fluid between the plates. The Stokes-type damping effect in the ambient fluid is (Cho et al. 1994)

$$c_d = \mu A \beta \quad (15)$$

The Stokes-type damping effect underneath the plate is

$$c_d = \mu A \beta \frac{\sinh(2\beta d) + \sin(2\beta d)}{\cosh(2\beta d) - \cos(2\beta d)} \quad (16)$$

where  $\beta = \sqrt{\pi f / \nu}$ ,  $f$  is the oscillation frequency, and  $\nu$  is the kinetic viscosity.

In Fig. 9, the penetration depth  $\delta$  is given by  $\delta = 1/\beta$ . When the operating pressure is reduced, the gas rarefaction must be included in the damping model by applying various boundary conditions such as the first-order or high-order slip boundary conditions. The behavior of the gas will change due to the gas rarefaction effect, so the effective viscosity is introduced as (Veijola and Turowski 2001)

$$\mu_{eff} = \frac{\mu}{1 + 2K_n} \quad (17)$$

where  $K_n$  is the Knudsen number and the ratio of the mean free path  $\lambda$  and gap height  $d$ :

$$K_n = \frac{\lambda}{d} \quad (18)$$

In the drive and sense mode, the slide film damping coefficient is the sum of the damping due to the proof mass and the damping due to the comb fingers.

$$c_{slide} = c_{mass} + c_{combs} \quad (19)$$

Then the damping ratio is

$$\xi_i = c_{slide} / 2\omega_i m_i \quad (20)$$

where  $\omega_i$  is the eigenfrequency of the  $i$ th modal and the  $m_i$  is the  $i$ th modal mass.

In order to calculate the Rayleigh damping coefficient  $\alpha$  and  $\beta$ , two frequencies which cover the concerned modes are often selected (Chowdhury and Dasgupta 2003). Here the drive frequency  $\omega_1$  and the sense frequency  $\omega_2$  which are the first two modes are chosen and the corresponding damping ratios can be obtained by Eq. (20). Then the Rayleigh damping coefficient can be solved out from

$$\begin{aligned} \alpha/2\omega_1 + \beta\omega_1/2 &= \xi_1 \\ \alpha/2\omega_2 + \beta\omega_2/2 &= \xi_2 \end{aligned} \quad (21)$$

That are,

$$\alpha = \frac{2(\omega_1\omega_2^2\xi_1 - \omega_1^2\omega_2\xi_2)}{\omega_2^2 - \omega_1^2} = \frac{c_{\text{slide}}(\omega_2^2/m_1 - \omega_1^2/m_2)}{\omega_2^2 - \omega_1^2} \tag{22}$$

$$\beta = \frac{2(\omega_1\xi_1 - \omega_2\xi_2)}{\omega_1^2 - \omega_2^2} = \frac{c_{\text{slide}}(1/m_1 - 1/m_2)}{\omega_1^2 - \omega_2^2}$$

**Electrostatic Effect**

The electrostatic force generated by a comb-drive actuator is

$$F_e = \frac{1}{2}V(t)^2 \frac{\partial C}{\partial x} \tag{23}$$

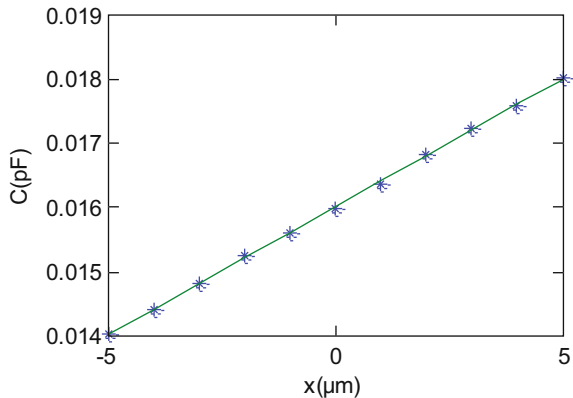
where  $V(t)$  is the driving voltage,  $C$  is the capacitance. In order to consider the fringing fields of the comb fingers, the capacitance-displacement relation is computed by a series of electrostatic computations of FEA. After least squares fitting, the capacitance-displacement relation is  $C = 0.000399x + 0.015998$  and shown in Fig. 10.

The total electrostatic force can be imposed on the mechanical structure by converting it to surface forces of the plates. Then the electrostatic force vector is

$$\mathbf{F}_e = \frac{N}{2}V(t)^2 \frac{\partial C}{\partial x} (\mathbf{B}_{dl} + \mathbf{B}_{dr} + \mathbf{B}_{su} + \mathbf{B}_{sd}) \tag{24}$$

where  $N$ ,  $\mathbf{B}_{dl}$ ,  $\mathbf{B}_{dr}$ ,  $\mathbf{B}_{su}$ ,  $\mathbf{B}_{sd}$  are the number of comb-drive fingers in one side, the distribution matrix of the electrostatic force on the left, the distribution matrix of the electrostatic force on the right, the distribution matrix of the electrostatic force in the above, the distribution matrix of the electrostatic force in the bottom, respectively.

**Fig. 10** The capacitance-displacement relation. The simulation results are shown as a line





### Temperature Effect

The effects of environmental thermal fluctuations on the gyroscope should be considered when estimating the actual performance of the microgyroscope. The effects of temperature changes on the frequency of a system can be decoupled into three different types: changes in the modulus of the elasticity, thermal expansion, and thermally induced stress (Painter and Shkel 2001, 2003).

The relation of Young's modulus and temperature is

$$E(T) = E(T_0) + E(T_0)TC_E\Delta T \quad (25)$$

where  $E(T)$  is Young's modulus at temperature  $T$ ,  $E(T_0)$  is the initial value of Young's modulus,  $TC_E$  is the temperature coefficient of Young's modulus.

The thermal load vector imposed on the structure is

$$\mathbf{F}_t = ETC_h\Delta T\mathbf{B}_t \quad (26)$$

where  $\mathbf{B}_t$  is the distribution matrix of the thermal load.

Residual stress usually happens in many MEMS fabrication processes and it will impact the performance of the micromachined devices. Since the thermal stress can be calculated by  $\sigma = ETC_h\Delta T$ , the effect of residual stress can be modeled by thermal load.

### Centrifugal Effect

Rotation of a structure will cause centrifugal load. The centrifugal load is

$$\mathbf{F}_r = \omega^2\rho\mathbf{B}_r \quad (27)$$

where  $\omega$  is angular velocity of rotation and  $\mathbf{B}_r$  is the distribution matrix of the centrifugal load.

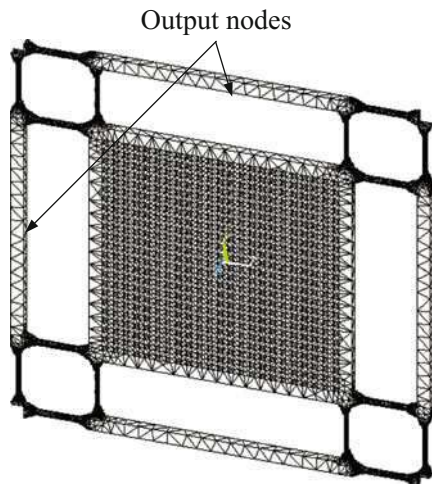
### Finite Element Modeling

In order to extract the matrices of the Eq. (12), the finite element model of the microgyroscope was constructed using ANSYS as shown in Fig. 11. As the combs do not significantly impact the eigenfrequency of the structure and will obviously increase the number of the nodes and cause the singular of the grids, here the model is simplified by removing the comb fingers. It consists of 45,092 tetrahedral structural elements (SOLID187) and 93,139 nodes. Dirichlet boundary conditions are applied to all degrees of freedom of the nodes belonging to the bottom surfaces of the suspension beams and 1656 degrees of freedom are constrained. Two nodes which are located in the middle of the suspension beams of driving and sense direction are selected as output nodes for observation of the dynamic response.

### Parametric Order Reduction

In order to preserve the parameters in Eq. (12), parametric order reduction method is used. Since the damping matrix  $\mathbf{D}$  in Eq. (12) does not have to take part in the process of generation of the subspace proved by Eid et al. (2007). The element mass matrix is (Guo and Chu 2001)

**Fig. 11** The finite element mesh and definition of output nodes. The number of elements (SOLID187) are 45,092



$$\mathbf{M}_e = \rho \int_V \mathbf{N}^T \mathbf{N} dv \quad (28)$$

where  $\mathbf{N}$  is the element shape function matrix.

And the element Coriolis matrix is

$$\mathbf{G}_e = \rho \int_V \mathbf{N}^T \boldsymbol{\Omega} \mathbf{N} dv \quad (29)$$

where  $\boldsymbol{\Omega}$  is the rotational matrix. Here the angle velocity around z-axis is only considered, so the  $\boldsymbol{\Omega}$  can be written as

$$\boldsymbol{\Omega} = \omega_z \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (30)$$

So the global matrices  $\mathbf{M}$  and  $\mathbf{G}$  can be an assembly of element matrices  $\mathbf{M}_e$  and  $\mathbf{G}_e$ . The Coriolis matrix can be obtained from the mass matrix. Therefore, here the Coriolis matrix does also take part in the process of generation of the subspace.

Then the parameterized equation of the micogyroscope mentioned above can be rewritten as

$$c_1 \mathbf{M} \ddot{\mathbf{u}}(t) + (\mathbf{K} + c_2 \mathbf{S}) \mathbf{u}(t) = \frac{1}{E} (\mathbf{F}_e + \mathbf{F}_t + \mathbf{F}_r) \quad (31)$$

where  $c_1 = \rho/E$ ,  $c_2 = TC_h \Delta T$ .

So a total of 2 parameters should be accounted during the reduction process. The stiffness matrix  $\mathbf{K}$  is selected as a constant matrix to generate the projection matrix.

The Laplace transform of Eq. (31) can be written as

$$(c_1\mathbf{M}s^2 + (\mathbf{K} + c_2\mathbf{S}))\mathbf{X}(s) = \frac{1}{E}\mathbf{B}\mathbf{U}(s) \quad (32)$$

Then the transfer function is

$$\begin{aligned} H(s) &= \frac{1}{E}\mathbf{L}^T[\mathbf{K} + \tilde{s}_1\mathbf{P}_1 + \tilde{s}_2\mathbf{P}_2]\mathbf{B} \\ &= \frac{1}{E}\mathbf{L}^T[\mathbf{K} + \tilde{s}_1\mathbf{P}_1 + \tilde{s}_2\mathbf{P}_2]^{-1}\mathbf{B} \\ &= \frac{1}{E}\mathbf{L}^T[\mathbf{I} - (-\tilde{s}_1\mathbf{K}^{-1}\mathbf{P}_1 - \tilde{s}_2\mathbf{K}^{-1}\mathbf{P}_2)]^{-1}\mathbf{K}^{-1}\mathbf{B} \\ &= \frac{1}{E}\mathbf{L}^T\sum_{n=1}^{\infty}(-\tilde{s}_1\mathbf{K}^{-1}\mathbf{P}_1 - \tilde{s}_2\mathbf{K}^{-1}\mathbf{P}_2)^n\mathbf{K}^{-1}\mathbf{B} \\ &\approx \frac{1}{E}\mathbf{L}^T\left[\sum_{n=1}^{\infty}(-\mathbf{K}^{-1}\mathbf{P}_1)^n\mathbf{K}^{-1}\mathbf{B}\tilde{s}_1^n + \sum_{n=1}^{\infty}(-\mathbf{K}^{-1}\mathbf{P}_2)^n\mathbf{K}^{-1}\mathbf{B}\tilde{s}_2^n\right] \end{aligned} \quad (33)$$

where  $\tilde{s}_1 = c_1s^2$ ,  $\tilde{s}_2 = c_2$ ,  $\mathbf{P}_1 = \mathbf{M}$ ,  $\mathbf{P}_2 = \mathbf{S}$ .

The pure moments are only computed and the two projection matrices can be generated separately by

$$\begin{aligned} \text{colspan}(\mathbf{V}_1) &= K_{r1}(-\mathbf{K}^{-1}\mathbf{P}_1, \mathbf{K}^{-1}\mathbf{B}) \\ \text{colspan}(\mathbf{V}_2) &= K_{r2}(-\mathbf{K}^{-1}\mathbf{P}_2, \mathbf{K}^{-1}\mathbf{B}) \end{aligned} \quad (34)$$

The final projection matrix  $\mathbf{V}$  is composed of the two projection matrices above as

$$\text{colspan}(\mathbf{V}) = \text{colspan}(\mathbf{V}_1, \mathbf{V}_2) \quad (35)$$

Then the reduced order system can be obtained by projecting Eq. (12) on the subspace  $\mathbf{V}$

$$\begin{aligned} \rho\mathbf{M}_r\ddot{\mathbf{x}}_r(t) + (\mathbf{D}_r + \omega\rho\mathbf{G}_r)\dot{\mathbf{x}}_r(t) + \\ (\mathbf{E}\mathbf{K}_r + ETC_h\Delta T\mathbf{S}_r)\mathbf{x}_r(t) &= \mathbf{F}_{er} + \mathbf{F}_{tr} + \mathbf{F}_{rr} \\ \mathbf{y}(t) &= \mathbf{L}_r^T\mathbf{x}_r(t) \end{aligned} \quad (36)$$

where  $\mathbf{M}_r = \mathbf{V}^T\mathbf{M}\mathbf{V}$ ,  $\mathbf{D}_r = \mathbf{V}^T\mathbf{D}\mathbf{V}$ ,  $\mathbf{G}_r = \mathbf{V}^T\mathbf{G}\mathbf{V}$ ,  $\mathbf{K}_r = \mathbf{V}^T\mathbf{K}\mathbf{V}$ ,  $\mathbf{S}_r = \mathbf{V}^T\mathbf{S}\mathbf{V}$ ,  $\mathbf{F}_{er} = \mathbf{V}^T\mathbf{F}_e$ ,  $\mathbf{F}_{tr} = \mathbf{V}^T\mathbf{F}_t$ ,  $\mathbf{F}_{rr} = \mathbf{V}^T\mathbf{F}_r$ ,  $\mathbf{L}_r = \mathbf{V}^T\mathbf{L}$ .

So some important parameters can be preserved, like material density, Rayleigh damping coefficient, angle velocity, Young's modulus, the coefficient of thermal expansion, etc., due to the projection matrix  $\mathbf{V}$  does not depend on these parameters.

### Simulation and Experiment

In this simulation, the parametric order reduction algorithm presented above was applied. The first subspace matrix was composed of 35 vectors and the second

subspace matrix was composed of 40 vectors. The final dimension of the reduced order model was 70.

In order to verify the designed interface circuit of the gyroscope, a numerical simulation was carried out. The effectiveness of the circuit was further tested on a fabricated MEMS vibratory gyroscope in conjunction with signal conditioning electronics.

### Simulation Results of the Mechanical

The slide film damping coefficient in the driving mode and the sense is the sum of the damping due to the center mass and the damping due to the comb fingers. The slide film damping coefficient for operating pressures of one atmosphere is  $1.54 \times 10^{-5}$  Ns/m, so the Rayleigh damping coefficient can be computed by Eq. (22). The large oscillating amplitude for driving mode and high sensitivity for sensing mode due to the slide-film damping effect, enable it a good performance at atmospheric pressure. A total of 190 frequency points are evaluated between 100 and 100 kHz. The harmonic response of the output node in the driving direction at atmospheric pressure is shown in Fig. 12a. It shows that the original model and the reduced order model match perfectly even at high frequency. The relative error is kept below 0.0007% and shown in Fig. 12b. The error for harmonic response was estimated as

$$\text{Error} = \frac{|\mathbf{H} - \hat{\mathbf{H}}|}{\mathbf{H}} \times 100\% \quad (37)$$

where  $\mathbf{H}$  is the harmonic solution of the full model (12) and  $\hat{\mathbf{H}}$  is the harmonic solution of the reduced order model (36).

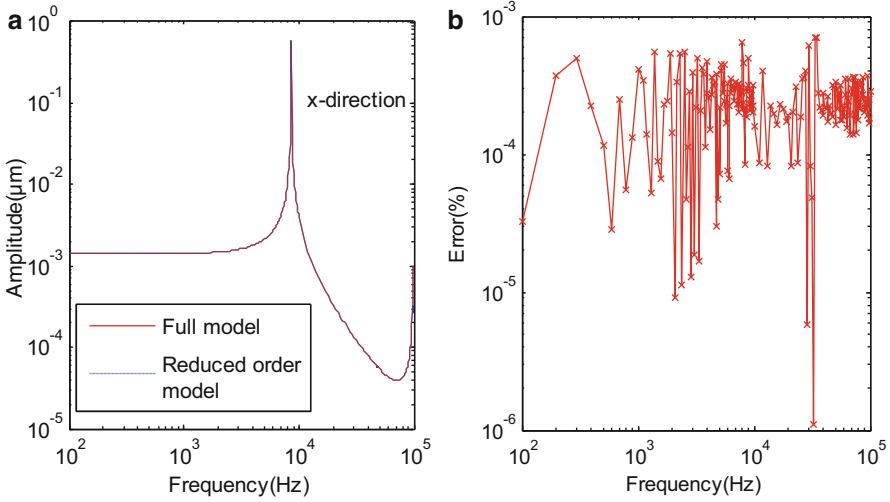
Figure 13a shows the harmonic response at different angular velocities in the sense direction and the relative error is shown in Fig. 13b. It can be seen that the maximum error is 0.2% when the angular velocity is 100 rad/s.

In order to research the effect of the temperature change, the harmonic response at 100 °C and -100 °C are shown in Figs. 14a and 15a, respectively. It shows that the resonate frequency of the gyroscope decreases as the temperature increases. Due to the symmetrical structure, the modal frequencies also match well. It can be seen from Figs. 14b and 15b that the relative error keeps below 0.42% when the temperature changes from -100 °C to 100 °C.

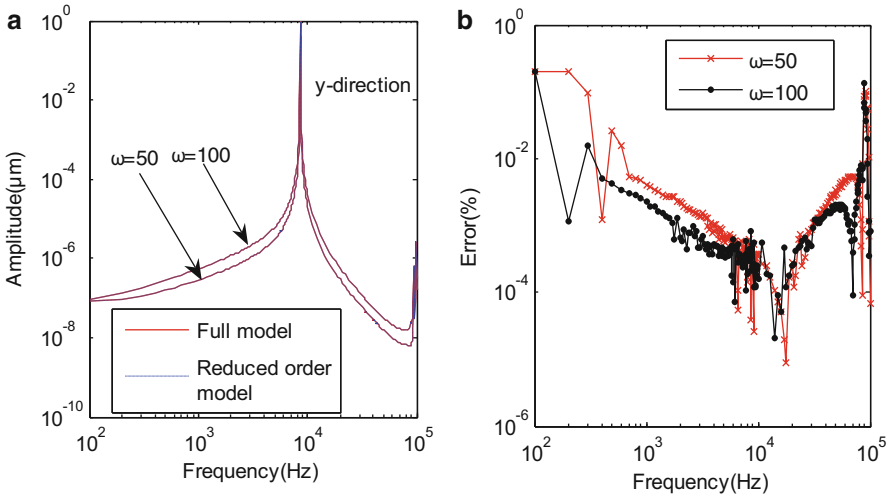
Table 4 shows the computation time between the original model and the reduced order model. It can be seen that the simulation speeds up 27 times, but the error is below 0.42%.

### Simulation Results of Mechanical with Interface Circuit

The circuit block diagram of the gyroscope is shown in Fig. 16. The circuit consists of a self-oscillation loop and a Coriolis signal channel. The gyroscope resonates in the direction of the driving axis using a self-oscillation technique which is a useful method for driving a mechanical resonator at the resonant frequency. The mechanical displacement causes a change in the capacitance of the electrodes, and the



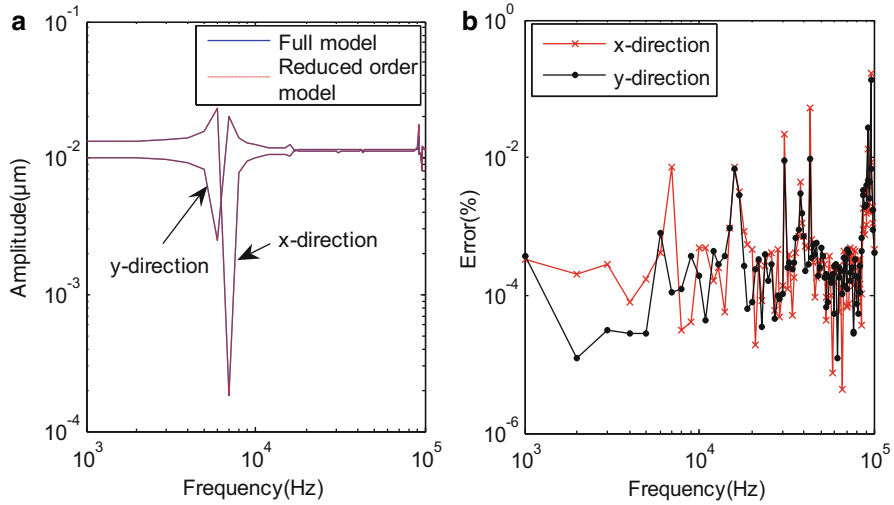
**Fig. 12** (a) Harmonic response in the driving direction at atmospheric pressure. (b) Error for harmonic response



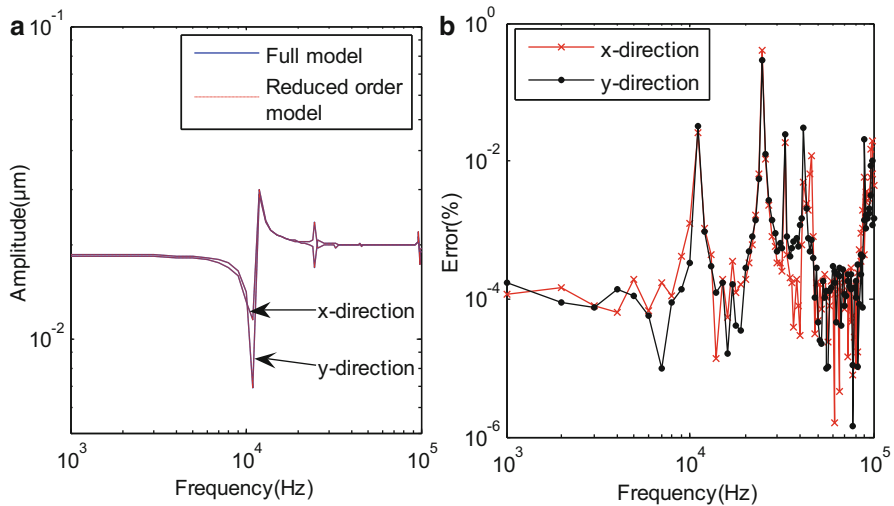
**Fig. 13** (a) Harmonic response in the sense direction at atmospheric pressure. (b) Error for harmonic response

change in capacitance is converted to an electrical signal by a charge amplifier. The Coriolis signal is filtered, demodulated, and amplified to pick-off the angular rate signal through the Coriolis signal channel.

The mechanical model and its interface circuit were described using hardware description language in Saber. The carrier wave used for modulation is 100 kHz.



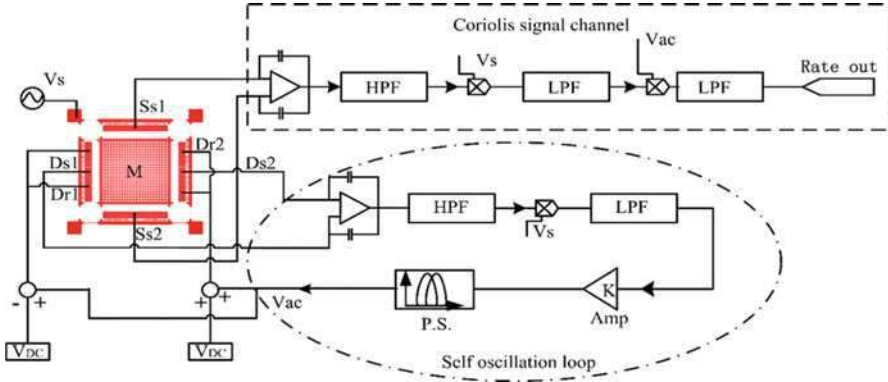
**Fig. 14** (a) Harmonic response at the output nodes at 100 °C. (b) Error for harmonic response



**Fig. 15** (a) Harmonic response at the output node at -100 °C. (b) Error for harmonic response

**Table 4** Comparison of simulation time between the original model and the reduced order model

Model	Size	Harmonic simulation time	Error
Original model	334,713	33,377 s	
Reduced order model	70	1220s	0.42%



**Fig. 16** The circuit block diagram of the gyroscope. The circuit consists of a self-oscillation loop and a Coriolis signal channel

The applied force on the sensing system is the sum of the Coriolis force and the electrostatic force. The displacement in the sensing direction makes a capacitance change in the sensing combs. The change in capacitance is converted into current and voltage by a charge amplifier. The current and voltage are calculated using:

$$I = V_s \left( \frac{\partial C}{\partial y} \right) \frac{dy}{dt}, V = - \frac{V_s}{C_f} \left( \frac{\partial C}{\partial y} \right) y \tag{38}$$

where  $V_s$  is the amplitude of the carrier wave,  $C$  is the capacitance change of the sensing electrode,  $y$  is the displacement in the sensing direction and  $C_f$  is the feedback capacitor of the charge amplifier.

In this simulation, on the one hand, to reduce the effect of the high-frequency transients and decrease the simulation time, minimum time step was set. Accounting for the carrier signal, the time step was 1us and the minimum step was 500 ns. On the other hand, to fast verify the designed interface circuit of the gyroscope, the effect of temperature change was not temporarily considered and the order of the reduced model was 35.

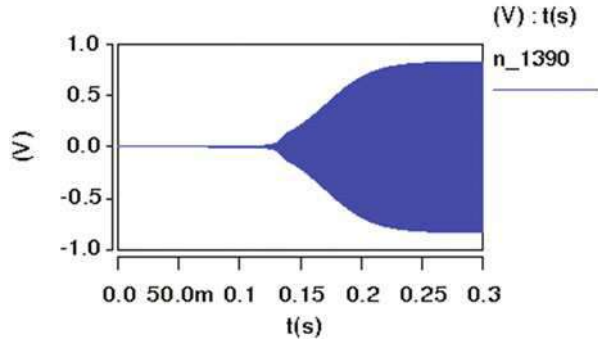
When the simulation time was 120 ms, the total computation time was 6960 s on a PC with 3 GB RAM and 2.53 GHz CPU.

If the reference voltage of the automatic gain control and time constant the low pass filter do not change, the response time of the driving loop is 250 ms when the driving gain is 1500 and shown in Fig. 17. Figure 18 shows that the response time is shorten to 80 ms when the driving gain becomes 1800. It can be concluded that the response time can be reduced by increasing the driving loop gain.

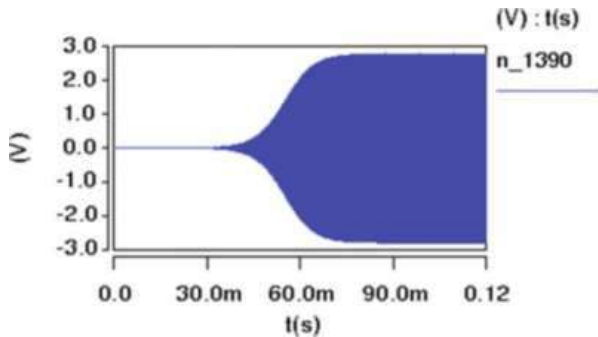
**Experiment**

In this section, the experiment results of the gyroscope using the proposed circuit are presented. Figures 19 and 20 show that the test results of the closed loop driving

**Fig. 17** The simulation result of the response time when the driving gain is 1500



**Fig. 18** The simulation result of the response time when the driving gain is 1800



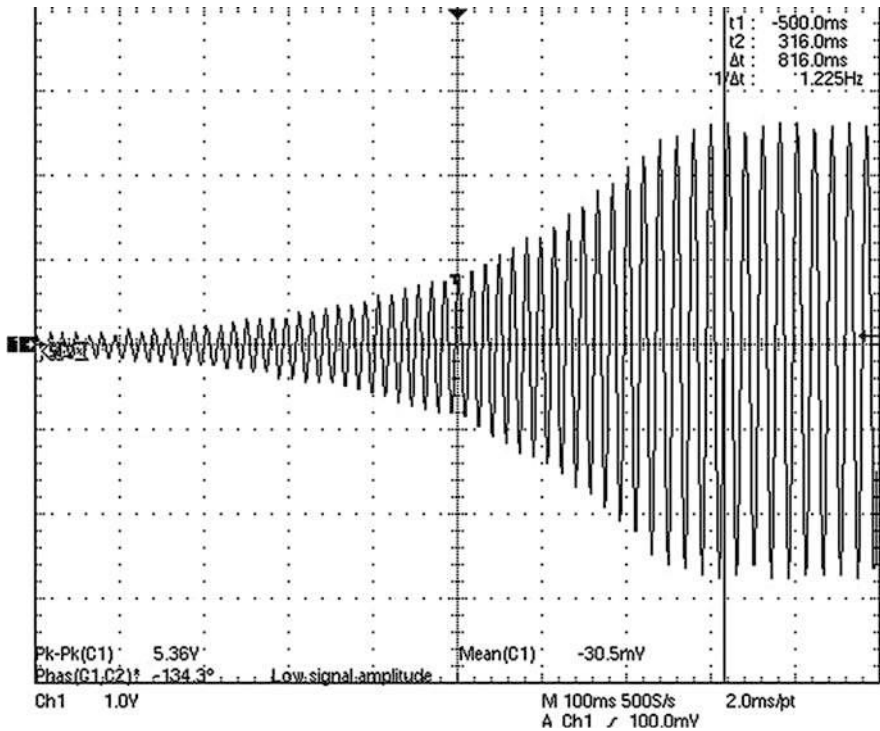
circuit. From the results it was observed that the response time is 816 ms when the driving gain is 1500 and response time is shorten to 554 ms when the driving gain is 1800. This agrees with the simulation results.

To calculate the scale factor value, we perform a dynamic rate experiment using a high precision rate table-FDT-01, a function Generator-AFG320, an oscilloscope-TDS 3014B and a data acquisition-PCI-6024E.

Figure 21 shows the gyroscope output signal together with the temperature and atmospheric pressure. The x-axis in the figure represents the reference signal and the y-axis represents the output voltage from the gyroscope system. The asterisk mark is the measured data and the dotted line is the first-order regressive model. The value of the scale factor is about  $17.8 \text{ mVdeg}^{-1} \text{ s}^{-1}$ . Compared with the simulation result, the relative error is within 1%.

In previous sections, we have demonstrated that the proposed method for a microgyroscope can be able to reduce the dimension of the full model considerably and at the same time to preserve some important parameters and the accuracy within engineering requirements. In this section, we will discuss a few important issues about system level modeling of the mirogyroscope using PMOR.

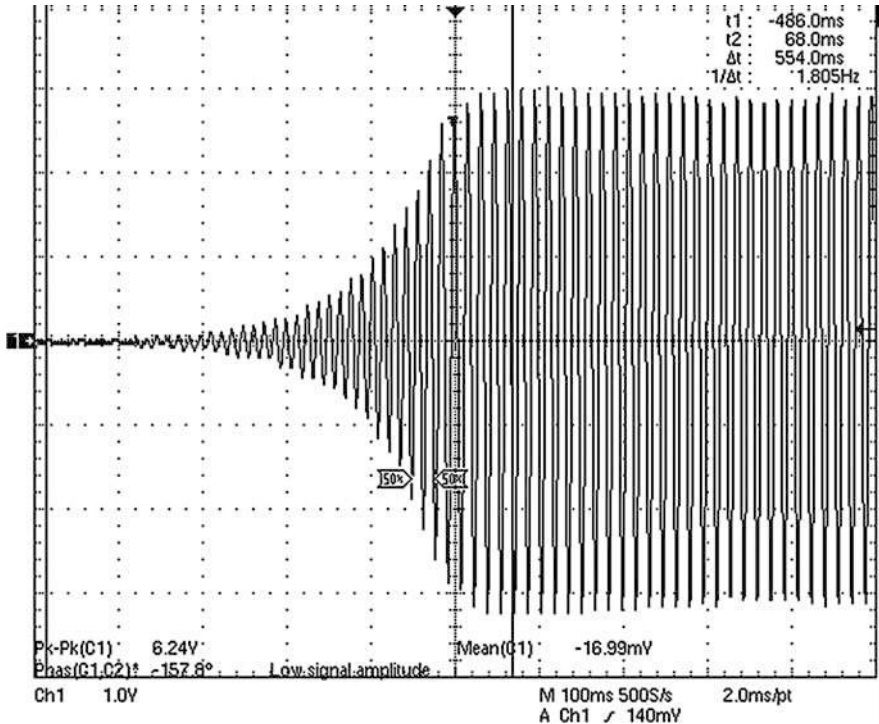




**Fig. 19** The measured result of the response time when the driving gain is 1500

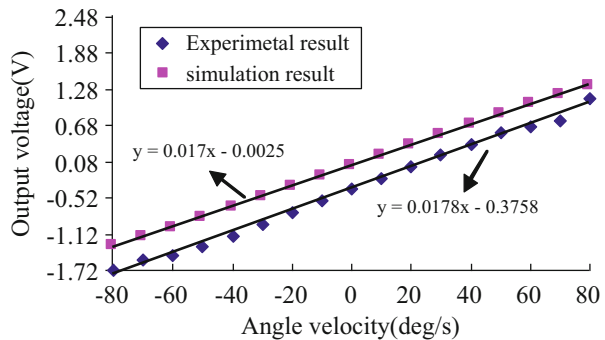
Different from the traditional model order reduction method, PMOR need Taylor expansion in different parameters direction. During the process, there will be a large number of mixed moments. The number of mixed moments grows linearly with each iteration. If many parameters need to be preserved, it usually leads to a high dimensional system and a rapid accumulation of the rounding errors. So it can neglect the mixed moments and perform separate Taylor expansions in all parameter directions. Then the number of moments can be selected according to their importance. However, the order of the reduced model is very high if considering many parameters. So under preserving the accuracy the order of the macromodel is the lower the better. The Rayleigh damping matrix does not have to be taken into account for generation of the subspace which is proven by Eid et al. (2007). Here we find that the Coriolis matrix does not also have to take part in the reduction. And it will reduce the dimension of the macromodel dramatically due to avoid expanding all the parameters direction.

However, when cosimulating with the interface circuit, the frequency of carrier wave in the interface circuit is usually larger than 100 kHz and the generated macromodel must consider the high frequency character. In order to ensure the



**Fig. 20** The measured result of the response time when the driving gain is 1800

**Fig. 21** The gyroscope output signal versus input angular rate



accuracy of macromodel at high frequency, more moments need to match along frequency which will lead to a larger dimensional system. On the other hand, the typical resonant frequency of the MEMS inertial sensors is about 1 to 20 kHz. The different time scale of the mechanical and the circuit lead that cosimulation will be very time-consuming and hard to be convergent. To speed the simulation, it demands that the order of the macromodel as low as possible. So it must make tradeoffs between accuracy and cost.

## Trajectory Piecewise-Linear (TPWL) Method

Compared to linear MOR methods, nonlinear MOR methods are still in the elementary stage and need to be developed further (Striebel and Rommes 2011). Among these, the two most representative methods are Proper Orthogonal Decomposition (POD) (Astrid 2004) and Trajectory Piecewise-Linear (TPWL) (Rewiński and White 2003; Bond 2010; He et al. 2011; Nahvi et al. 2012). In POD, the basis functions are derived from the collected snapshots, and then the original nonlinear system is directly projected onto a low dimension subspace expanded by the basis functions. This nonlinear MOR method is very accurate, but its efficiency is greatly limited by the simulation of the original nonlinear system to extract enough snapshots. Moreover, because evaluating the reduced nonlinear term in each step is time-consuming (Rewiński and White 2003), the simulation cost of compact model is not significantly reduced. Alternatively, the TPWL method first approximates the nonlinear model piecewise-linearly by a group of local linearized models and then reducing these local linearized models by some linear MOR method (such as Krylov-subspace methods (Simoncini and Szyld 2007) and Truncated Balanced Realization (TBR) methods (Vasilyev et al. 2003)). Finally, the compact model is generated by weighting all the local linearized reduced models. It has been demonstrated that the TPWL method is capable of capturing strongly nonlinear behavior of the original system (Rewiński and White 2003). And as a result of a cost-efficient representation of system's nonlinearity, the simulation cost of the TPWL compact model is less than that of the POD compact model (Voss et al. 2008). This superiority is quite important in MOR techniques. In this section, a micromachined thermal sensor is taken as an example to show the modeling process of TPWL method.

## Thermoelectric Governing Differential Equations

Micromachined thermal sensors are based on the interactions of the electric, thermal, and fluid fields. However, it is well known that fluid simulation needs a high computational cost, so we exclude the fluid field from the computational domain and use a forced convection boundary condition for the thermal field instead. Generally, thermoelectric effects include the Seebeck effect, Peltier effect, and Thomson effect. In this chapter, the Seebeck and Peltier effects can be neglected and set to zero. Thus, refer to the system of coupled equations of thermoelectricity given in reference (Antonova and Looman 2005), the thermoelectric interaction of micromachined thermal sensors can be described by the following governing equations:

$$\begin{cases} -\nabla \cdot (\sigma(T)\nabla\varphi) = 0 \\ \rho C_h \frac{\partial T}{\partial t} - \nabla \cdot (k\nabla T) = \mathbf{J} \cdot \mathbf{E} \end{cases} \quad (39)$$

where  $\varphi$  is the applied voltage,  $T$  is the temperature,  $\sigma(T)$  is the electrical conductivity,  $\rho$  is the density,  $C_h$  is the specific heat and  $k$  is the thermal conductivity.  $\mathbf{E}$  and  $\mathbf{J}$  are the electric field intensity and the electric current density, respectively, and are calculated as follows:

$$\mathbf{E} = -\nabla\varphi, \mathbf{J} = \sigma\mathbf{E}$$

Denote the computation domain by  $\Omega$ , and then the boundary conditions of electric field can be represented as

$$\begin{cases} \varphi(x, y, z) = \varphi_0 & (x, y, z) \in \Gamma_D^\varphi \\ \mathbf{n} \cdot \mathbf{J}|_{(x, y, z)} = J_0 & (x, y, z) \in \Gamma_N^\varphi \end{cases} \quad (40)$$

where  $\Gamma_D^\varphi \cap \Gamma_N^\varphi = \emptyset$ ,  $\Gamma_D^\varphi \cup \Gamma_N^\varphi = \partial\Omega$  and  $\mathbf{n}$  is the unit outward normal to boundary surface  $\Gamma_N^\varphi$ . In particular,  $\varphi_0$  is equal to zero at the grounded part of  $\Gamma_D^\varphi$ , and  $J_0$  is equal to zero at the electrically insulated part of  $\Gamma_N^\varphi$ .

For thermal field, the boundary conditions are represented as:

$$\begin{cases} T(x, y, z) = T_w & (x, y, z) \in \Gamma_D^T \\ \mathbf{n} \cdot k\nabla T = h(T_\infty - T) & (x, y, z) \in \Gamma_N^T \end{cases} \quad (41)$$

where  $\Gamma_D^T \cap \Gamma_N^T = \emptyset$ ,  $\Gamma_D^T \cup \Gamma_N^T = \partial\Omega$ ,  $\mathbf{n}$  is the unit outward normal to boundary surface  $\Gamma_N^T$ ,  $T_\infty$  is the environmental temperature,  $T_w$  is the wall temperature,  $h$  is the convection intensity coefficient. When  $h$  is equal to zero, the corresponding part of  $\Gamma_N^T$  is thermally isolated. And otherwise the corresponding part of  $\Gamma_N^T$  is surrounded by fluid and forced by a convection condition.

The initial condition of thermal field is:

$$T(x, y, z, 0) = T_0(x, y, z) \quad (x, y, z) \in \Omega \quad (42)$$

## Trajectory Piecewise-Linear Method

The linearized model only has good approximation at the states which are very close to the linearization point, so the basic idea behind the TPWL method is to select multiple linearization points, and then to linearize the nonlinear model near each linearization point. A global reduced subspace can be created by these generated local linearized models. And each local linearized model can be reduced by projecting it into the global reduced subspace. A TPWL compact model is constructed as a weighted sum of all local linearized reduced models. In the following subsections, we will detailedly discuss the TPWL method from three aspects.

Finite element discretization of Eq. (39) yields:

$$\mathbf{M} \frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t)) + \mathbf{B}\mathbf{u}(t) \quad (43)$$

where  $\mathbf{x}(t) = (\varphi_1(t), \varphi_2(t), \dots, \varphi_n(t), T_1(t), T_2(t), \dots, T_n(t))^T$  is a state vector at time  $t$ ,  $\varphi_i(t)$  and  $T_i(t)$  are respectively the potential and temperature at the  $i$ th discrete point,  $n$  is the number of discrete points,  $\mathbf{u}(t)$  is an input signal vector,  $\mathbf{B}$  is an input matrix that describes  $\mathbf{u}(t)$ 's distribution,  $\mathbf{M}$  is a  $2n \times 2n$  matrix,  $\mathbf{f} : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$  is a nonlinear vector-valued function.

### Creating the Local Linearized Models

Given an input signal  $\mathbf{u}(t)$  and an initial state  $\mathbf{x}_0$ , we need to simulate the nonlinear system (43) to select linearization points from a ‘‘training’’ trajectory. The initial state  $\mathbf{x}_0$  is selected as the first linearization point and the corresponding linearized model is as follows:

$$\mathbf{M} \frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}_0) + \mathbf{A}_0(\mathbf{x} - \mathbf{x}_0) + \mathbf{B}\mathbf{u}(t) \quad (44)$$

where  $\mathbf{A}_0$  is the Jacobian of  $\mathbf{f}(\mathbf{x})$  evaluated at  $\mathbf{x}_0$ . For a given maximum allowable deviation  $\alpha$ , if the current state  $\mathbf{x}$  is outside the neighborhoods of radius  $\alpha$  of all existing  $i$  linearization points  $\mathbf{x}_i$ , i.e.,  $\min_{0 \leq j < i} \|\mathbf{x} - \mathbf{x}_j\| \geq \alpha$ , a new linearized model should be generated:

$$\mathbf{M} \frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}_i) + \mathbf{A}_i(\mathbf{x} - \mathbf{x}_i) + \mathbf{B}\mathbf{u}(t) \quad (45)$$

where  $\mathbf{x}_i = \mathbf{x}$  is the  $(i + 1)$ th linearization point,  $\mathbf{A}_i$  is the Jacobian of  $\mathbf{f}(\mathbf{x})$  evaluated at  $\mathbf{x}_i$ . Because the original nonlinear system needs to be simulated, the above process is a time-consuming task. In order to reduce computational cost, the newly generated local linearized model is reduced, and then the local linearized reduced model is simulated until the next linearization point is found. This process named fast TPWL method computes an approximate trajectory and obtains a set of approximate linearization points. Obviously, the computation efficiency is greatly improved at the expense of precision. In section ‘‘Analysis of the Simulation Results,’’ the performances of the TPWL method and the fast TPWL method on a micromachined thermal sensor will be compared.

Further, the maximum allowable deviation  $\alpha$  is estimated as follows:

$$\alpha = \|\mathbf{x}_\infty - \mathbf{x}_0\| / s$$

where  $s$  is the number of linearization points,  $\mathbf{x}_\infty$  is the steady state which is obtained through simulating the linearized reduced model related to the initial state  $\mathbf{x}_0$ .

### Generating the Global Projection Matrix

In a sense, any linear MOR method, such as Krylov-subspace methods (Simoncini and Szyld 2007) and Truncated Balanced Realization (TBR) methods (Vasilyev et al. 2003), can be applied to the linearized models to generate local projection matrices  $\mathbf{V}_i, i = 0, \dots, s - 1$ . In this chapter, a block Krylov-subspace method with deflation

procedure (Simoncini and Szyld 2007; Freund 2000) is employed. In this process, the column number of each local projection matrix  $\mathbf{V}_i$  is bounded by a predetermined reduced basis order  $r$  from above. The aggregate projection matrix  $\mathbf{V}_g$  is obtained through merging all local projection matrices  $\mathbf{V}_i$ ,  $i = 0, \dots, s - 1$ , that is,  $\mathbf{V}_g = [\mathbf{V}_0, \dots, \mathbf{V}_{s-1}]$ . Generally, the columns of  $\mathbf{V}_g$  are linearly dependent which is against an invertible linear projection matrix, so they must be orthogonalized before use. The singular value decomposition (SVD) algorithm is commonly used in this stage due to its excellent performance. Finally, the global projection matrix  $\mathbf{V}$  is formed by collecting the orthogonalized columns of  $\mathbf{V}_g$  corresponding to the largest singular values. Besides, the deflation Gram-Schmidt method (Freund 2000) can also be used to generate  $\mathbf{V}$ . The advantages of this method include easy to use, low memory cost and high efficiency. Therefore, we will use the deflation Gram-Schmidt method rather than the SVD algorithm to  $\mathbf{V}_g$ .

### Constructing the TPWL Compact Model

The TPWL compact model can be represented as a weighted combination of all local linearized reduced models:

$$\mathbf{M}_r \frac{d\mathbf{z}(t)}{dt} = \sum_{i=0}^{s-1} \tilde{w}_i(\mathbf{z}) \mathbf{A}_{ir} \mathbf{z} + \sum_{i=0}^{s-1} \tilde{w}_i(\mathbf{z}) \mathbf{f} \mathbf{A}_{ir} + \mathbf{B}_r \mathbf{u}(t) \quad (46)$$

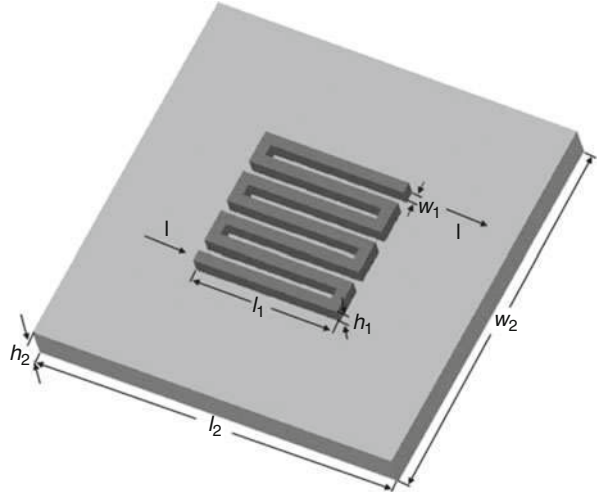
where  $\mathbf{z} = \mathbf{V}^T \mathbf{x}$ ,  $\mathbf{M}_r = \mathbf{V}^T \mathbf{M} \mathbf{V}$ ,  $\mathbf{A}_{ir} = \mathbf{V}^T \mathbf{A}_i \mathbf{V}$ ,  $\mathbf{f} \mathbf{A}_{ir} = \mathbf{V}^T (\mathbf{f}(\mathbf{x}_i) - \mathbf{A}_i \mathbf{x}_i) \mathbf{V}$ ,  $\mathbf{B}_r = \mathbf{V}^T \mathbf{B} \mathbf{V}$ . For any  $\mathbf{z}$ , the weight functions  $\tilde{w}_i(\mathbf{z}) = e^{-\beta d_i/m} / \sum_{i=0}^{s-1} e^{-\beta d_i/m}$ ,  $i = 0, \dots, s - 1$ , and  $\sum_{i=0}^{s-1} \tilde{w}_i(\mathbf{z}) = 1$ , where  $d_i = \|\mathbf{z} - \mathbf{z}_i\|_2$ ,  $m = \min_{0 \leq i \leq s-1} d_i$  and  $\beta$  is a positive constant to control the compactness of weight functions, and usually taken as  $\beta = 25$ . There are still other ways to construct the weight functions. To use them, the readers can refer to references (Tiwary and Rutenbar 2005; Bond and Daniel 2007).

## Compact Model Extraction and Result Analysis

### Geometric and Material Parameters

A micromachined thermal sensor shown in Fig. 22 is comprised of a Pt thermal sensitive resistor and a polyimide flexible substrate, as our research subject. The thermal sensitive resistor is loaded by a constant current  $I$ . Specifically, the thermal sensitive resistor is grounded at one end and charged at the other end. This results in the boundary conditions for the electric field that are  $\varphi = 0$  at the grounded end of the thermal sensitive resistor,  $\mathbf{n} \cdot \mathbf{J} = J_0$  at the charged end of the thermal sensitive resistor, and  $\mathbf{n} \cdot \mathbf{J} = 0$  at the rest of boundaries. Here  $J_0 = I/S$  and  $S$  is the related cross-sectional area. For the thermal field, forced convective heat transfer is applied on the upper surfaces of the micromachined thermal sensor, i.e.,  $\mathbf{n} \cdot k \nabla T = h(T_\infty - T)$ . Because the thermal conductivity of polyimide is very small, the remaining boundaries are set as thermal insulation, i.e.,  $\mathbf{n} \cdot k \nabla T = 0$ . Tables 5, 6, and 7 show the relevant

**Fig. 22** Schematic of the micromachined thermal sensor made of a platinum structure on polyimide



**Table 5** The geometric dimensions of the micromachined thermal sensor

Geometric dimensions	$l_1$	$w_1$	$h_1$	$l_2$	$w_2$	$h_2$
Length( $\mu\text{m}$ )	400	30	0.5	1000	1000	25

**Table 6** The material properties of the micromachined thermal sensor

Material properties	Pt	Polyimide
Density ( $\text{kg m}^{-3}$ )	21,450	1420
Thermal conductivity ( $\text{W m}^{-1} \text{K}^{-1}$ )	73	0.1
Specific heat ( $\text{J kg}^{-1} \text{K}^{-1}$ )	132.51	1090
Resistivity( $\Omega \text{ m}$ )	$1.06 \times 10^{-7}$ (20 °C)	$1 \times 10^{16}$
	$4.3 \times 10^{-7}$ (1000 °C)	

**Table 7** The boundary condition parameters of the micromachined thermal sensor

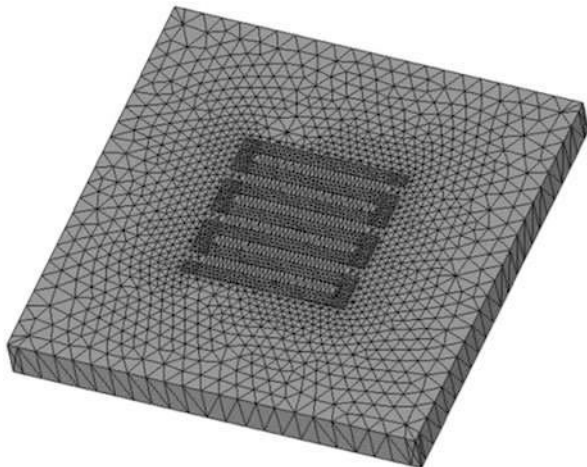
Parameter	Current	Environmental temperature	Convection intensity coefficient
Value	8 (mA)	20 (°C)	78.62 ( $\text{W m}^{-2} \text{K}^{-1}$ )

parameters of geometric dimensions, material properties, and boundary conditions of the micromachined thermal sensor, respectively.

### Compact Model Extraction

For spatial discretization of the thermoelectric coupled Eq. (39) by finite element method, the micromachined thermal sensor is discretized by 4954 nodes and 15,792 tetrahedral elements, as shown in Fig. 23. After spatial discretization, a system of nonlinear ordinary differential equations can be obtained with the form of Eq. (43). If we linearize the nonlinear term around  $\mathbf{x}(t^i)$  and regard the current  $I$  and the

**Fig. 23** Discretization of the micromachined thermal sensor



convection intensity coefficient  $h$  as input signals, i.e.,  $u_1(t) = I$  and  $u_2(t) = h$ , then we obtain a system of linearized ordinary differential equations:

$$\begin{pmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{M}^{TT} \end{pmatrix} \begin{pmatrix} \dot{\varphi} \\ \dot{T} \end{pmatrix} + \begin{pmatrix} \mathbf{A}^{\varphi\varphi} & \mathbf{A}^{\varphi T} \\ \mathbf{A}^{T\varphi} & \mathbf{A}^{TT} \end{pmatrix} \begin{pmatrix} \varphi \\ T \end{pmatrix} = \begin{pmatrix} \mathbf{F}^{\varphi} \\ \mathbf{F}^T \end{pmatrix} + \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{O} \end{pmatrix} u_1(t) + \begin{pmatrix} \mathbf{O} \\ \mathbf{B}_2 \end{pmatrix} u_2(t) \quad (47)$$

where

$$\begin{aligned} M_{ij}^{TT} &= \int_{\Omega} \rho C_h N_i N_j d\Omega, \\ A_{ij}^{\varphi\varphi} &= \int_{\Omega} \sigma(T^n) (\nabla N_i \cdot \nabla N_j) d\Omega, \\ A_{ij}^{\varphi T} &= \int_{\Omega} \sigma'(T^n) N_j (\nabla N_i \cdot \nabla \varphi^n) d\Omega, \\ A_{ij}^{T\varphi} &= - \int_{\Omega} N_i \sigma(T^n) (2 \nabla \varphi^n \cdot \nabla N_j) d\Omega, \\ A_{ij}^{TT} &= \int_{\Omega} k \nabla N_i \cdot \nabla N_j d\Omega - \int_{\Omega} N_i \sigma'(T^n) N_j (\nabla \varphi^n \cdot \nabla \varphi^n) d\Omega, \\ F_i^{\varphi} &= - \int_{\Omega} \sigma'(T^n) (\nabla N_i \cdot \nabla \varphi^n) T^n d\Omega, \\ F_i^T &= - \int_{\Omega} N_i (\sigma(T^n) + \sigma'(T^n) T^n) (\nabla \varphi^n)^2 d\Omega, \\ B_{1,i} &= 1 / s \int_{\Gamma_N^{\varphi}} N_i d\Gamma, \\ B_{2,i} &= \int_{\Gamma_N^T} N_i (T_0 - T^n) d\Gamma, \end{aligned}$$

where  $N_i(x, y, z)$ , ( $i = 1, 2, \dots, n$ ) are shape functions.  $T^n$  and  $\nabla \varphi^n$  are the temperature and electric potential gradient at time  $t_n$ , respectively, and can be calculated as follows:

$$T^n(x, y, z) = \sum_{k=1}^n N_k(x, y, z) T_k^n$$



$$\nabla\varphi^n(x, y, z) = \sum_{k=1}^n \nabla N_k(x, y, z)\varphi_k^n$$

where  $T_k^n$  and  $\varphi_k^n$  are the temperature and electric potential of the  $k$ th node at time  $t_n$ . For the TPWL method, we need to calculate Eq. (47) at each linearization point  $\mathbf{x}_m$ , ( $m = 0, 1, \dots, s - 1$ ), and then obtain the compact model (46) by projection and weighting.

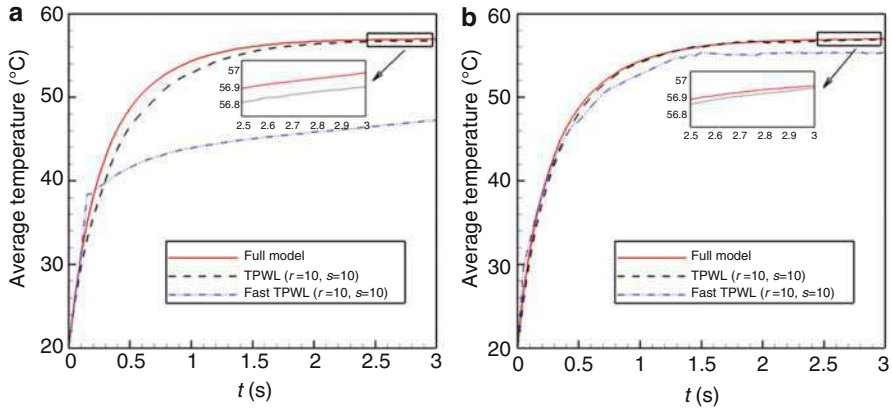
### Analysis of the Simulation Results

In this subsection, we first compare the efficiency and accuracy of the TPWL method and the fast TPWL method to find out which method is more suitable for the current micromachined thermal sensor. Next, the effects of the local reduced basis order  $r$  and the linearization point number  $s$  on the TPWL compact models are studied. Finally, the expandability of the TPWL compact models is also demonstrated.

In our simulations, the time step size is fixed as  $\Delta t = 0.05$ s and the total simulation time is  $t = 3$ s. All the calculations are made on a Lenovo notebook computer (CPU: Intel i3-M350, 2.27 GHz). Because the geometric dimensions and thermoelectric parameters of the thermal sensitive resistor and the flexible substrate have considerable differences in magnitude, the final system matrix is ill-conditioned. However, the procedures of simulating the full model and generating the projection matrix both need to solve the system of linear algebraic equations. Common sparse iteration linear system solvers, such as generalized minimum residual (GMRES), conjugate gradient (CG), biconjugate gradient (BICG) and biconjugate gradient stabilized (BICGSTAB) (Saad 2003), are unable to obtain convergent results. Therefore, the sparse LU decomposition method (Davis and Duff 1997) which has very good adaptability to ill-conditioned matrix is used in this chapter.

### Comparison of Two MOR Methods

Using the TPWL method and the fast TPWL method, we first generate compact models with the local reduced basis order  $r = 10$  and the linearization point number  $s = 10$  and  $s = 20$ . Figure 24 shows average temperatures of the thermal sensitive resistor (i.e., the heater structure on the upper surface of the micromachined thermal sensor), computed by the TPWL method, the fast TPWL method and the full model. Table 8 shows comparisons of extraction time, model size (i.e., degree of freedom of the model), and run time. Evidently, the results of the TPWL compact models show better agreement with the full model by comparison with the compact models generated by the fast TPWL method, so the TPWL method has higher precision than the fast TPWL method. Although when the linearization point number is increased, the simulation results of the fast TPWL method have some improvement. But at the same time, the extraction time, model size, and run time are also increased. As a result, we choose the TPWL method as MOR method for the present micromachined thermal sensor.



**Fig. 24** Comparisons of accuracy: **(a)** local reduced basis order  $r = 10$ , linearization point number  $s = 10$ ; **(b)** local reduced basis order  $r = 10$ , linearization point number  $s = 20$

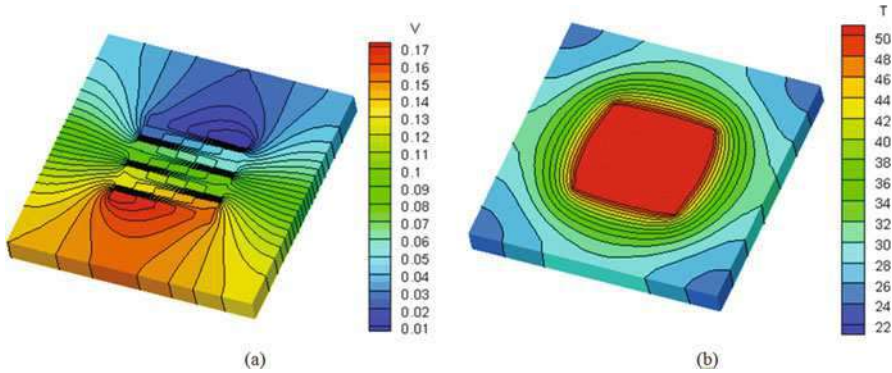
**Table 8** Comparisons of extraction time, model size (i.e., degree of freedom of the model) and run time of the TPWL and fast TPWL compact models

	MOR method	Extraction time	Model size	Run time
Full model	—	—	9908	5400 s
$r = 10$	TPWL	6342.3 s	40	2.075 s
$s = 10$	Fast TPWL	945.6 s	33	2.014 s
$r = 10$	TPWL	7253.8 s	54	2.482 s
$s = 20$	Fast TPWL	1851.2 s	41	2.341 s

It should be noted that the fast TPWL method still has a lot of merits. In fact, the fast TPWL method is effective for many applications. Because it does not need to simulate the full nonlinear model, its efficiency is far higher than other nonlinear model order reduction methods. But unfortunately, much accuracy will be lost in the situation of strong nonlinearity, unless much more linearization points are involved. Since we are more concerned with accuracy, the fast TPWL method is not adopted here.

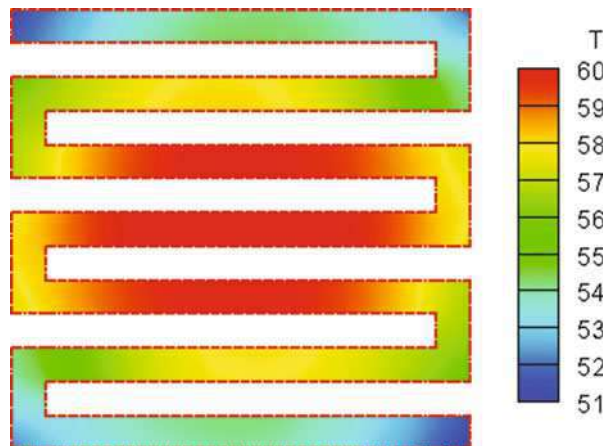
Figure 25 shows the electric potential and the temperature of the micromachined thermal sensor at  $t = 2$  s, computed by the compact model with the TPWL method ( $r = 10, s = 10$ ). It is seen that the computational results are smooth and continuous. In Fig. 26, only the temperature of the resistor is shown, and the temperature variations are very obvious. Hence, the interaction between the electrical field and thermal field cannot be ignored for the micromachined thermal sensor.

It is worth mentioning that we need to generate various TPWL compact models of the same nonlinear system for the sake of research. The generation of each TPWL compact model needs the simulation result of the nonlinear full model. As can be seen from Table 8, simulating the nonlinear full model is very time consuming. However, it is unnecessary to simulate the nonlinear full model every time.



**Fig. 25** Contour lines of electric potential and temperature computed by TPWL compact model: (a) Contour lines of electric potential; (b) contour lines of temperature

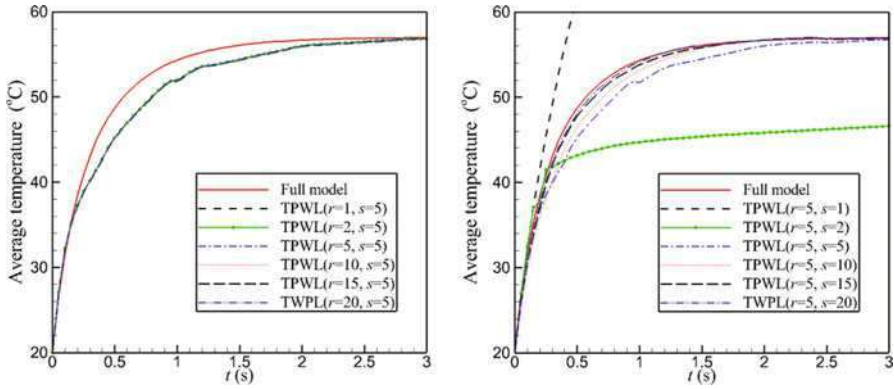
**Fig. 26** Calculated temperature distribution of the thermal sensitive resistor by the TPWL compact model



Throughout the study, the nonlinear full model only needs to be simulated once. If the simulation result is stored, then it can be invoked repeatedly to generate various TPWL compact models. According to this way, the efficiency of the TPWL method will be raised greatly.

### **Influence of Local Reduced Basis Order and Linearization Point Number**

Figure 27 presents average temperatures of the thermal sensitive resistor computed by various models. When the linearization point number is fixed and the local reduced basis order is increased, the performances of the TPWL compact models are not improved. On the contrary, as we gradually increase the linearization point number, the computational results of the TPWL compact models get closer to that of the nonlinear full model. This also indicates that the nonlinearity of the original full model is visible for the micromachined thermal sensor, and the piecewise linearization technique of the TPWL method has the advantages for extracting the compact



**Fig. 27** Influence of the local reduced basis order  $r$  and the linearization point number  $s$ : (a) linearization point number  $s = 5$  with varying local reduced basis order  $r$ ; (b) local reduced basis order  $r = 5$  with varying linearization point number  $s$

**Table 9** Relative errors measured in  $L2$  norm between the nonlinear full model and the TPWL compact models

	$r = 1$	$r = 2$	$r = 5$	$r = 10$	$r = 15$	$r = 20$
$s = 1$	6.6364e-1	1.3005	1.3005	1.3005	1.3005	1.3005
$s = 2$	1.1960e-1	1.1744e-1	1.1667e-1	1.1679e-1	1.1678e-1	1.1678e-1
$s = 5$	1.5727e-2	1.5223e-2	1.5257e-2	1.5261e-2	1.5261e-2	1.5261e-2
$s = 10$	6.4792e-3	6.4843e-3	6.4852e-3	6.4848e-3	6.4833e-3	6.4845e-3
$s = 15$	1.6950e-3	1.6976e-3	1.6985e-3	1.6991e-3	1.6989e-3	1.6990e-3
$s = 20$	8.7771e-4	8.7714e-4	8.7693e-4	8.7693e-4	8.7700e-4	8.7693e-4

**Table 10** The model size (i.e., degree of freedom of the model) of the TPWL compact models

	$r = 1$	$r = 2$	$r = 5$	$r = 10$	$r = 15$	$r = 20$
$s = 1$	2	3	6	11	16	21
$s = 2$	4	6	12	21	30	39
$s = 5$	8	11	21	30	44	55
$s = 10$	11	17	28	40	55	69
$s = 15$	15	22	34	47	65	78
$s = 20$	15	28	39	54	73	88

model of the micromachined thermal sensor. For quantitative studies, Table 9 lists the relative errors measured in  $L2$  norm between the nonlinear full model and the above TPWL compact models.

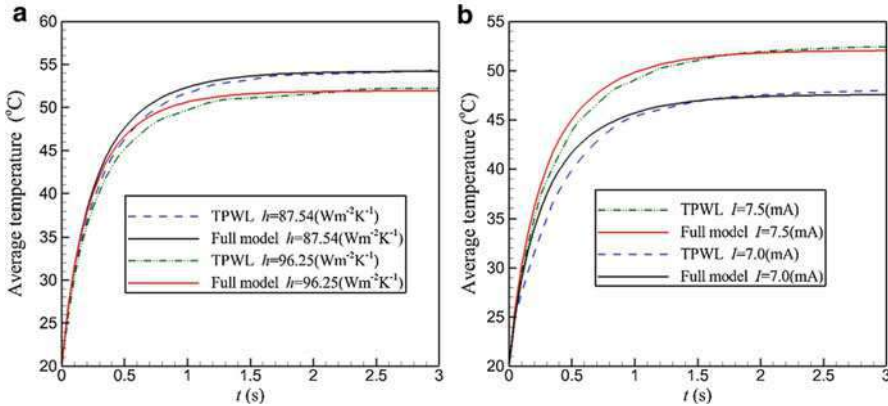
Besides the accuracy of the TPWL compact models, the local reduced basis order and the linearization point number affect the size of the TPWL compact models and further have effect on the run time of the TPWL compact models. The size of the TPWL compact models for different  $s$  and  $r$  are listed in Table 10. It is seen that the

**Table 11** The run time of the TPWL compact models (unit: s)

	$r = 1$	$r = 2$	$r = 5$	$r = 10$	$r = 15$	$r = 20$
$s = 1$	1.856	1.622	1.716	1.747	1.872	1.872
$s = 2$	1.700	1.700	1.981	1.700	2.013	1.966
$s = 5$	1.638	1.872	1.997	1.763	1.903	2.450
$s = 10$	1.919	1.653	1.919	2.075	2.122	2.278
$s = 15$	1.637	1.747	1.872	2.606	2.669	2.153
$s = 20$	2.122	2.028	2.028	2.482	2.560	2.278

**Table 12** Convection intensity coefficients of three different wind speeds

Wind speed (m/s)	17.5	20	22.5
Convection intensity coefficient ( $W/m^2K$ )	78.62	87.54	96.25



**Fig. 28** Expandability of the TPWL compact model: (a) Results of convection intensity coefficients  $h = 87.54$  and  $h = 96.25$  ( $W \cdot m^{-2} \cdot K^{-1}$ ) simulated by the TPWL compact model generated for the input convection intensity coefficient  $h = 78.62$  ( $W \cdot m^{-2} \cdot K^{-1}$ ); (b) Results of currents  $I = 7.0$  and  $I = 7.5$  (mA) simulated by the TPWL compact model generated for the input current  $I = 8.0$  (mA)

size of the TPWL compact models increases as the increase of the linearization point number and the local reduced basis order. Tables 8 and 11 show that the corresponding run time of the TPWL compact models are about 2 s, and that of the nonlinear model is 5400 s, so the TPWL compact models have greatly raised computational efficiency.

In practical application, for a given accuracy, the local reduced basis order should be as small as possible, but as the local reduced basis order becomes too small, some important information may be lost, and the accuracy of the TPWL compact model may be cut down. Thus, generally the local reduced basis order  $r = 5$  is selected, and then the accuracy of the TPWL compact model is handled by adding linearization points.

### Expandability of the TPWL Compact Model

In optimization and system simulation, input parameters of a compact model may be changed in certain range, so the compact model requires expandability. This means that the TPWL compact model generated for a special input parameter needs good approximation to other different input parameters.

Table 12 lists the convection intensity coefficients of three different wind speeds on the surface of cylinder with 80 mm diameter at an ambient temperature of 20°C (Ma et al. 2008, 2010). The TPWL compact model is generated for convection intensity coefficient 78.62, and then it is used to simulate for convection intensity coefficient 87.54 and 96.25. The simulated results shown in Fig. 28a demonstrate that the TPWL compact model has good expandability. And Fig. 28b also shows the expandability about electric current.

---

## Conclusions

In this chapter, we focus on a parametric model order reduction method based on implicit moment matching and a trajectory piecewise-linear method. The two methods deal with the parametric macromodeling and nonlinear macromodeling of complex MEMS devices, respectively. Numerical examples have demonstrated the effectiveness of these methods. However, in order to overcome the overall difficulties in the behavior modeling and simulation of MEMS inertial sensors, more efficient macromodeling techniques need further study.

---

## References

- Antonova EE, Looman DC (2005) Finite elements for thermoelectric device analysis in ANSYS. In: 24th international conference on thermoelectrics, pp 215–218
- Astrid P (2004) Reduction of process simulation models: a proper orthogonal decomposition approach. Dissertation, Department of Electrical Engineering, Eindhoven University of Technology
- Bedyk W, Niessner M, Schrag G, Wachutka G, Margesin B, Faes A (2008) Automated extraction of multi-energy domain macromodels demonstrated on capacitive MEMS microphones. *Sensors Actuators A Phys* 145–146:263–270
- Bond BN (2010) Stability-preserving model reduction for linear and nonlinear systems arising in analog circuit applications. Dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology
- Bond BN, Daniel L (2007) Stabilizing schemes for piecewise-linear reduced order models via projection and weighting functions. In: 2007 IEEE/ACM international conference on computer-aided design, pp 860–867
- Chang H, Zhang Y, Xu J, Yuan W (2009) Parametric model order reduction for squeeze film damping in perforated microstructures. *Microsyst Technol* 15(6):893–898
- Chang H, Zhang Y, Xie J, Zhou Z, Yuan W (2010) Integrated behavior simulation and verification for a MEMS vibratory gyroscope using parametric model order reduction. *J Microelectromech Syst* 19(2)
- Chatterjee AN, Aluru NR (2005) Combined circuit/device modeling and simulation of integrated microfluidic systems. *J Microelectromech Syst* 14(1):81–95

- Chen J, Kang S-M(S), Zou J, Liu C, Schutt-Ainé JE (2005) Macromodeling of weakly nonlinear MEMS devices with Taylor-series expansion and Arnoldi approach. *J Microelectromech Syst* 14(4):441–451
- Cho YH, Pisano AP, Howe RT (1994) Viscous damping model for laterally oscillating microstructures. *J Microelectromech Syst* 3:81–87
- Chowdhury I, Dasgupta SP (2003) Computation of Rayleigh damping coefficients for large systems. *Electron J Geotech Eng* 8:1–11
- Clark JV, Pister KSJ (2007) Modeling, simulation, and verification of an advanced micromirror using SUGAR. *J Microelectromech Syst* 16(6):1524–1536
- Coventor Inc coventorWare ARCHITECT 2008
- Davis TA, Duff IS (1997) An unsymmetric-pattern multifrontal method for sparse LU factorization. *SIAM J Matrix Anal Appl* 18(1):140–158
- Eid R, Salimbahrami B, Lohmann B, Rudnyi EB, Korvink JG (2007) Parametric order reduction of proportionally damped second-order systems. *Sens Mater* 19(3):149–164
- Freund RW (2000) Krylov-subspace methods for reduced-order modeling in circuit simulation. *J Comput Appl Math* 123(1–2):395–421
- Guo D, Chu FL (2001) The influence of rotation on vibration of a thick cylindrical shell. *J Sound Vib* 242(3):487–505
- He J, Sætrom J, Durlafsky LJ (2011) Enhanced linearized reduced-order models for subsurface flow simulation. *J Comput Phys* 230(23):8313–8341
- Jing Q, Mukherjee T, Fedder GK (2002) Large-deflection beam model for schematic behavioral simulation in NODAS. *Model Simul Microsyst*:136–139
- Liu G, Wang A, Jiang T, Jiao J, Jang JB (2008) Effects of environmental temperature on the performance of a micromachined gyroscope. *Microsyst Technol* 14:199–204
- Liu Y, Yuan W, Chang H (2014a) A global maximum error controller-based method for linearization point selection in trajectory piecewise-linear model order reduction. *IEEE Trans Comput Aided Des Integr Circuits Syst* 33(7):1100–1104
- Liu Y, Yuan W, Chang H, Ma B (2014b) Compact thermoelectric coupled models of micromachined thermal sensors using trajectory piecewise-linear model order reduction. *Microsyst Technol* 20(1):73–82
- Ma BH, Zhou BQ, Deng JJ, Yuan WZ (2008) On heat insulation of micro thermal sensor using FEA. *Chin J Sensors Actuators* 21(6):933–937
- Ma BH, Ren JZ, Yuan WZ (2010) Flexible thermal sensor array on PI film substrate for underwater applications. In: *Proceedings of 23rd IEEE international conference MEMS*, pp 679–682
- Mukherjee T, Fedder GK, Ramaswamy D, White J (2000) Emerging simulation approaches for micromachined devices. *IEEE Trans Comput Aided Des Integr Circuits Syst* 19(12):1572–1589
- Nahvi SA, Nabi M, Janardhanan S (2012) Trajectory based methods for nonlinear mor: review and perspectives. In: 2012 I.E. ISPC, pp 1–6
- Nayfeh H, Younis MI, Abdel-Rahman EM (2005) Macromodels for MEMS applications. *Nonlinear Dyn* 41(1–3):211–236
- Painter CC, Shkel AM (2001) Structural and thermal analysis of a mems angular gyroscope. *Proc SPIE Int Soc Opt Eng* 4334:86–94
- Painter CC, Shkel AM (2003) Structural and thermal modeling of a z-axis rate integrating gyroscope. *J Micromech Microeng* 13:229–237
- Reitz S et al (2003) System level modeling of microsystems using order reduction methods. *Analog Integr Circ Sig Process* 37(1):7–16
- Rewienski M, White J (2003) A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices. *IEEE Trans CAD* 22(2):155–170
- Rudnyi EB, Feng LH, Salleras M, Marco S, Korvink JG (2005) Error indicator to automatically generate dynamic compact parametric thermal models. In: *THERMINIC 2005, 11th international workshop on thermal investigations of ICs and systems, Belgirate*, pp 139–145
- Saad Y (2003) *Iterative methods for sparse linear systems*, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia

- Schrag G, Wachutka G (2002) Physically based modeling of squeeze-film damping by mixed-level system simulation. *Sensors Actuators A Phys* 97–98:193–200
- Schrag G, Wachutka G (2004) Accurate system-level damping model for highly perforated micro-mechanical devices. *Sensors Actuators A Phys* 111(2–3):222–228
- Senturia SD (2001) *Microsystem design*. Kluwer Academic Publishers, Boston
- Simoncini V, Szylid DB (2007) Recent computational developments in Krylov subspace methods for linear systems. *Numer Linear Algebra* 14(1):1–59
- Striebel M, Rommes J (2011) Model order reduction of nonlinear systems in circuit simulation: status and applications. *Lecture Notes in Electrical Engineering* 74(2):289–301. In: Benner P, Hinze M, ter Maten EJW (eds)
- Tilmans H (1996) Equivalent circuit representation of electromechanical transducers: I. Lumped-parameter systems. *J Micromech Microeng* 6:157–176
- Tiwary SK, Rutenbar RA (2005) Scalable trajectory methods for on-demand analog macromodel extraction. In: *Proceedings of the 42nd annual design automation conference*, pp 403–408
- Vandemeer JE 1998 *Nodal design of actuators and sensors (NODAS)*. M.S. Thesis, Carnegie Mellon University
- Vasilyev D, Rewiński M, White J (2003) A TBR-based trajectory piecewise-linear algorithm for generating accurate low-order models for nonlinear analog circuits and MEMS. In: *Proceedings of the 40th design automation conference*, pp 490–495
- Veijola T, Kuisma H et al (1995) Equivalent-circuit model of the squeezed gas film in a silicon accelerometer. *Sensors Actuators A Phys* 48(3): 239–248
- Veijola T, Turowski M (2001) Compact damping models for laterally moving microstructures with gas-rarefaction effects. *J Microelectromech Syst* 10:263–273
- Voss T, Verhoeven A, Bechtold T, ter Maten EJW (2008) Model order reduction for nonlinear differential algebraic equations in circuit simulation. In: Bonilla LL, Moscoso M, Platero G, Vega JM (eds) *Progress in industrial mathematics at ECMI 2006*. Series mathematics in industry, vol 12. Heidelberg, Springer, pp 518–523
- Yang Y-JJ, Yen P-C (2004) An efficient macromodeling methodology for lateral air damping effects. *J Microelectromech Syst* 13(3):812–828
- Yang YJ, Yu CJ 2002 *Technical proceedings of the 2002 international conference on modeling and simulation of microsystems*, pp 178–181
- Zhang Y, Kamalian R, Agogino AM et al (2005) Hierarchical MEMS synthesis and optimization. In: *Proceedings of 12th SPIE annual international symposium: smart structures and materials*, #5763–12
- Zhang R, Wang W, Dounavis A, Jullien GA (2008) Passive reduced-order macromodeling algorithm for microelectromechanical systems. *J Microelectromech Syst* 17(3):678–687
- Zhou N, Agogino AM (2002) Automated design synthesis for micro-electro-mechanical systems (MEMS). In: *Proceedings of the ASME design automation conference, Montreal, 29 Sept–2 Oct*



---

## **Part II**

# **Characterization and Test of Materials and Microstructures in MEMS**



# Online Test Microstructures of the Mechanical Properties for Micromachined Multilayered Films

Zai-Fa Zhou, Qing-An Huang, Xin-Ge Guo, and Yi-Fan Gu

## Contents

Overview .....	198
Extraction by Electrostatic Drive Method .....	200
Model .....	200
Extraction of Material Properties .....	206
Simulation .....	208
Pattern of Anchor .....	209
Test Structures' Design for Material Properties .....	210
Extraction by Resonance Frequency Measurements .....	214
Model .....	214
Extraction by Multilayered Doubly Clamped Beams .....	216
Extraction by Multilayered Cantilevers .....	224
Extraction by Combination of Multilayer Cantilevers and Double-Clamped Beams .....	232
Conclusion .....	233
References .....	234

## Abstract

Recently, multilayered structures have been utilized in MEMS applications, including infrared focal plane arrays, radio-frequency (RF) components, micro-machined mirrors, etc. It is well known that MEMS devices are highly dependent on material parameters such as Young's modulus and residual stress of the multilayered films. These properties determine both the final shape and the functionality of released microstructures and should therefore be accurately

Z.-F. Zhou (✉) · Q.-A. Huang

Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, Jiangsu, China

e-mail: [zfzhou@seu.edu.cn](mailto:zfzhou@seu.edu.cn); [hqa@seu.edu.cn](mailto:hqa@seu.edu.cn)

X.-G. Guo · Y.-F. Gu

Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, China

e-mail: [guoxinge93@outlook.com](mailto:guoxinge93@outlook.com); [yfgu@seu.edu.cn](mailto:yfgu@seu.edu.cn)

evaluated. Young's modulus and residual stress for single-layer films have been widely studied by the cantilever deflections, wafer curvatures, displacements of variously designed microstructures, buckling lengths, membrane deflections, resonance frequency, pull-in voltages, and double-clamped beam deflections. However, these methods are not easily extended to multilayered films. Thus it is significantly expected to directly measure both Young's modulus and residual stress for multilayer films simultaneously. This chapter presents some methods to characterize the material properties of the composite films by electrostatic pull-in testing and the resonance frequency testing approaches adopting the composite double-clamped beam or the cantilever beam. The analytical models are presented and test structures with different lengths and widths are designed. In situ methods for simultaneously extracting material properties (Young's modulus and residual stress) of each layer for the composite films are reported. The extracting methods have been confirmed by FEM simulations and experiments.

---

**Keywords**

In situ test · Test structures · Multilayer films · Double-clamped beam · Cantilever · Pull-in voltage · Resonance frequency · Young's modulus · Residual stress

---

**Overview**

With the development and modification of MEMS technology, a number of micro-mechanical components, such as cantilever, double-clamped beam, and microbridge, have been fabricated using surface micromachining technology or bulk silicon micromachining process. Due to relative small size, these micromechanical components are frequently considered as thin-film structures in macroscopic view. Differences of mechanical behavior between micromechanical components and bulk mechanical material are pretty great, so that mechanical parameters of macroscopic mechanical material are no longer suitable to measure mechanical properties of thin-film structures (Osterberg and Senturia 1997; Van Drieënhuizen et al. 1993; Petersen 1978; Thomas et al. 1975). Thin films are required to be capable of bearing mechanical load and delivering force and movement. To fully utilize mechanical performance of thin films, optimize MEMS device design, enlarge suitable material range, and enhance reliability of MEMS devices, it is crucial to grasp the relationship between manufacturing technique and mechanical properties. In addition, mechanical properties of thin films have great difference from that of bulk material with the same chemical component. Thus, conventional testing technique and devices cannot be applied to thin films directly. In a word, it is significant to control the mechanical parameters of thin films, such as residual stress, Young's modulus, and Poisson's ratio, in surface micromachining process. Studying and testing mechanical properties of thin films, which is popular these days, has intrigued researchers in micro-electronics, mechanics, and physics domain. Although a lot of researches have been done and the differences of mechanical properties between thin films and bulk

materials have been realized in a certain extent, researches are not in-depth and more problems need to be solved. With the fast development of MEMS technology, it is a fact that higher demand will be proposed for test techniques of thin films.

At present, compatibility with IC technology is one of the main development trends of MEMS development. In situ monitoring technique is important to guarantee consistency, repeatability, and realization of products (Schweitz 1992). However, in the practical fabrication of some MEMS devices, consistency of products cannot be certainly guaranteed. Because material has influence on structure and properties of MEMS devices, tests of material parameters are appeared to be especially crucial. Many materials, especially crystalline material, will have obviously different thermal parameters in the same fabrication process but different fabrication environment. In addition, the same structure but different fabrication method will also lead to different mechanical parameters. If some mechanical parameters, such as density, Young's modulus, and residual stress, are known, static or dynamic response of sensors and actuators can be estimated by them. Thus, in situ monitoring mechanical parameters of thin films is vital to MEMS devices (Van Drieënhuizen et al. 1993; Petersen 1978; Najafi and Suzuki 1989).

Mechanical parameters, which are not obtained by in situ monitoring tests, have a lot of problems. First, data, which are obtained under different fabrication processes, spacemen size, and test instruments, lack of universality and authority. Second, data obtained are incomplete in a certain extent. Third, with the appearance of new materials and process, lack of quick responding mechanism to collect new data is the biggest drawback of it.

Mechanical, thermal, and electrical parameters of MEMS thin films are closely related to process conditions. Thin-film deposition frequently causes various defects, and the density of defects is different from that of microstructures depending on different deposition methods. The discrepancy of results is caused by different measure and deposition methods. Therefore, not only suitable test methods and structures can obtain physical parameters required, but also test structures can be used to monitor process. It is required for in situ monitoring to be simple, feasible, and repeatable in the test environment. In addition, test structures should be simple-designed and away from destructiveness and occupy small area of chips.

Considering all the prerequisites for in situ testing mentioned above, two types of method have shown their potential: electrostatic drive method and resonance frequency method. Pull-in voltage testing (Najafi and Suzuki 1989; Nie et al. 2005; Osterberg and Senturia 1997; Petersen 1978; Zou et al. 1995) and capacity-voltage testing (Chu et al. 2002; Chan et al. 1999; Wang et al. 1992) are mostly used currently. While the capacity of MEMS structure is always extremely small, the latter method has stringent requirements for testing equipments as well as the testing environment. Although pull-in phenomenon is unstable and always destructive, proper design of the testing structures would provide desirable results before it ever happens. The method of using resonance frequency to extract mechanical properties has been proposed back in 1971 (Petersen and Guarnieri 1979). From then on, researchers have developed various testing methods using similar method (De Coster et al. 2011; Ilic et al. 2010; Li et al. 2003). Tilmans et al. presented micro-resonant force gauges (Tilmans et al. 1992), which are

capable of testing multiple mechanical properties. Resonance frequency method has been proven accurate, while the measurement is relatively convenient. As a matter of fact, however, the fabricated structures always suffer from various defects, among them the stress-induced buckling. Therefore, further investigation is still needed in improving these methods.

Moreover, in practical MEMS devices such as RF switches and acoustic resonators, it is evitable to use multilayer structures. Generally, one or several layers are developed on the thin film by bonding or sputtering. When electrostatic actuation is required, conducting layer should be plated on insulating thin-film materials. In some cases, tensile stress and compressive stress can make the whole thin film have deflection. To counteract this kind of effect, a layer of thin film, which has the opposite stress, can be developed on it. From the above, multilayer structures are widely applied to many MEMS devices. Therefore, the chapter is mainly about in situ extraction of material parameters for multilayer thin films.

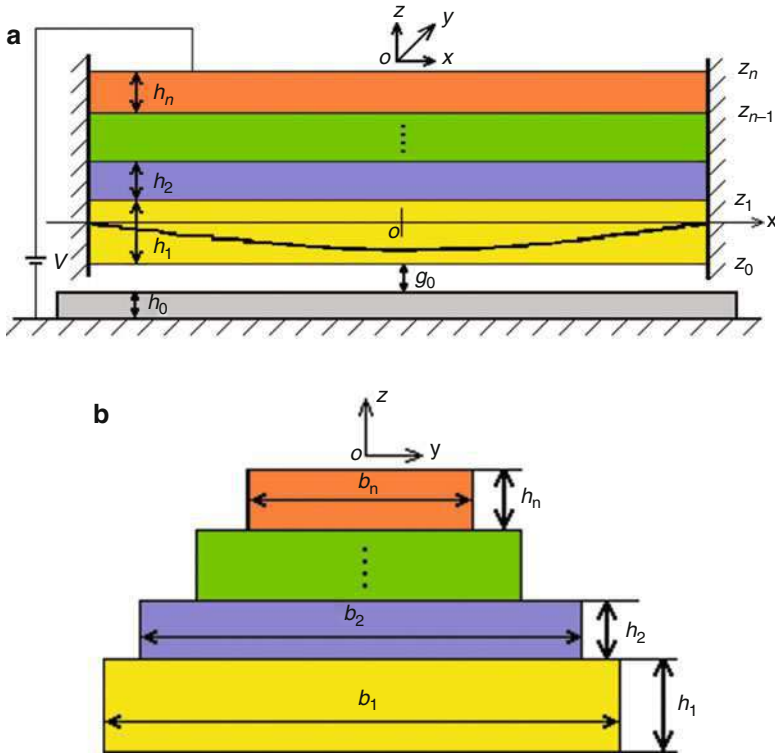
---

## Extraction by Electrostatic Drive Method

### Model

For an  $n$ -layer doubly clamped composite beam, as shown in Fig. 1, only the top layer of the beam is temporarily assumed to be a conductor, while other left layers are all dielectric. A fixed planar electrode (ground) coated by a dielectric layer underlies the beam. It is assumed that the microbeam is flat and is not bent without applied voltage. When a DC voltage is applied between the beam and ground, the beam will bend due to the attractive electrostatic force. The length of the beam is  $l$ ; the width, thickness, Young's modulus, and Poisson's ratio of the  $i$ th layer are  $b_i$ ,  $h_i$ ,  $E_i$ , and  $\nu_i$ , respectively. The effective modulus is  $\tilde{E}_i$ ; it is related to the beam width. The beam is considered to be wide when the width of the beam is not smaller than five times thickness of the beam. The wide beams exhibit plane-strain conditions, and, therefore,  $\tilde{E}_i$  becomes the plate modulus  $E/(1 - \nu^2)$ , and  $\tilde{\sigma}_i$  becomes the plate residual stress  $\sigma(1 + \nu)$ . The beam is considered to be narrow when the width of the beam is smaller than five times thickness of the beam, and in this case,  $\tilde{E}_i$  reduces simply Young's modulus  $E$ , and  $\tilde{\sigma}_i$  reduces simply Young's modulus  $\sigma$ . The relative permittivity of the  $i$ th layer is  $\epsilon_{ri}$  (since the top layer is a conductor, no relative permittivity should be considered),  $\epsilon_0$  is the permittivity of vacuum, and the thickness and relative permittivity of the coating dielectric layer on the fixed electrode are  $h_0$  and  $\epsilon_{r0}$ , respectively.  $g_0$  is the initial gap between the bottom of the beam and the top of the coating dielectric.  $V$  is the applied voltage.

Assuming there is residual stress in the  $i$ th layer, since both ends are fixed, axial stretching of the beam will yield when the beam bends. Then bending of the beam will cause the bending strain, stretching strain along the length direction of the beam, and the strain caused by the residual stress along the length direction. On the other hand, the difference between the single-layer beam and the composite beam is that



**Fig. 1** (a) An n-layer doubly clamped composite beam under an electrostatic load. (b) A cross-sectional view of an n-layer doubly clamped composite beam

lateral residual stress along the width direction occurs in the composite beam, which will affect the behavior of the beam.

Assuming the lateral strain along the width direction of the beam is zero in order to maintain continuum in the plate during bending, it yields (Timoshenko and Woinowsky-Krieger 1959):

$$\epsilon_x = \frac{\sigma_x}{E} - \frac{\nu\sigma_y}{E}, \quad \epsilon_y = \frac{\sigma_y}{E} - \frac{\nu\sigma_x}{E} = 0, \quad \sigma_y = \nu\sigma_x \tag{1}$$

where  $\epsilon_x$  and  $\epsilon_y$  are the strains along the length and the width of the beam, respectively.  $\sigma_x$  and  $\sigma_y$  are the residual stresses along the length and the width of the beam, respectively.

Then the effect caused by the lateral residual stress can be denoted by the axial residual stress. So the total strain  $\epsilon_{Ti}$  in the  $i$ th layer of the composite beam can be defined as follows:

$$\epsilon_{Ti} = \epsilon_{bi} + \epsilon_{si} + \frac{(1 + \nu_i)\sigma_i}{\tilde{E}_i} \tag{2}$$

where  $\sigma_i$ ,  $\varepsilon_{bi}$ , and  $\varepsilon_{si}$  represent the residual stress, bending strain, and stretching strain of the  $i$ th layer along the length/axial direction of the beam, respectively. X-axis is taken to be the neutral axis. The distance between the neutral axis and the bottom of the beam is  $z_c$ . Since the beam is assumed to be in the plane,  $z_c$  can be given by (Nie et al. 2005)

$$z_c = \frac{1}{2} \cdot \frac{\sum_{i=1}^n \tilde{E}_i b_i \left( \left( \sum_{j=1}^i 2h_j \right) - h_i \right) \cdot h_i}{\sum_{i=1}^n \tilde{E}_i b_i h_i} \quad (3)$$

$$z_0 = -z_c \quad (4)$$

$$z_i = \left( \sum_{j=1}^i h_j \right) - z_c, (i = 1 \cdots n) \quad (5)$$

where  $z_0$  and  $z_n$  are the distances between the top and bottom of the beam and the neutral axis, respectively.  $z_1, \dots, z_{n-1}$  are the distances between interfaces of the corresponding layers and the neutral axis, respectively.

When the beam bends due to the electrostatic force, a trial function in the form of a cosine is still taken here (Nie et al. 2005):

$$w(x) = \frac{c}{2} \left( 1 + \cos \frac{2\pi x}{l} \right) \quad (6)$$

Then the total strain  $\varepsilon_{Ti}$  in the  $i$ th layer can be written as

$$\varepsilon_{Ti} = -z \frac{d^2 w}{dx^2} + \frac{1}{l} \int_{-\frac{l}{2}}^{\frac{l}{2}} \frac{1}{2} \left( \frac{dw}{dx} \right)^2 dx + \frac{\tilde{\sigma}_i}{\tilde{E}_i} \quad (7)$$

The total elastic energy of the beam is

$$W_{elastic} = \sum_{i=1}^n \frac{\tilde{E}_i b_i}{2} \int_{z_{i-1}}^{z_i} \int_{-\frac{l}{2}}^{\frac{l}{2}} \varepsilon_{Ti}^2 dz dx \quad (8)$$

then

$$\begin{aligned} W_{elastic} = c^2 \left[ \sum_{i=1}^n \frac{\tilde{E}_i b_i}{3} \cdot \frac{\pi^4}{l^3} \cdot (z_i^3 - z_{i-1}^3) + \sum_{i=1}^n \frac{\pi^2}{4l} \cdot b_i \tilde{\sigma}_i h_i \right] \\ + c^4 \cdot \sum_{i=1}^n \frac{\tilde{E}_i b_i \pi^4}{32l^3} h_i + \sum_{i=1}^n \frac{b_i}{2} \cdot \frac{\tilde{\sigma}_i^2}{\tilde{E}_i} h_i \end{aligned} \quad (9)$$

The co-energy can also be adopted as previous (Rong et al. 2004):

$$W_e^* = \frac{V^2}{2} \epsilon_0 b_e \int_{-\frac{l}{2}}^{\frac{l}{2}} \frac{dx}{\tilde{g}_0 - \frac{c}{2} \left( 1 + \cos \frac{2\pi x}{l} \right)} = \frac{V^2}{2} \epsilon_0 b_e \frac{l}{\sqrt{(\tilde{g}_0 - c)\tilde{g}_0}} \tag{10}$$

Let

$$\tilde{g}_0 = g_0 + \frac{h_0}{\epsilon_{r0}} + \frac{h_1}{\epsilon_{r1}} + \frac{h_2}{\epsilon_{r2}} + \dots + \frac{h_{n-1}}{\epsilon_{r(n-1)}} \tag{11}$$

where  $\tilde{g}_0$  is the effective gap and  $b_e$  is the width of beam electrode.

The total energy  $W$  consists of elastic energy  $W_{elastic}$  and electrostatic energy  $W_e$ :

$$W = W_{elastic} + W_e \tag{12}$$

By the principle of virtual work, we obtain

$$\frac{\partial W}{\partial c} = \frac{\partial W_{elastic}}{\partial c} + \frac{\partial W_e}{\partial c} = 0 \tag{13}$$

For energy domains using co-energy, we shall make substitution (Gabbay 1998):

$$\frac{\partial W_e}{\partial c} = - \frac{\partial W_e^*}{\partial c} \tag{14}$$

therefore,

$$\frac{\partial W}{\partial c} = \frac{\partial W_{elastic}}{\partial c} - \frac{\partial W_e^*}{\partial c} = 0 \tag{15}$$

substituting Eqs. (9) and (10) into Eq. (15) yields

$$\frac{V^2}{4} \cdot \frac{\epsilon_0 b_e}{\sqrt{\tilde{g}_0} \cdot (\tilde{g}_0 - c)^{3/2}} = \tag{16}$$

$$2c \left[ \sum_{i=1}^n \frac{\tilde{E}_i b_i}{3} \cdot \left( \frac{\pi}{l} \right)^4 \cdot (z_i^3 - z_{i-1}^3) + \sum_{i=1}^n \frac{\pi^2}{4l^2} \cdot b_i \tilde{\sigma}_i h_i \right] + \frac{\pi^4}{8l^4} \cdot c^3 \sum_{i=1}^n \tilde{E}_i h_i b_i$$

with

$$A = 2 \left[ \sum_{i=1}^n \frac{\tilde{E}_i b_i}{3} \cdot \left( \frac{\pi}{l} \right)^4 \cdot (z_i^3 - z_{i-1}^3) + \sum_{i=1}^n \frac{\pi^2}{4l^2} \cdot b_i \tilde{\sigma}_i h_i \right] \tag{17}$$

$$B = \frac{\pi^4}{8l^4} \sum_{i=1}^n \tilde{E}_i h_i b_i \tag{18}$$



We can obtain

$$Ac + Bc^3 = \frac{V^2}{4} \cdot \frac{\varepsilon_0 b_e}{\sqrt{\tilde{g}_0} \cdot (\tilde{g}_0 - c)^{\frac{3}{2}}} \quad (19)$$

We can obtain  $c$  from Eq. (19) by making use of the iteration method.

At the pull-in state, the beam is in the balance under all of the forces, and then Eq. (19) can be written as

$$Ac_{PI} + Bc_{PI}^3 = \frac{V_{PI}^2}{4} \cdot \frac{\varepsilon_0 b_e}{\sqrt{\tilde{g}_0} \cdot (\tilde{g}_0 - c_{PI})^{\frac{3}{2}}} \quad (20)$$

where  $c_{PI}$  is the pull-in displacement of the center of the beam.

At the pull-in state, the acceleration of the beam should be zero, that is, the following equation should also be satisfied besides Eq. (20) (Gupta 1998):

$$\frac{\partial^2 W}{\partial c^2} = 0 \quad (21)$$

that is

$$A + 3Bc_{PI}^2 = \frac{V_{PI}^2}{4} \cdot \frac{3}{2} \cdot \frac{\varepsilon_0 b_e}{\sqrt{\tilde{g}_0} \cdot (\tilde{g}_0 - c_{PI})^{\frac{5}{2}}}. \quad (22)$$

From Eqs. (20) and (22), it yields

$$\frac{Ac_{PI} + Bc_{PI}^3}{A + 3Bc_{PI}^2} = \frac{2}{3} \cdot (\tilde{g}_0 - c_{PI}) \quad (23)$$

From Eq. (23), we can obtain

$$9\beta^3 - 6\beta^2 + 5D\beta - 2D = 0 \quad (24)$$

where  $\beta = c_{PI}/\tilde{g}_0$ ,  $\beta$  is known as the normalized displacement:

$$D = \frac{A}{B\tilde{g}_0^2} \quad (25)$$

The real root solution to Eq. (24) is given by

$$\beta = \frac{1}{9} \cdot \left( 36D + 8 + 3\sqrt{375D^3 - 156D^2 + 144D} \right)^{\frac{1}{3}} - \frac{15D - 4}{9 \cdot \left( 36D + 8 + 3\sqrt{375D^3 - 156D^2 + 144D} \right)^{\frac{1}{3}}} + \frac{2}{9} \quad (26)$$

From Eq. (26), the pull-in displacement of the center of the beam can be obtained:

$$c_{PI} = \beta \tilde{g}_0 \quad (27)$$

Substituting Eq. (27) into Eq. (20), we can obtain the pull-in voltage:

$$V_{PI} = \sqrt{4(AC_{PI} + BC_{PI}^3) \cdot \sqrt{\tilde{g}_0} \cdot (\tilde{g}_0 - c_{PI})^{3/2} / \varepsilon_0 / b_e} \quad (28)$$

From Eq. (9), it yields

$$\sum_{i=1}^n \tilde{E}_i b_i (z_i^3 - z_{i-1}^3) = \tilde{E}_{\text{eff}} \cdot \frac{h_{\text{eff}}^3}{4} \cdot b_e \quad (29)$$

$$\sum_{i=1}^n \tilde{\sigma}_i h_i b_i = \tilde{\sigma}_{\text{eff}} \cdot h_{\text{eff}} \cdot b_e \quad (30)$$

$$\sum_{i=1}^n \tilde{E}_i h_i b_i = \tilde{E}_{\text{eff}} \cdot h_{\text{eff}} \cdot b_e \quad (31)$$

Combining Eqs. (29), (30), and (31) yields

$$h_{\text{eff}} = 2 \cdot \sqrt{\frac{\sum_{i=1}^n \tilde{E}_i b_i (z_i^3 - z_{i-1}^3)}{\sum_{i=1}^n \tilde{E}_i h_i b_i}} \quad (32)$$

$$\tilde{\sigma}_{\text{eff}} = \sum_{i=1}^n \tilde{\sigma}_i h_i b_i / h_{\text{eff}} / b_e \quad (33)$$

$$\tilde{E}_{\text{eff}} = \sum_{i=1}^n \tilde{E}_i h_i b_i / h_{\text{eff}} / b_e \quad (34)$$

where  $h_{\text{eff}}$ ,  $\tilde{\sigma}_{\text{eff}}$ , and  $\tilde{E}_{\text{eff}}$  are the total effective thickness, the total effective residual stress, and the total effective modulus of the composite fixed-fixed beam, respectively.

The analytical expression in Eq. (28) is modified by considering fringing effect and additional fitting parameters, in which the additional fitting parameters are obtained by the least-square method. The final analytical model for pull-in voltage of the composite film is written as

$$V_{PI} = \gamma \sqrt{\frac{4(AC_{PI} + BC_{PI}^3) \cdot \sqrt{\tilde{g}_0} \cdot (\tilde{g}_0 - c_{PI})^{3/2}}{\varepsilon_0 \cdot b_e \cdot (1 + 0.42\tilde{g}_0/b_e)}} \quad (35)$$

where  $\gamma = 1/[1.1454(1 - \gamma_2 e^{\gamma_3})]$ ,  $\gamma_1 = 1.0874 \times 10^{-2} \cdot \tilde{g}_0/h_{\text{eff}}$ ,  $\gamma_2 = 1 - 0.8894 \cdot e^{\gamma_1}$ , and  $\gamma_3 = -40551(\tilde{\sigma}_{\text{eff}}^2)^{0.3115} / (\tilde{E}_{\text{eff}} \cdot h_{\text{eff}})$ .

**Table 1** Comparison of pull-in voltages obtained by CoventorWare<sup>®</sup>, the present pull-in voltage model, and the previous pull-in voltage model

l (μm)	Vc (V)	V2D (V)	V1D (V)
100	322.5	313.2	303.3
110	275	267.9	258.2
120	239	233.1	223.8
130	208.9	205.8	196.8
140	185.1	184	175.2
150	165	166.1	157.6
200	106.2	111.3	104.1

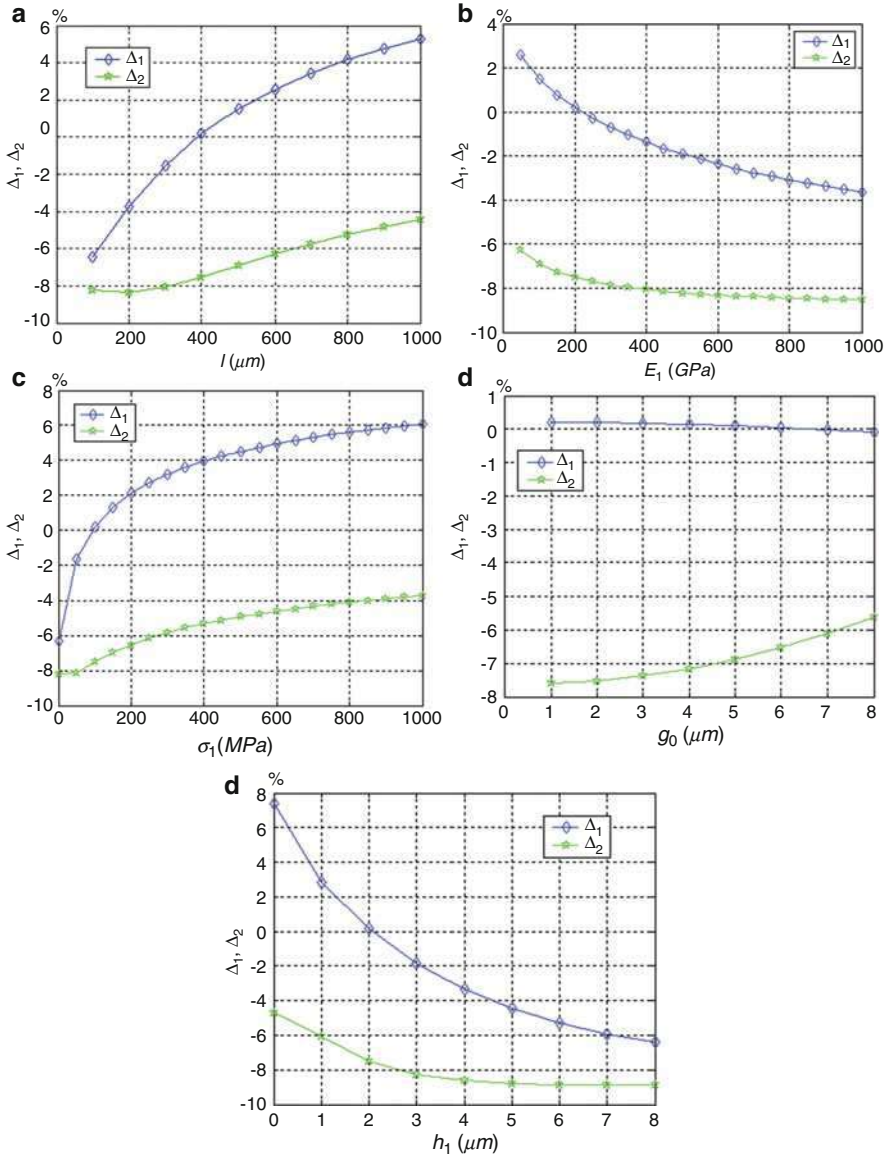
$b_1 = b_2 = b_e = 100 \mu\text{m}$ ,  $h_1 = 1 \mu\text{m}$ ,  $h_2 = 0.8 \mu\text{m}$ ,  $g_0 = 2 \mu\text{m}$ ,  
 $E_1 = 201 \text{ GPa}$ ,  $E_2 = 57 \text{ GPa}$ ,  $\sigma_1 = 100 \text{ MPa}$ ,  $\sigma_2 = 12 \text{ MPa}$ ,  $\nu_1 = 0.22$ ,  
 $\nu_2 = 0.3$ .

In the above, an assumption has been made that only the top layer of the beam is a conductor. If the conductor layer is not only the top layer, or there are many conductor layers in the composite beams, the effective gap  $\tilde{g}_0$  and the width of beam electrode should be calculated simply according to the actual situation.

Table 1 is the comparison of pull-in voltages, where Vc, V2D, and V1D are the pull-in voltages obtained by CoventorWare<sup>®</sup>, the present pull-in voltage model, and our previous pull-in voltage model, respectively. Figure 2a–e shows relations between relative errors and geometries or material properties of the test structures.  $\Delta_1$  is defined as the relative error between the pull-in voltages obtained from Eq. (35) and the CoventorWare<sup>®</sup>;  $\Delta_2$  is defined as the relative error between the pull-in voltages obtained from our previous work without the lateral residual stress of the beam and CoventorWare<sup>®</sup> (Nie et al. 2005). It is clear from Fig. 2 that the improved model for the pull-in voltage has higher accuracy than the previous model. From the analysis results, it can be seen that the thickness of the microbeam has an important effect on the error.

## Extraction of Material Properties

Since the pull-in voltage is dependent of material properties and geometries of the beam, Young's modulus and residual stress of each layer of the beam can be extracted by using a set of pull-in voltages. In this chapter, a method to extract Young's modulus and residual stresses of the composite films is presented by making a set of beams with different geometries. For an  $n$ -layer film, at least  $n$  kinds of beams with different lengths should be designed, and each kind of beams with the same length needs two different types: one has the same width of each layer, while the other has different widths of each layer to extract Young's modulus and residual stress of each layer.  $2n$  pull-in voltages can be measured from these  $2n$  beams. Through Eq. (35),  $(2n)$  nonlinear equations can be obtained as follows:



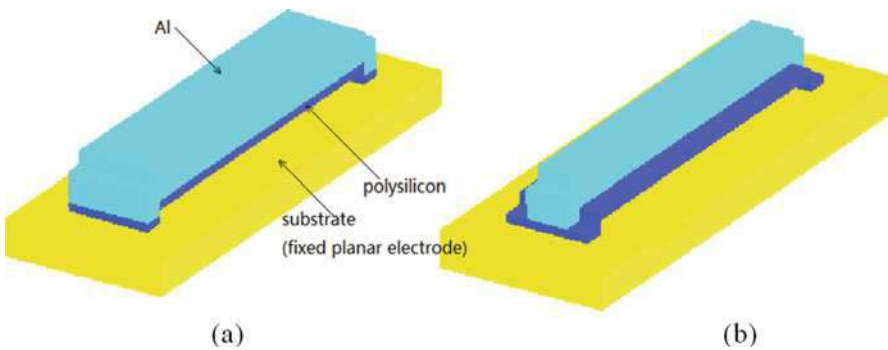
**Fig. 2** (a) Relative error vs. length of the beam ( $b_1 = b_2 = b_e = 100 \mu\text{m}$ ,  $h_1 = 1 \mu\text{m}$ ,  $h_2 = 0.8 \mu\text{m}$ ,  $g_0 = 2 \mu\text{m}$ ,  $E_1 = 201 \text{ GPa}$ ,  $E_2 = 53 \text{ GPa}$ ,  $\sigma_1 = 100 \text{ MPa}$ ,  $\sigma_2 = 12 \text{ MPa}$ ,  $\nu_1 = 0.22$ ,  $\nu_2 = 0.3$ ). (b) Relative error vs. residual stress in the bottom layer ( $l = 400 \mu\text{m}$ ,  $b_1 = b_2 = b_e = 100 \mu\text{m}$ ,  $h_1 = 2 \mu\text{m}$ ,  $h_2 = 0.8 \mu\text{m}$ ,  $g_0 = 2 \mu\text{m}$ ,  $E_1 = 201 \text{ GPa}$ ,  $E_2 = 53 \text{ GPa}$ ,  $\sigma_2 = 12 \text{ MPa}$ ,  $\nu_1 = 0.22$ ,  $\nu_2 = 0.3$ ). (c) Relative error vs. Young's modulus in the bottom layer ( $l = 400 \mu\text{m}$ ,  $b_1 = b_2 = b_e = 100 \mu\text{m}$ ,  $h_1 = 2 \mu\text{m}$ ,  $h_2 = 0.8 \mu\text{m}$ ,  $g_0 = 2 \mu\text{m}$ ,  $E_2 = 53 \text{ GPa}$ ,  $\sigma_1 = 100 \text{ MPa}$ ,  $\sigma_2 = 12 \text{ MPa}$ ,  $\nu_1 = 0.22$ ,  $\nu_2 = 0.3$ ). (d) Relative error vs. initial gap ( $l = 400 \mu\text{m}$ ,  $b_1 = b_2 = b_e = 100 \mu\text{m}$ ,  $h_1 = 2 \mu\text{m}$ ,  $h_2 = 0.8 \mu\text{m}$ ,  $E_1 = 201 \text{ GPa}$ ,  $E_2 = 53 \text{ GPa}$ ,  $\sigma_1 = 100 \text{ MPa}$ ,  $\sigma_2 = 12 \text{ MPa}$ ,  $\nu_1 = 0.22$ ,  $\nu_2 = 0.3$ ). (e) Relative error vs. thickness of the bottom layer ( $l = 400 \mu\text{m}$ ,  $b_1 = b_2 = b_e = 100 \mu\text{m}$ ,  $h_2 = 0.8 \mu\text{m}$ ,  $g_0 = 2 \mu\text{m}$ ,  $E_1 = 201 \text{ GPa}$ ,  $E_2 = 53 \text{ GPa}$ ,  $\sigma_1 = 100 \text{ MPa}$ ,  $\sigma_2 = 12 \text{ MPa}$ ,  $\nu_1 = 0.22$ ,  $\nu_2 = 0.3$ )

$$\begin{cases} V_{PI1}(\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n, \tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_n) - V_{PI1t} = 0 \\ \dots\dots\dots \\ V_{PIn}(\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n, \tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_n) - V_{PInt} = 0 \\ \dots\dots\dots \\ V_{PI2n}(\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n, \tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_n) - V_{PI2nt} = 0 \end{cases} \quad (36)$$

where  $\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n, \tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_n$  are the effective modulus and the effective residual stress of each layer to be extracted,  $V_{PIi}(i = 1, 2, \dots, 2n)$  are the nonlinear function of  $\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n, \tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_n$  given by Eq. (35), and  $V_{PIi}(i = 1, 2, \dots, 2n)$  are the measured pull-in voltages. Material properties can be calculated by using the Newton iteration and least-square method. From the equations three solutions can be obtained. However, according to the practical condition that the material properties such as polysilicon, metal, etc., are always within a reasonable range, two of the solutions greatly deviate from the range, and they are omitted.

**Simulation**

A doubly clamped bilayer beam is taken here to verify the proposed method for extracting effective modulus and residual stresses. Figure 3 shows the doubly clamped bilayer beam made of Al and polysilicon. The top layer is Al, while the bottom layer is polysilicon. The Al layer and polysilicon layer have the same width in Fig. 3a. The Al layer is a half width of the polysilicon layer in Fig. 3b. There is no dielectric layer coated on the fixed planar electrode (ground). In order to extract the effective modulus and residual stress of these two layers, four beams with two different kinds of lengths (200, 250 μm) are needed. The geometry and material properties of the beams are shown in Table 2. Pull-in voltages were calculated first by CoventorWare<sup>®</sup> using parameters in Table 2. The simulated pull-in voltages and the geometric dimensions were then substituted into Eq. (35). The effective modulus and residual stresses may be calculated by Eq. (35). Table 3 gives the comparison of



**Fig. 3** (a) A doubly clamped bilayer beam with the same width of layers. (b) A doubly clamped bilayer beam with the different widths of layers

**Table 2** Geometries and material properties of the doubly clamped bilayer beams

Length $l$ ( $\mu\text{m}$ )		200,250
Width ( $\mu\text{m}$ )	$b_1$	50
	$b_2$	50,25
Gap $g_0$ ( $\mu\text{m}$ )		1.0
Thickness ( $\mu\text{m}$ )	$h_1$	0.5
	$h_2$	1.8
Effective modulus (GPa) (set)	$\tilde{E}_1$	201
	$\tilde{E}_2$	80
Effective residual stress (MPa) (set)	$\tilde{\sigma}_1$	100
	$\tilde{\sigma}_2$	0

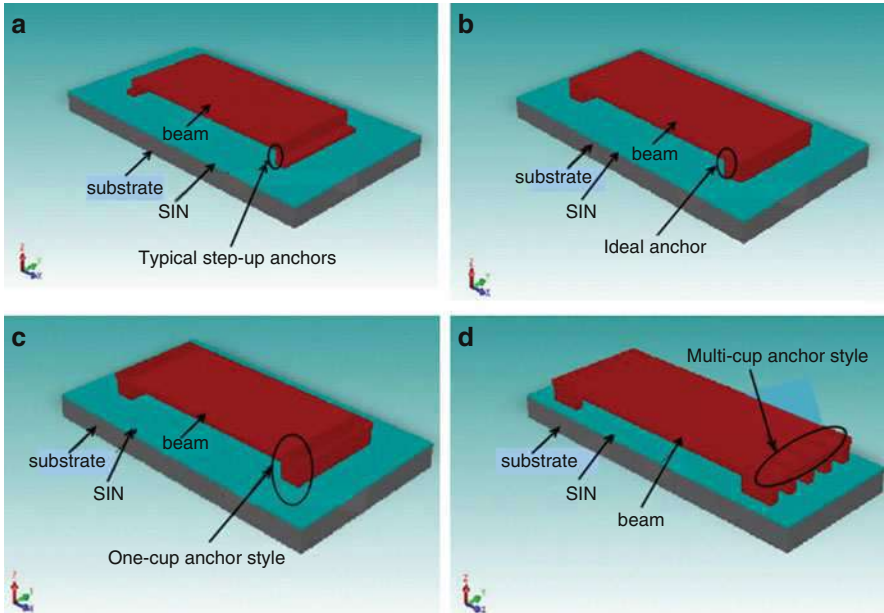
**Table 3** Calculated results of material properties

Material properties	Set values	Calculated results	Error
$\tilde{E}_1$ (GPa)	201	201.80	0.4%
$\tilde{E}_2$ (GPa)	80	81.52	1.9%
$\tilde{\sigma}_1$ (MPa)	100	103.33	3.3%
$\tilde{\sigma}_2$ (MPa)	0	-0.19	~

material properties calculated from Eq. (35) with the set parameters, where the error is defined as the relative error of the modulus and residual stress between the effective values calculated by Eq. (35) and the set one.

### Pattern of Anchor

The anchors fabricated by surface micromachining technology are usually typical step-up anchors (Fig. 4a), which are often not well-defined fixed ends (Fig. 4b) and may cause considerable testing errors in reality. Since surface micromachined beams with nonideal compliant supports are not suitable for the model, to increase the connecting area between the substrate and the anchor, a novel anchor that is constituted by some cavum cup can be designed (Hua et al. 2004). In order to prove the validity of the multi-cup anchor in improving the boundary fixed condition, compared with the one-cup anchor style and typical step-up anchor style, the FEM method by CoventorWare<sup>®</sup> simulations is used. The one-cup and multi-cup anchor styles are shown in Fig. 4c, d. The length, width, thickness, Young’s modulus, and Poisson’s ratio of the microbeam are 100, 50, 0.5  $\mu\text{m}$ , 165 GPa, and 0.23, respectively. The initial gap between the beam and the ground is 1.5  $\mu\text{m}$ . A pressure force 0.015 MPa is loaded on the microbeam vertically. The results are shown in Table 4, where the error and error are defined as the relative errors between each kind anchor style and the ideal anchor with different residual stresses, respectively. From simulation results, it can be seen that the multi-cup anchor is more



**Fig. 4** (a) Typical step-up anchor fabricated by surface micromachining technology. (b) Ideal fixed-end anchor. (c) One-cup anchor style. (d) Multi-cup anchor style

**Table 4** Central deflection of the double-clamped beam and error between each anchor style and ideal anchor

Style of anchor	Displacement <sup>a</sup> ( $\mu\text{m}$ )	Error <sup>a</sup>	Displacement <sup>b</sup> ( $\mu\text{m}$ )	Error <sup>b</sup>
Ideal fixed-end anchor	0.8153	–	0.3431	–
Typical step-up anchor	1.0453	28.2%	0.5820	69.6%
One-cup anchor	0.8541	4.7%	0.4776	39.2%
Multi-cup anchor	0.8385	2.8%	0.3564	3.9%

<sup>a</sup>Residual stress in the beam: 0 MPa

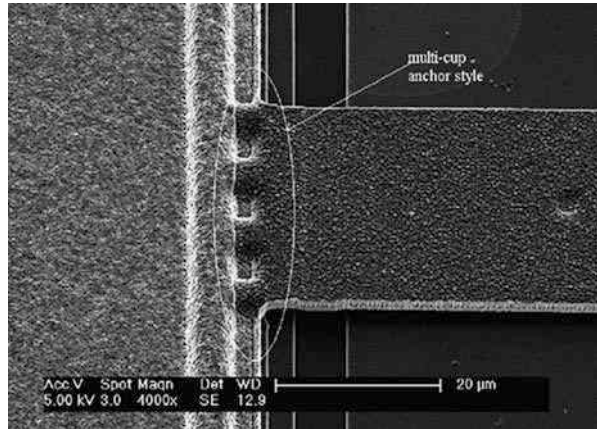
<sup>b</sup>Residual stress in the beam: 100 MPa

proper for the fixed condition assumption. Figure 5 shows the SEM of the micro-beam with the multi-cup anchor style.

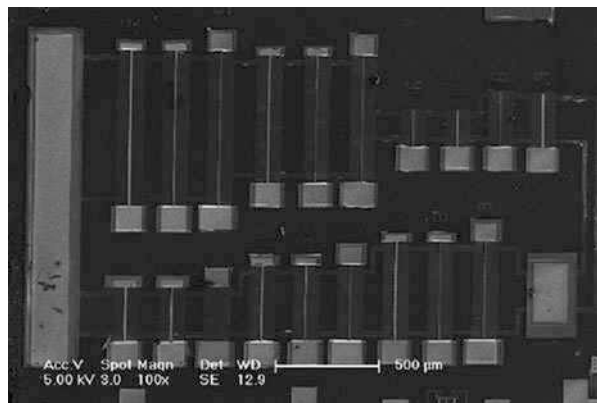
## Test Structures' Design for Material Properties

It can be seen from Fig. 2a that the influence of the lateral residual stress decreases as the ratio of length to width of the composite beam increases from the analysis of the improved model. Based on this improved model and the compliance of the anchor, we designed the suitable geometry dimension and anchor of test structures to extract material properties. As stated in section “[Extraction of Material Properties](#),” to extract Young’s modulus and residual stresses of each layer, each kind of beam

**Fig. 5** SEM of a microbeam with multi-cup anchor style



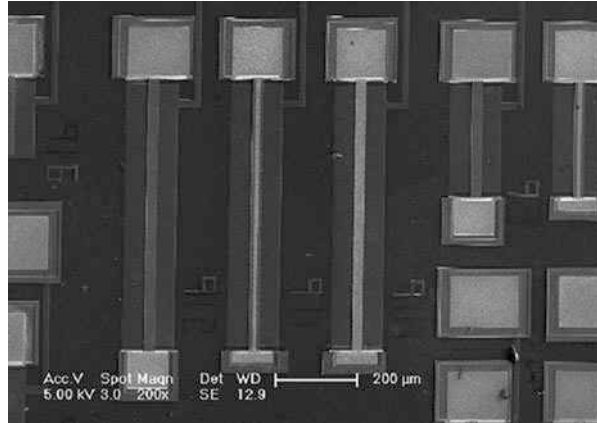
**Fig. 6** SEM photograph 1: test structures for extracting material properties



with the same length needs two different types: one has the same width of each layer, while the other has different widths of each layer. Five kinds of beam with different lengths (300, 420, 550, 700, 850 μm) were fabricated. Test structures were fabricated by IME of PKU. double-clamped beam are the bilayer. The top layer is Al, and the bottom layer is polysilicon. Figures 6 and 7 show the SEM photographs of the test structures. Parameters of the test structures are shown in Table 5. All the test experiments were processed at 25 °C. The experimental setup for the pull-in measurement is shown in Fig. 8. A TZ-103 Manual Probe Station and a HP4145B Semiconductor Parametric Analyzer were used to measure pull-in voltages of the test structures. The HP4145B was programmed to slowly ramp the voltage at 1 volt over a specified voltage range until pull-in was detected on the HP4145B I/V screen output interface as a sudden step in the current, which will increase from zero to several microamperes. The measured pull-in voltages are given in Table 6.  $V_{PI1}$ ,

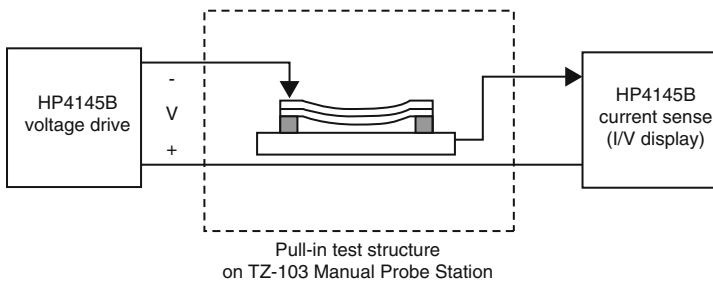


**Fig. 7** SEM photograph 2: test structures for extracting material properties



**Table 5** Parameters of the bilayer double-clamped beams

Width ( $\mu\text{m}$ )		Gap $g_0$ ( $\mu\text{m}$ )	Thickness ( $\mu\text{m}$ )	
$b_1$	$b_2$		$h_1$	$h_2$
20	20,10,0	2	2	0.8



**Fig. 8** Pull-in test setup

**Table 6** Measured pull-in voltages and calculated pull-in voltages from Eq. (35) with the extracted material properties

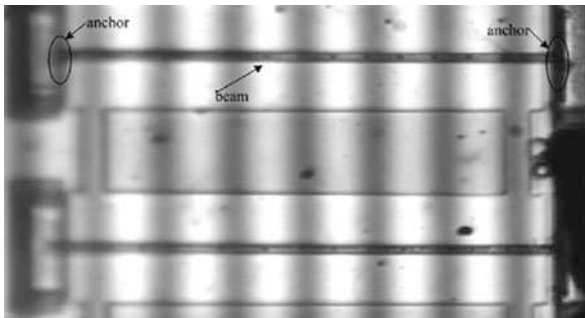
Length $l$ ( $\mu\text{m}$ )	$V_{PI1}$ (V)	$V_{PI2}$ (V)	$V_{PI3}$ (V)	$V'_{PI1}$ (V)	$V'_{PI2}$ (V)	$V'_{PI3}$ (V)
300	68.3	64.2	59.6	67.7	63.4	58.4
420	42	39.9	37.6	42.4	40.1	37.4
550	26.7	28.3	26.9	30.2	28.7	27
700	22	21.1	20.3	22.8	21.7	20.6
850	17.5	16.9	16.3	18.3	17.5	16.7

$V_{PI2}$ , and  $V_{PI3}$  are the pull-in voltages when the widths of Al layer are 20, 10, 0  $\mu\text{m}$ , respectively. The ramp speed of 1 volt is slow enough to insure no dynamic effect during the bending of the test structures. The test structures were maintained in their static regime during the voltage ramp. It should be pointed out that, for no dielectric layer between the doubly clamped conductive microbeams, a resistor with a large value in the series circuit is used to prevent circuit short during the testing.

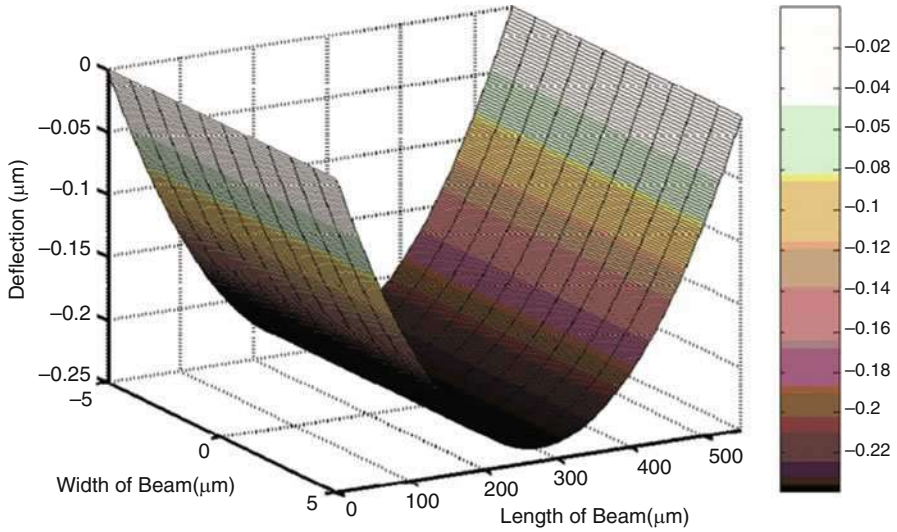
By using the measured geometric dimensions provided by IME of PKU (Table 5) and measured pull-in voltages (Table 6), the effective modulus and residual stresses are solved simultaneously by Eq. (35). For the beams with two different lengths, there are four corresponding test structures. The measured four pull-in voltages are used to extract the effective modulus and the residual stresses of each layer. In our design, there are beams with four different lengths, and hence, there are four different sets of pull-in voltages. The results are as follows: Young's modulus of polysilicon is 145 GPa, residual stress of polysilicon is 31.3 MPa, Young's modulus of Al is 41.8 GPa, and residual stress of Al is 13.2 MPa. To confirm the pull-in model, the optical method was used to measure the maximum displacement of the bilayer doubly clamped beam.

Figure 9 is the fringe projection figure of the beam by micro-fringe projection system. Figure 10 shows the bending of the beam when a 22 V DC voltage was applied to a bilayer fixed-fixed beam with a length of 550  $\mu\text{m}$ . When the voltage was applied to the same test structure, the maximum displacement of the beam (0.25  $\mu\text{m}$ ) was simulated by CoventorWare<sup>®</sup> with the measured material properties.

Compared with the available data in Table 7, it is verified that the experimental results are reasonable. Substituting the extracted material properties into Eq. (35), the theoretical pull-in voltages can be calculated. They are also listed in Table 6, and are the calculated pull-in voltages when the widths of Al layer are 20, 10, and 0  $\mu\text{m}$ , respectively. Comparing the measured pull-in voltages and the calculated ones in Table 6, it indicates that this method for extracting material properties is proper.



**Fig. 9** Fringe projection figure of the beam



**Fig. 10** Bending of the beam measured by micro-fringe projection system

**Table 7** Experimental results compared with data from literatures

	$\tilde{E}_{\text{poly}}$ (GPa)	$\tilde{\sigma}_{\text{poly}}$ (MPa)	$\tilde{E}_{Al}$ (GPa)		$\tilde{\sigma}_{Al}$ (MPa)
Experiments	145	31.3	41.8		13.2
References <sup>a</sup>	120–201	10–290	47.2–74.1		–

<sup>a</sup>Petersen 1978; Najafi and Suzuki 1989; Tilmans and Legtenberg 1994; Zou et al. 1995; Osterberg and Senturia 1997; Chan et al. 1999; Pamidighantam et al. 2002; O'Mahony et al. 2003; Nemirovsky and Bochobza-Degani 2001

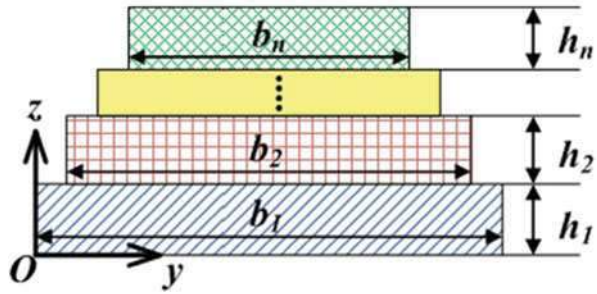
## Extraction by Resonance Frequency Measurements

### Model

This method uses multilayered micromachined double-clamped beams that are excited to resonance and are measured using a scanning laser Doppler vibrometer (LDV). The resonance frequency predicted by a composite beam model is used to extract Young's modulus and stresses of all the individual thin-film layers in the beam. To easily extract these material parameters, test structures for the doubly clamped beams with different widths are designed. Figure 11 shows the cross section of the beam with different widths. The length of the beam is  $l$ . The width, thickness, density, Young's modulus, residual stress, and Poisson's ratio of the  $i$ th layer film are  $b_i$ ,  $h_i$ ,  $\rho_i$ ,  $E_i$ ,  $\sigma_i$ , and  $\nu_i$ , respectively. Here a positive stress represents the tensile, and a minus stress represents the compressive.

In the designed test structures, since the width of the beam is bigger than five times thickness of the beam, the beam is considered to be wide, Young's modulus  $E_i$

**Fig. 11** Cross section of an n-layer beam with different widths



is replaced by the effective Young’s modulus  $\tilde{E}_i$ , and residual stress  $\sigma_i$  is replaced by the effective residual stress  $\tilde{\sigma}_i$ , i.e.,

$$\tilde{E}_i = \frac{E_i}{1 - \nu_i^2}, \quad \tilde{\sigma}_i = (1 - \nu_i)\sigma_i \tag{37}$$

In this part, effective Young’s modulus  $\tilde{E}_i$  is used instead of Young’s modulus  $E_i$ , the effective residual stress  $\tilde{\sigma}_i$  is used instead of the residual stress  $\sigma_i$ . If the height direction of the beam is assumed to be along the  $z$  axis, and the position of the beam bottom,  $z_0$ , is 0, while the position of the top of the  $i$ th layer is  $z_i$ , and the distance between the neutral axis and the bottom of the beam is  $z_c$ , it yields

$$z_i = \sum_{j=1}^i h_j, \quad (i = 1, 2, \dots, n) \tag{38}$$

According to the static equilibrium condition of the forces,  $z_c$  can be written as (Gere and Goodno 2009)

$$z_c = \frac{\sum_{i=1}^n \tilde{E}_i b_i (z_i^2 - z_{i-1}^2)}{2 \sum_{i=1}^n \tilde{E}_i b_i (z_i - z_{i-1})} \tag{39}$$

Then the moment of inertia of  $i$ th layer,  $I_i$ , with respect to the neutral axis of the beam is given by

$$I_i = \int_i (z - z_c)^2 dA_i = \frac{1}{3} b_i \left[ (z_i - z_c)^3 - (z_{i-1} - z_c)^3 \right], \quad (i = 1, 2, \dots, n) \tag{40}$$

where  $A_i$  is the cross-sectional area of the  $i$ th layer. For a multilayered beam, the bending stiffness,  $\overline{EI}$ , the axial stiffness,  $\overline{EA}$ , the linear density,  $\overline{\rho A}$ , and the axial load,  $\overline{\sigma A}$ , are defined by the sum of the individual terms for each layer in the beam as follows (Marshall et al. 2007):

$$\overline{EI} = \sum_{i=1}^n \tilde{E}_i I_i \quad (41)$$

$$\overline{EA} = \sum_{i=1}^n \tilde{E}_i A_i \quad (42)$$

$$\overline{\rho A} = \sum_{i=1}^n \rho_i A_i \quad (43)$$

$$\overline{\sigma A} = \sum_{i=1}^n \tilde{\sigma}_i A_i \quad (44)$$

As is well known (Nayfeh et al. 1995), in the beam theory, there is an  $n$ th critical buckling load  $P^{(n)}$  defined as the axial force sufficient to maintain the beam in a slightly bent form. When the axial load  $P$  is greater than the critical value  $P^{(n)}$ , the beam remains straight. And if  $P < P^{(n)}$  (i.e.,  $|P| > |P^{(n)}|$ ), the beam is initially statically deflected, and a post-buckling displacement  $b\varphi_n(\hat{x})$  is induced, where  $b$  is the dimensional scaling constant and  $\varphi_n(\hat{x})$  is the  $n$ th buckling mode shape. These two cases will be discussed below.

## Extraction by Multilayered Doubly Clamped Beams

### Initially Flat Beam

For an unbuckled beam, the axial load  $P = \overline{\sigma A}$  is greater than the critical value  $P^{(n)}$ , so the beam remains flat. For the undamped beam with a uniform cross section along its length, the natural frequencies  $\omega$  for the initially flat beam are given by the eigenvalues (Weaver et al. 1990):

$$|\mathbf{S}| = \begin{vmatrix} 0 & 1 & 0 & 1 \\ \lambda_1 & 0 & \lambda_2 & 0 \\ \sin(\lambda_1 l) & \cos(\lambda_1 l) & \sinh(\lambda_2 l) & \cosh(\lambda_2 l) \\ \lambda_1 \cos(\lambda_1 l) & -\lambda_1 \sin(\lambda_1 l) & \lambda_2 \cosh(\lambda_2 l) & \lambda_2 \sinh(\lambda_2 l) \end{vmatrix} = 0 \quad (45)$$

where

$$\beta_1 = \frac{\overline{\sigma A}}{\overline{EI}}, \beta_2 = \frac{\overline{\rho A}}{\overline{EI}} \omega^2 \quad (46)$$

$$\lambda_1 = \sqrt{\left(\sqrt{\beta_1^2 + 4\beta_2} - \beta_1\right)/2}, \lambda_2 = \sqrt{\left(\sqrt{\beta_1^2 + 4\beta_2} + \beta_1\right)/2}$$

here the matrix in Eq. (35) is considered to be  $\mathbf{S}$ .

### Buckled Beam

For the buckled beam,  $P < P^{(n)}$ , the beam will buckle, and a post-buckling displacement  $b\varphi_n(\hat{x})$  is induced. The natural frequencies  $\omega$  for the buckled beam are given by the eigenvalues (Nayfeh et al. 1995; Nicu and Bergaud 1999):

$$|\mathbf{B}| = \begin{vmatrix} 0 & 1 & 0 & 1 & 1 \\ \lambda_1 & 0 & \lambda_2 & 0 & 0 \\ \sin(\lambda_1 l) & \cos(\lambda_1 l) & \sinh(\lambda_2 l) & \cosh(\lambda_2 l) & 1 \\ \lambda_1 \cos(\lambda_1 l) & -\lambda_1 \sin(\lambda_1 l) & \lambda_2 \cosh(\lambda_2 l) & \lambda_2 \sinh(\lambda_2 l) & 0 \\ \frac{\alpha\lambda_1}{l^2}(\cos(\lambda_1 l) - 1) & \frac{-\alpha\lambda_1}{l^2} \sin(\lambda_1 l) & \frac{\alpha\lambda_2}{l^2}(1 - \cosh(\lambda_2 l)) & \frac{-\alpha\lambda_2}{l^2} \sinh(\lambda_2 l) & \beta_2 - \frac{\pi\alpha}{l} \\ \lambda_1^2 - (\frac{2\pi}{l})^2 & \lambda_1^2 - (\frac{2\pi}{l})^2 & \lambda_2^2 + (\frac{2\pi}{l})^2 & \lambda_2^2 + (\frac{2\pi}{l})^2 & 2\pi \end{vmatrix} = 0 \tag{47}$$

where

$$\alpha = \frac{8\pi^3 b^2 \overline{EA}}{l^3}, \beta_1 = \frac{P^{(1)}}{EI}, \beta_2 = \frac{\overline{\rho A}}{EI} \omega^2 \tag{48}$$

$$\lambda_1 = \sqrt{\frac{\sqrt{\beta_1^2 + 4\beta_2} - \beta_1}{2}}, \quad \lambda_2 = \sqrt{\frac{\sqrt{\beta_1^2 + 4\beta_2} + \beta_1}{2}}$$

Here the matrix in Eq. (47) is considered to be  $\mathbf{B}$ .

### Extraction of Material Properties

The matrices in Eqs. (45) and (47) are defined as  $\mathbf{S}$  and  $\mathbf{B}$ , respectively, Eqs. (45) and (47) can be rewritten as

$$F_S = F_S(\omega, \tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n, \tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_n) = \det(\mathbf{S}) = 0 \tag{49}$$

$$F_B = F_B(\omega, \tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n, \tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_n) = \det(\mathbf{B}) = 0 \tag{50}$$

where  $\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n, \tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_n$  represent the effective modulus and effective residual stresses of each layer, respectively, and  $\omega$  is the natural frequency. Since Eqs. (49) and (50) indicate that the natural frequencies are dependent of the material properties and geometries of the beam, Young's modulus and residual stress of each layer of the composite beam can be extracted by using natural frequencies of a set of beams with different geometries. For an  $n$ -layer composite beam, at least two kinds of beams with different lengths are needed, and each kind of beam with the same length needs at least  $n$  different types in widths. If the widths of each layer in the  $j$ th beam are assumed to be  $b_{j1}, b_{j2}, \dots, b_{jn} (1 \leq j \leq n)$ , and those in the  $k$ th beam with the same length as the  $j$ th beam are  $b_{k1}, b_{k2}, \dots, b_{kn} (1 \leq k \leq n, k \neq j)$ , it is required that the vector  $(b_{j1}, b_{j2}, \dots, b_{jn})$  and vector  $(b_{k1}, b_{k2}, \dots, b_{kn})$  must be linearly independent. It is well known that, for a single-layer beam when  $n$  equals to 1, it is not restricted by this principle. So at least  $2n$  kinds of beams are needed. The natural frequencies of

these beams are required, and deflection profiles of these beams need to be observed so as to determine Eq. (49) or Eq. (50) to be used. Once the natural frequencies of the designed test structures are measured, Young's modulus and residual stress of each layer for the composite beam can be extracted by numerical methods such as the Newton iteration method or least-square method. Generally, there are infinite solutions (Young's modulus  $E$  and residual stress  $\sigma$ ) to the above equation. However, for an engineering problem, for example, Young's modulus of the polysilicon has been reported in the range of 120–201 GPa. Therefore, only one solution will be in this range. This solution is what we want. For the bilayered beam, there are four unknowns (two Young's modulus and two residual stresses); four beams with different (length and width) dimensions are needed. If it is assumed that Young's modulus and residual stress are the same for each layer of the four beams because they are located close to each other, we have

$$\begin{cases} F_{11} = (\omega_{11}, \tilde{E}_1, \tilde{E}_2, \tilde{\sigma}_1, \tilde{\sigma}_2) = 0 \\ F_{21} = (\omega_{21}, \tilde{E}_1, \tilde{E}_2, \tilde{\sigma}_1, \tilde{\sigma}_2) = 0 \\ F_{12} = (\omega_{12}, \tilde{E}_1, \tilde{E}_2, \tilde{\sigma}_1, \tilde{\sigma}_2) = 0 \\ F_{22} = (\omega_{22}, \tilde{E}_1, \tilde{E}_2, \tilde{\sigma}_1, \tilde{\sigma}_2) = 0 \end{cases} \quad (51)$$

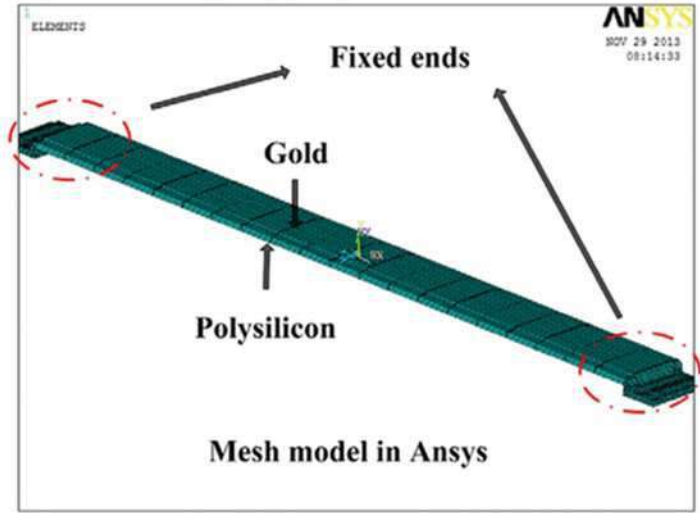
In the experiments, we measured and obtained four different fundamental frequencies ( $\omega_{11}$ ,  $\omega_{21}$ ,  $\omega_{12}$ ,  $\omega_{22}$ ) for the four beams by a scanning LDV. Then, we can extract the two Young's modulus and two residual stresses for the bilayered beam according to the solving flow. In the experiments, by using a digital holographic microscope (DHM, R2200), initially flat or buckled beams were checked so that Eq. (49) for the initially flat or Eq. (50) for the buckled beams could be selected.

## Simulation

To verify the proposed approach, analysis of doubly clamped bilayer beams as an example was simulated by FEM. Figure 12 shows a model for the doubly clamped bilayer beam, where the bottom layer is polysilicon and the top layer is gold. Four different beams in length and width are taken so that Young's modulus and residual stresses of both layers can be extracted. Input parameters of the beams for the simulation are listed in Table 8. To distinguish between the initially flat and buckled conditions, the simulations are carried out three times with two different pairs of residual stresses acting on the beams. Using the parameters in Table 8, the fundamental frequencies and deflection profiles of the beams are obtained in FEM simulations by ANSYS<sup>®</sup>. The results are then put back into Eq. (49) or Eq. (50) to figure out Young's modulus and residual stresses. Table 9 lists the comparison between material properties set to the FEM analysis and those calculated from FEM analysis. It is clear from Table 9 that the relative errors are less than 23%.

## Experiments and Discussions

Test structures were fabricated using MEMSCAP<sup>®</sup> PolyMUMPs process. Layer poly2 and layer metal were used as structural materials. Both bilayer and single-



**Fig.12** A doubly clamped bilayer beam in ANSYS®

**Table 8** Parameters of a bilayer beam

Parameters	Polysilicon	Gold
Young's modulus (GPa)	158	79
Residual stress (MPa)	20	20
	-20	-20
Poisson's ratio	0.22	0.44
Density (kg/m <sup>-3</sup> )	2230	19,300
Beam length (μm)	400	400
	300	300
Width (μm)	20	14
	20	10
Thickness (μm)	1.5	0.5

**Table 9** Comparison of the set and calculated values

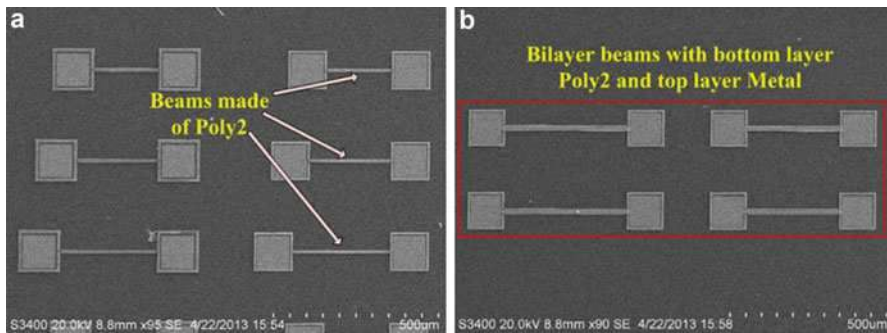
Material properties	Set values	Calculated results	Relative error (%)
Young's modulus (GPa) for polysilicon	158	149.2	5.57
Young's modulus (GPa) for gold	79	82.7	4.68
Residual stress (MPa) for polysilicon	20	24.5	22.5
Residual stress (MPa) for gold	20	17.9	10.5
Young's modulus (GPa) for polysilicon	158	132.8	15.9
Young's modulus (GPa) for gold	79	85.5	8.22
Residual stress (MPa) for polysilicon	-20	-16.9	15.5
Residual stress (MPa) for gold	-20	-22.3	11.5



**Table 10** Test results for the bilayer beams<sup>a</sup>

Parameter	Length $l$ ( $\mu\text{m}$ )	Poly2 width ( $\mu\text{m}$ )	Metal width ( $\mu\text{m}$ )	Initially status	Frequency (Hz)
Beam 1	200	10	0	Flat	110,250
Beam 2	250	10	0	Buckled	111,875
Beam 3	300	10	0	Buckled	137,688
Beam 4	400	20	14	Buckled	N/A
Beam 5	400	20	10	Buckled	39,688
Beam 6	300	20	14	Flat	64,531
Beam 7	300	20	10	Flat	52,969

<sup>a</sup>The thickness, density, and Poisson's ratio for poly2 are 1.5177  $\mu\text{m}$ , 2330 kg/m<sup>3</sup>, and 0.22, while the ones for metal are 0.5447  $\mu\text{m}$ , 19,300 kg/m<sup>3</sup>, and 0.44, respectively (MEMSCAP® 2016)

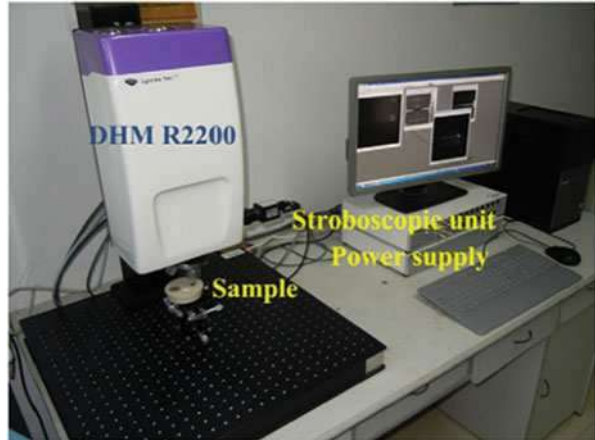


**Fig. 13** SEM photographs of the test structures. (a) Single-layer beams fabricated in layer poly2. (b) Bilayer beams made up of poly2 and metal

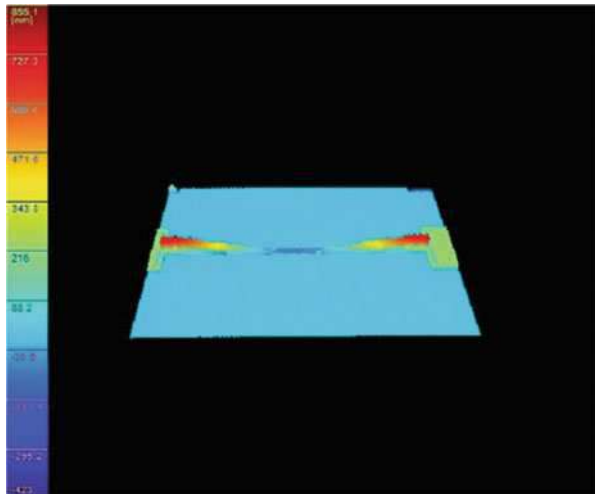
layer double-clamped beams were fabricated. For the bilayer beams, two lengths were applied for the beams with the same length but different widths of layer metal. And additional single-layer beams were also provided, which can be treated as the 0 width of layer metal. Parameters of the test structures are listed in Table 10. Scanning electron microscopy (SEM) photographs of the test structures are shown in Fig. 13. All the measurements were performed at room temperature.

To determine the deformation of the beams, a DHM (R2200) from Lyncée Tec was used. The setup of the DHM system is shown in Fig. 14. The DHM can produce an off-axis hologram through the interference between the wave reflected by the sample and a reference wave. Figure 15 shows a 3D image of one buckled beam. The profiles of the beams extracted from the DHM measurement are depicted as shown in Fig. 16. It can be observed clearly that some profiles of the beams are similar to a cosine curve, while the other profiles gently bend like bows. Due to the step-up support of the beams introduced by the surface micromachined fabrication process, the double-clamped beams applied with a slightly compressive load before buckling will have a small deformation rather than keep straight (Bouwstra and Geijselaers 1991). Hence, the beams with profiles like cosine curves are considered as buckled,

**Fig. 14** DHM system setup



**Fig. 15** A 3D image of one buckled beam



and others slightly bent are treated as initially flat beams. The observation results are summarized in Table 10.

After the deformation profiles were tested, resonant frequencies of the beams were then measured. A Polytec scanning LDV system with hardware modules of MSV-400-M2 was used. The setup of the LDV system is shown in Fig. 17. The test structures were excited by a piezoelectric plumbum zirconate titanate (PZT) ceramic slice with a periodical chip signal. By applying driven signal with frequencies ranging from 1 kHz to 1 MHz, the resonant frequencies were extracted from the frequency response functions (Song et al. 2009). Figure 18 shows the frequency response curve of a bilayer beam with length of 300  $\mu\text{m}$ . The fundamental frequencies of six beams were measured. The measured fundamental frequencies are listed in Table 10.

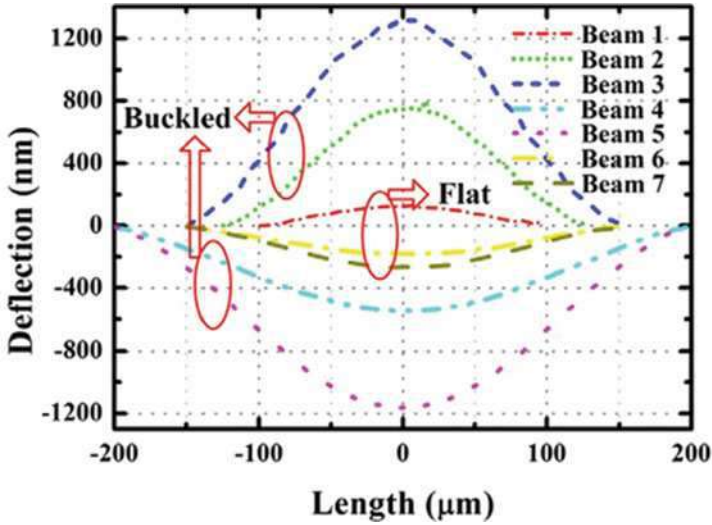


Fig. 16 Profiles of the beams extracted from the DHM measurement

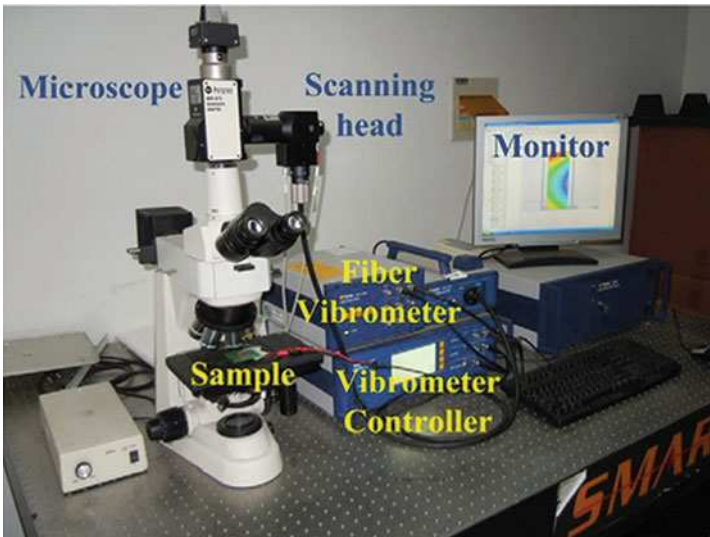
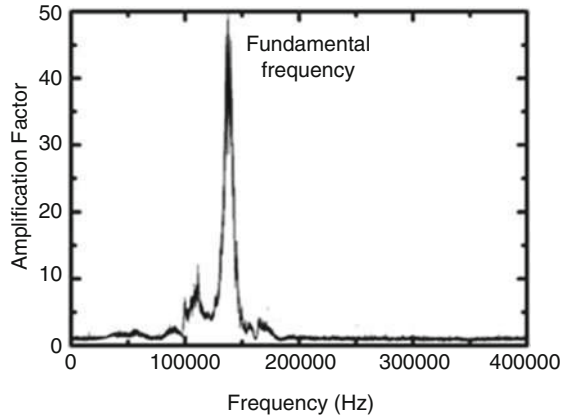


Fig. 17 LDV system setup

With the measured results, the effective modulus and residual stresses were extracted simultaneously by Eq. (49) or Eq. (50). There are 12 types of combinations in total, made up of four beams out of the whole six ones, which forms 12 groups of solvable equations from Eq. (49) or Eq. (50). Young's modulus and residual stresses of both layers are then extracted. The beam combinations and calculated results are listed in Table 11. The results show the proposed method produces results that agree

**Fig. 18** The frequency response curve of a bilayer beam with length of 300  $\mu\text{m}$



**Table 11** Extracted results from the measurements

Combinations of beams	Young's modulus of poly2 (GPa)	Young's modulus of metal (GPa)	Residual stress of poly2 (MPa)	Residual stress of metal (MPa)
1, 2, 5, 6	133.2	99.0	-13.3	12.2
1, 2, 5, 7	133.2	87.9	-13.3	14.6
1, 2, 6, 7	133.2	57.2	-13.3	27.7
1, 3, 5, 6	135.5	90.3	-13.7	15.9
1, 3, 5, 7	135.5	84.6	-13.7	17.2
1, 3, 6, 7	135.5	68.1	-13.7	24.2
1, 5, 6, 7	138.4	80.6	-14.3	20.4
2, 3, 5, 6	139.3	82.7	-14.1	18.1
2, 3, 5, 7	139.3	79.4	-14.1	18.9
2, 3, 6, 7	139.3	68.6	-14.1	23.6
2, 5, 6, 7	142.8	74.9	-14.5	21.4
3, 5, 6, 7	146.9	69.8	-14.7	22.4
Averages	137.7	78.6	-13.9	19.7
Standard deviation	4.2	11.5	0.5	4.4
References <sup>a</sup>	120~201	78	-18~-1	5~166

<sup>a</sup>Petersen 1978; Najafi and Suzuki 1989; Tilmans and Legtenberg 1994; Zou et al. 1995; Osterberg and Senturia 1997; Chan et al. 1999; Pamidighantam et al. 2002; O'Mahony et al. 2003; Nemirovsky and Bochobza-Degani 2001

with the values most often given in the literature. For example, Young's modulus of gold in this experiment coincides with 78 GPa as reported (WebElements™ 2016).

There are some factors that may affect the accuracy of the extraction model. First of all, squeezed damping in air may lead to the measured resonance frequency to shift from the expected value (Song et al. 2009). To reduce this deviation, the amplification factor for the frequency response functions at points of the resonant frequencies is required to be more than 20, and in this situation, the error caused by

the frequency shift will be less than 0.01%. Secondly, compliance of step-up support fabricated by the surface micromachined fabrication process may cause a relaxation of residual stresses. It is reported that under a compressive biaxial residual stress of 11 MPa, the relaxed stress is about 9 MPa as beam length is 1000  $\mu\text{m}$ , while the relaxed stress decreases to 3 MPa as beam length decreases to 500  $\mu\text{m}$  (Kobrinisky et al. 2000). In this study, lengths of double-clamped beams are less than 400  $\mu\text{m}$ . Hence, the relaxation of residual stress in the test structures is small and acceptable.

## Extraction by Multilayered Cantilevers

### The Theoretical Deduction for Material Parameters

If ignoring the longitudinal gradient residual stress, the single-layer cantilever will keep unbuckled after release. And the strain resulting from the residual stress tends to have less effect on the vibration. Therefore, it is considered that the resonance frequency has nothing to do with the residual stress when the single-layer cantilever keeps unbuckled after release (Osterberg and Senturia 1997). However, due to unmatched thermal stress or other reasons, in most cases, the multilayer cantilever will be buckled after release. And the deflection has effect on the resonance frequency of cantilever after release (Shen et al. 2001; Henrych 1981).

For unbuckled multilayer cantilevers after release, the approximate analytic formula of its first resonance frequency is shown in Eq. (52):

$$f_{1ucf} = \frac{1.875^2}{2\pi l^2} \sqrt{\frac{EI}{\rho A}} \quad (52)$$

Here subscript 1 represents the first resonance frequency,  $u$  represents that the cantilever is unbuckled, and  $cf$  represents that the boundary condition is cantilever.

When the cantilever is buckled, the influence of deflection to resonance frequency cannot be ignored. Assuming that  $R$  is the curvature radius of cantilever after release, a function describing its first resonance frequency is shown in Eq. (53):

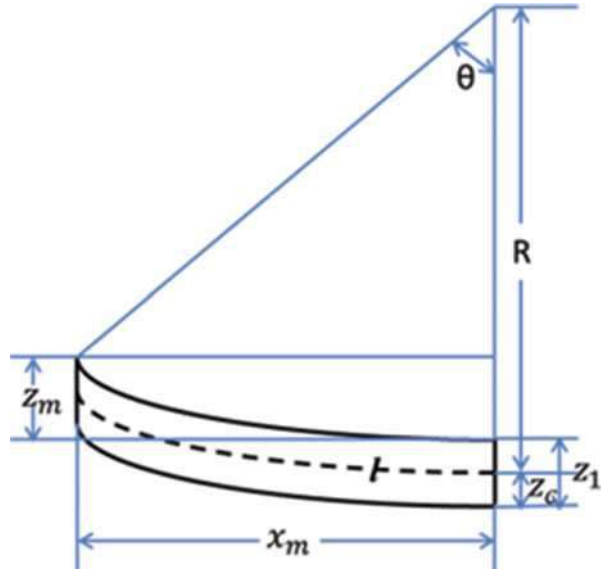
$$\begin{aligned} & \frac{1}{\sqrt{\chi_1^2 - 1}} \sin\left(\sqrt{\chi_1 + 1} \frac{l}{R}\right) \sinh\left(\sqrt{\chi_1 - 1} \frac{l}{R}\right) \\ & + \cos\left(\sqrt{\chi_1 + 1} \frac{l}{R}\right) \cosh\left(\sqrt{\chi_1 - 1} \frac{l}{R}\right) + 1 = 0 \end{aligned} \quad (53)$$

where

$$\chi_1^2 = \overline{\rho A R^4} (2\pi f_1)^2 / \overline{EI}. \quad (54)$$

If the curvature radius of the cantilever tends to infinity, take limit to Eq. (53), and the result is the same as Eq. (52), which means when the curvature radius of the cantilever tends to infinity, the vibrations of both the buckled and unbuckled

**Fig. 19** Schematic of a buckled multilayer beam



multilayer cantilever have little difference. Thus, in this situation, Eq. (52) can describe its resonance frequency approximatively.

**The Deduction of the Curvature Radius**

For a multilayer cantilever, if the compressive stress of the top layer is predominant, the cantilever will have an upward deflection. If the tensile stress of top layer is predominant, the cantilever will have a downward deflection. Both the upward and downward deflections have the same effect to modeling and results. However, if the downward deflection is too big to adhere to the substrate, results from these structures are invalid and discarded.

The chapter takes the upward deflection, for example, and assumes that the buckled multilayer cantilever has a uniform curvature radius, as shown in Fig. 19. Figure 19 shows the relation between the curvature radius and the maximum deflection of the multilayer cantilever, as expressed by Eq. (55):

$$\begin{cases} \theta = l/R \\ \cos(\theta) = [R - (z_1 - z_c) - z_m] / [R - (z_1 - z_c)] \end{cases} \quad (55)$$

In Fig. (19),  $R$  is the curvature radius of the multilayer cantilever,  $z_m$  is the maximum deflection of cantilever,  $l$  is the length of the unbuckled cantilever,  $x_m$  is the length of buckled cantilever,  $z_1$  is the thickness of the multilayer cantilever, and  $z_c$  is the height of neutral axis. By solving Eq. (55), values of curvature radius can be obtained accurately. Generally, the maximum deflection of the beam is much smaller than the length of the beam, so  $x_m$  can be approximated to  $l$ . Furthermore,  $z_1 - z_c$  is

also much smaller than the curvature radius of cantilever, so after simplification and ignoring  $z_j - z_c$ , a simplified equation, which is shown in Eq. (56), can be obtained:

$$R = l^2 / (2z_m). \quad (56)$$

### Extraction of Material Parameters

Under buckled or unbuckled situation of n-layer cantilevers, it only needs n kinds of cantilevers to get Young's modulus of each layer. Widths of the  $j$ th cantilever are  $w_{j1}, w_{j2}, \dots, w_{jn}$ , and widths of the  $k$ th cantilever are  $w_{k1}, w_{k2}, \dots, w_{kn}$ . If they have the same length, the vector  $(w_{j1}, w_{j2}, \dots, w_{jn})$  and the vector  $(w_{k1}, w_{k2}, \dots, w_{kn})$  should be linearly independent (Sun et al. 2014). DHM was used to distinguish the buckled from the unbuckled and determine to use Eq. (52) or Eq. (54). Changed form of Eq. (54) is shown in Eq. (57):

$$f_{1bcf} = \frac{1}{2\pi} \cdot \left( \frac{\chi_1^2 \cdot \overline{EI}}{\rho A \cdot R^4} \right)^{1/2} \quad (57)$$

Here  $b$  represents that the cantilever is buckled. Assuming that the first resonance frequencies measured is  $f_1, f_2, \dots, f_n$ , a set of equations is deduced from the theoretical formulas, as shown in Eq. (58):

$$\begin{cases} f_{1xcf,1}(\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n) - f_1 = 0 \\ f_{1xcf,2}(\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n) - f_2 = 0 \\ \dots \\ f_{1xcf,n}(\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_n) - f_n = 0 \end{cases}, \quad (58)$$

where

$$x = \begin{cases} b, & \text{when the cantilever is buckled.} \\ u, & \text{when the cantilever is unbuckled.} \end{cases} \quad (59)$$

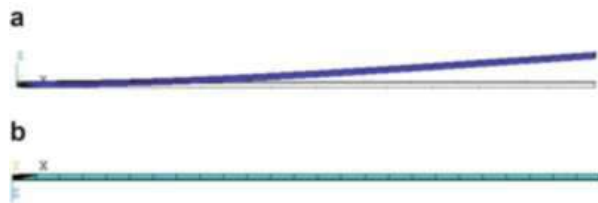
Effective Young's modulus  $\tilde{E}_i$  will be gotten from Eq. (58) rather than Young's modulus  $E_i$  because the Poisson's ratio  $\nu_i$  in each layer is unknown. To make results more intuitive, the chapter sets exact values to Poisson's ratio in each layer. The way to obtain Poisson's ratio  $\nu_i$  is shown in Sun et al. (2011), Namazu et al. (2013), and Li et al. (2009). Poisson's ratio of each layer in a certain rational range has little effect on results of Young's modulus. For example, the rational range of Poisson's ratio in poly layer is 0.2–0.25, and the chapter assumes 0.22 as Poisson's ratio in poly layer. It is obvious that results obtained from it will have little differences among results obtained from Poisson's ratio in poly layer which is 0.2 or other values in this rational range.

Solving Eq. (58) will get many sets of results, but only one of them is required. Thus it is necessary to select the only one rational result from them. For example, Young's modulus of polysilicon has been reported in the range of 120–201 GPa (Sun et al. 2014). Thus only one result in this range is correct.

**Table 12** Dimensions and parameters of cantilevers

	Poly layer	Metal layer
Young’s modulus $E_i$ (GPa)	–	57
Residual stress $\sigma_i$ (MPa)	10	–20
Poisson ratio $\nu_i$	0.22	0.35
Density $\rho_i$ (kg/m <sup>3</sup> )	2330	19,300
Length $l$ ( $\mu\text{m}$ )	200	200
Width $w_i$ ( $\mu\text{m}$ )	30	30
Thickness $h_i$ ( $\mu\text{m}$ )	1.5	0.5

**Fig. 20** Pictures of a poly-metal double-layer cantilever structure. (a) The cantilever is buckled. (b) The cantilever is unbuckled



**Finite Element Modeling**

The double-layer cantilever adopted in this chapter can simplify the situation. To verify the validity of the theoretical model, a set of double-layer cantilevers were analyzed by FEM method. Since ANSYS® cannot express the residual stress after release, heating the structure in advance simulates the buckled cantilever. From bottom to up, the first layer of cantilevers is designed to be polysilicon material, and the second layer is designed to be metal material: gold. Assuming that the parameters of the metal layer are fixed, only change Young’s modulus of the poly layer.

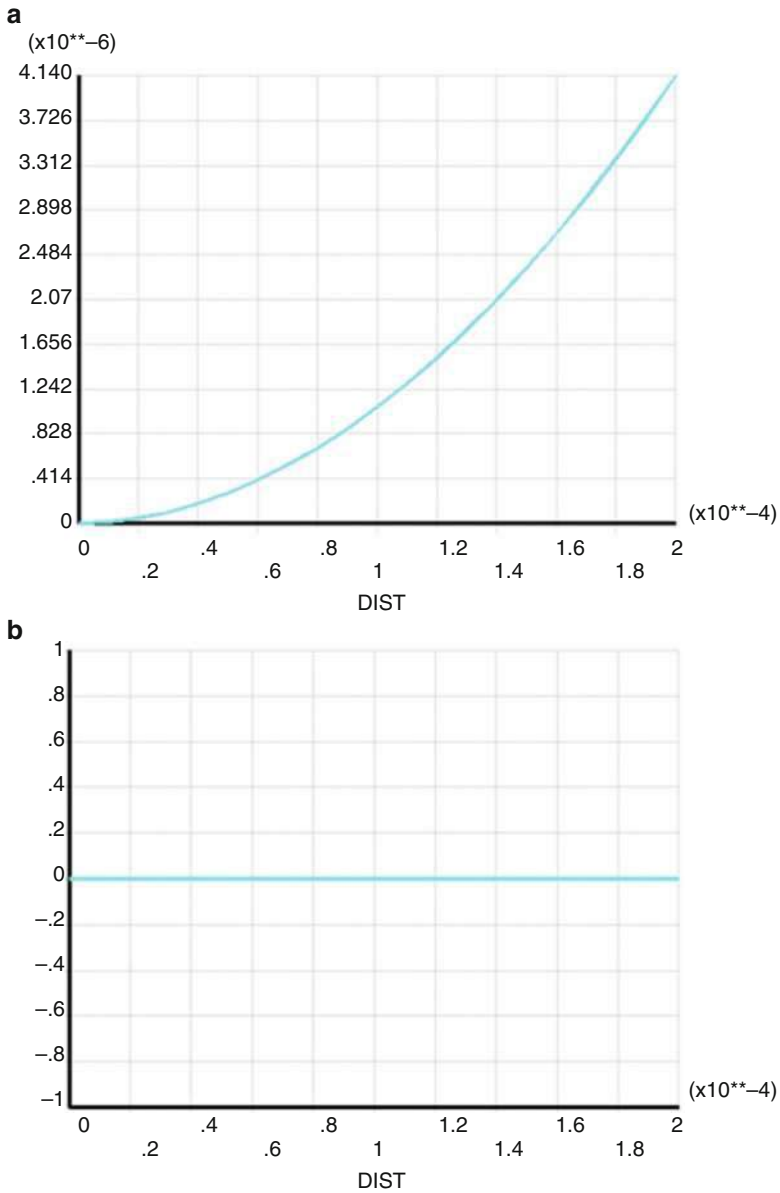
One hundred and twenty to 180 GPa is an available range of Young’s modulus of polysilicon. Thirteen values between it with the interval of 5 GPa are taken, and thirteen values of the first resonance frequency and the maximum deflection of cantilevers are gotten, respectively, for theoretical deduction. Dimensions and parameters of cantilevers (except Young’s modulus of poly layer) are shown in Table 12.

Taking Young’s modulus in poly layer as 120 GPa as an example, pictures of a poly-metal double-layer cantilever structure (shown in Fig. 20) and pictures of deflection values of the cantilever (shown in Fig. 21) can be obtained by using ANSYS® in both unbuckled and buckled situation.

**Simulation Results**

When the difference between the curvature radius of cantilever and the length of cantilever is not obvious, the cantilever has a large deflection. In this case, through its curvature radius and the first resonance frequency which is from ANSYS®, theoretical values of Young’s modulus in poly layer can be deduced accurately by solving Eqs. (53) and (54).





**Fig. 21** Pictures of deflection values of the cantilevers. (a) The cantilever is buckled. (b) The cantilever is unbuckled

When the cantilever has a small deflection, the deflection can be ignored approximately. In this case, the theoretical values of Young's modulus in poly layer can be obtained by using Eq. (52) through its first resonance frequency of cantilever.

**Table 13** Results when the cantilever is buckled

Setting values of Young's modulus in poly layer E (GPa)	The first resonance frequency from ANSYS® (Hz)	The maximum deflection of cantilever from ANSYS® (µm)	The curvature radius (µm)	Theoretical values of Young's modulus in poly layer E'(GPa)	Error (%) $\left  \frac{E-E'}{E} \right $
120	30,076	4.1403	4830.57	110.591	7.841
125	30,466	4.0334	4958.6	115.353	7.718
130	30,848	3.9237	5085.56	120.114	7.605
135	31,223	3.8376	5211.59	124.877	7.499
140	31,593	3.7476	5336.75	129.664	7.383
145	31,956	3.6622	5461.2	134.443	7.281
150	32,313	3.5810	5585.03	139.222	7.185
155	32,665	3.5038	5708.09	144.008	7.092
160	33,013	3.4301	5830.73	148.813	6.992
165	33,355	3.3598	5952.74	153.604	6.907
170	33,693	3.2926	6074.23	158.406	6.82
175	34,027	3.2283	6195.21	163.216	6.734
180	34,356	3.1667	6315.72	168.016	6.658

**Table 14** Results when the cantilever is unbuckled

Setting values of Young's modulus in poly layer E (GPa)	The first resonance frequency from ANSYS® (Hz)	Theoretical values of Young's modulus in poly layer E'' (GPa)	Error (%) $\left  \frac{E-E''}{E} \right $
120	30,074	110.65	7.792
125	30,464	115.413	7.670
130	30,846	120.173	7.560
135	31,221	124.937	7.454
140	31,590	129.711	7.349
145	31,954	134.503	7.239
150	32,311	139.282	7.145
155	32,663	144.069	7.052
160	33,011	148.874	6.954
165	33,353	153.665	6.870
170	33,691	158.468	6.784
175	34,025	163.278	6.698
180	34,354	168.078	6.623

In both buckled and unbuckled situation of cantilevers, the first resonance frequency obtained by ANSYS®, the maximum deflection of cantilevers, the curvature radius deduced, theoretical values of Young's modulus in poly layer, and the comparison between setting values of Young's modulus in poly layer and ones in theory model are shown in Tables 13 and 14.

From Tables 13 and 14, differences between setting values of Young's modulus in poly layer and values in theoretical model are both not obvious. As Young's modulus in poly layer increases, the curvature radius of cantilever increases. For relatively big curvature radius, errors from calculations have less effect on its veracity. Thus, final errors decrease as Young's modulus in poly layer increases. In addition, comparing the buckled situation with the unbuckled situation, little differences between theoretical values of Young's modulus in poly layer and errors in unbuckled situation are all less than ones in buckled situation. It means that the cantilever does not have much deflection and using Eq. (52), therefore, is more accurate. The reliability and feasibility of the theoretical model is demonstrated by little errors which are all less than 8 percent.

There are some reasons why setting values and theoretical values have differences. Firstly, conditions in theory model and the actual ones are not exactly the same. Secondly, when simulating, there might exist some uncontrollable factors which lead to the simulation environment that is different from the ideal environment. Thirdly, the deduction of the curvature radius of cantilever will bring in inevitable errors.

## Experiments and Discussions

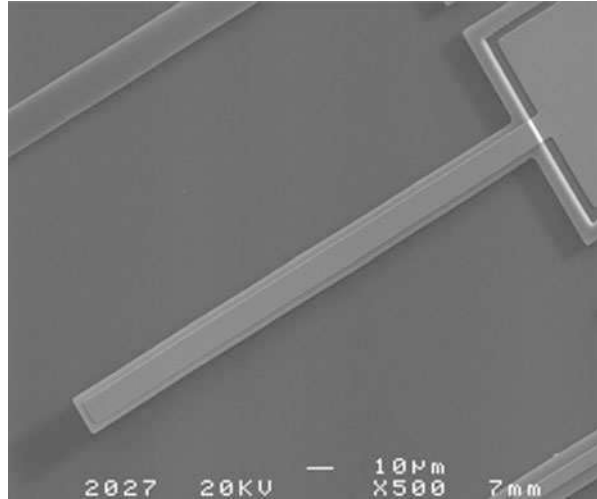
Test structures, which are double-layer cantilevers for simplification, were fabricated using MEMSCAP<sup>®</sup> PolyMUMPs process. Before measuring the first resonance frequency by LDV system with hardware modulus of MSV-400-M2, DHM (R2200) was used to determine whether cantilevers had deflection and to obtain values of curvature radius.

For the double-layer cantilever, there are two unknown values of Young's modulus to be obtained. Thus, two double-layer cantilevers with the same length but different widths should be a group to calculate. The double-layer cantilevers were designed in two different lengths. And with the same length, widths of two double-layer cantilevers should be linearly independent. Cantilever 4, as an example, was shown in Fig. 22 in scanning electron microscopy (SEM) photographs. The dimension, the deformation by DHM, and the first resonance frequency measurement of cantilevers were shown in Table 15. To avoid the situation that the bottom layer of cantilever adheres to the substrate, a test structure with suitable dimensions should be designed. Simulated by ANSYS<sup>®</sup>, cantilever 1, whose residual stress is  $-20$  MPa in poly layer and  $20$  MPa in metal layer, has a downward deflection of  $-1.4304$   $\mu\text{m}$ . Residual stress of  $20$  MPa is enough to have a big deflection. Thus, provided that the thickness of sacrifice layer is  $2$   $\mu\text{m}$ , cantilevers, whose dimension is little different from that of cantilever 1, may guarantee no possibility of substrate adhesion.

Assuming that the cantilever has a uniform curvature radius  $R$ , curvature radiuses are obtained by numerical fitting from the deflection curves. Curvature radiuses obtained are shown in Table 16.

Assuming that Poisson's ratio and density of the poly layer are  $0.22$  and  $2330$   $\text{kg/m}^3$ , respectively, and Poisson's ratio and density of metal are  $0.44$  and  $19,300$   $\text{kg/m}^3$ , respectively, Young's modulus of each layer can be calculated by the first resonance frequencies of test structures. Here is an example to use cantilever 1 and cantilever 2 to extract Young's modulus of each layer. After bringing the material parameters

**Fig. 22** An SEM picture of cantilever 4



**Table 15** Parameters of the cantilevers

Cantilever <i>i</i>	Length $l_i$ (µm)	Width of poly layer $w_{1i}$ (µm)	Width of metal layer $w_{2i}$ (µm)	Thickness of poly layer $h_1$ (µm)	Thickness of metal layer $h_2$ (µm)	Initial status	The first resonance frequency $f_i$ (kHz)
Cantilever 1	150	15	5	1.5	0.5	Buckled	74.38
Cantilever 2	150	15	9	1.5	0.5	Buckled	68.28
Cantilever 3	200	15	5	1.5	0.5	Buckled	41.72
Cantilever 4	200	15	9	1.5	0.5	Buckled	38.75

and the first resonance frequencies in Eqs. (53) and (54), a system of two element equations can be gotten, as shown in Eq. (60):

$$\begin{cases} f_{1bcf,1}(E_1, E_2) - f_1 = 0 \\ f_{1bcf,2}(E_1, E_2) - f_2 = 0 \end{cases} \quad (60)$$

Solve Eq. (60) and find the only one answer which meets the practical range of material parameters. The reasonable answer is that the values of Young’s modulus in poly layer and in metal layer are 156.774 GPa and 68.5414 GPa, respectively. Similarly, Young’s modulus of each layer can also be extracted by cantilever 3 and cantilever 4. The results are shown in Table 16. From Table 17, it is obvious that results agree with the practical range of material parameters and values reported in MEMSnet® (2016).

There are some reasons to affect the accuracy of the results. Firstly, in practical process, photoetching, mask, and self-aligned process can lead to the dimension deviation between designed structures and processed ones. Generally, in surface micromachining technology, dimension deviations in plane are no more than 0.5 µm, and dimension deviations in thickness are no more than 0.02 µm. It will have an

**Table 16** Results of numerical fitting for cantilevers

Structure	Curvature radius R ( $\mu\text{m}$ )	Standard deviation $\sigma_R$ ( $\mu\text{m}$ )	Abscissa of the anchor end $x_0$ ( $\mu\text{m}$ )	Standard deviation $\sigma_x$ ( $\mu\text{m}$ )	Ordinate of the anchor end $z_0$ ( $\mu\text{m}$ )	Standard deviation $\sigma_z$ ( $\mu\text{m}$ )
Cantilever 1	12997.32	24.60	2.98	$1.40 \times 10^{-1}$	$5.06 \times 10^{-3}$	$3.34 \times 10^{-4}$
Cantilever 2	9690.63	14.89	-1.37	$1.20 \times 10^{-1}$	$1.73 \times 10^{-3}$	$4.09 \times 10^{-4}$
Cantilever 3	13011.4	12.10	2.27	$0.93 \times 10^{-1}$	$3.31 \times 10^{-3}$	$3.03 \times 10^{-4}$
Cantilever 4	9587.3	7.33	-0.94	$0.79 \times 10^{-1}$	$1.86 \times 10^{-3}$	$3.61 \times 10^{-4}$

**Table 17** Results in different sets of cantilevers

Structure	Young's modulus in poly layer $E_1$ (GPa)	Young's modulus in, metal layer $E_2$ (GPa)
Cantilevers 1 and 2	156.7	68.66
Cantilevers 3 and 4	146.8	82.96
Average value	151.75	75.81
Reference (MEMSnet <sup>®</sup> 2016)	120~201	78

inevitable influence on minimize devices. Secondly, squeezed damping in air may lead to the inaccuracy of the first resonance frequency of cantilevers measured (Song et al. 2009; Kobrinsky et al. 2000). Thus, the amplification factor for the frequency response functions at points of the resonance frequencies is needed to more than 20 (Sun et al. 2014).

## Extraction by Combination of Multilayer Cantilevers and Double-Clamped Beams

Though only Young's modulus of multilayer beams can be obtained by Eq. (58), residual stress can also be obtained using a combination of n kinds of n-layer cantilevers and n kinds of n-layer double-clamped beams. First, using Eq. (58), Young's modulus of each layer can be calculated by the first resonance frequency. Then bringing Young's modulus obtained into Eq. (51) and using Eq. (51), the residual stress of each layer can be obtained. After obtaining Young's modulus of each layer, there is only one group of unknown parameters, residual stress of each layer, to be solved. Thus, the advantage of this method is simplified computation. However, if the radius of curvature of multilayer cantilevers is pretty small or the multilayer cantilevers have too huge downward deflection to touch the substrate, the method is invalid.

Test structures, which are double-layer cantilevers and double-clamped beams for simplification, were fabricated using MEMSCAP<sup>®</sup> PolyMUMPs process. The dimensions, the deflection, and the first resonance frequency of them are shown in Table 18.

According to the method above, Young's modulus  $E_1$ ,  $E_2$  and residual stress  $\sigma_1$   $\sigma_2$  of test structures can be obtained by combination of double-layer cantilevers and double-clamped beams, assuming that Poisson's ratio and density of the poly layer

**Table 18** Parameters of double-layer cantilevers and double-clamped beams

Structure <i>i</i>	Length $l_i$ ( $\mu\text{m}$ )	Width of poly layer $w_{1i}$ ( $\mu\text{m}$ )	Width of metal layer $w_{2i}$ ( $\mu\text{m}$ )	Thickness of poly layer $h_1$ ( $\mu\text{m}$ )	Thickness of metal layer $h_2$ ( $\mu\text{m}$ )	Initial status	The first resonance frequency $f_i$ (kHz)
Cantilever 11	150	15	5	1.5	0.5	Buckled	74.38
Cantilever 12	150	15	9	1.5	0.5	Buckled	68.28
Cantilever 13	200	15	5	1.5	0.5	Buckled	41.72
Cantilever 14	200	15	9	1.5	0.5	Buckled	38.75
Double-clamped beam 21	200	15	5	1.5	0.5	Unbuckled	224.69
Double-clamped beam 22	200	15	9	1.5	0.5	Unbuckled	216.72
Double-clamped beam 23	400	15	5	1.5	0.5	Buckled	22.97
Double-clamped beam 24	400	15	9	1.5	0.5	Unbuckled	27.97

**Table 19** Results in different combinations of double-layer cantilevers and double-clamped beams

Step 1	Step 2	$E_1$ (GPa)	$E_2$ (GPa)	$\sigma_1\sigma_1$ (MPa)	$\sigma\sigma_2$ (MPa)
Cantilever 11, Cantilever 12	Double-clamped beam 21	156.7	68.66	-17.36	30.77
	Double-clamped beam 22				
	Double-clamped beam 23			-15.44	25.62
	Double-clamped beam 24				
Cantilever 13, Cantilever 14	Double-clamped beam 21	146.8	82.96	-15.37	9.387
	Double-clamped beam 22				
	Double-clamped beam 23			-15.20	23.54
	Double-clamped beam 24				

are 0.22 and 2330 kg/m<sup>3</sup>, respectively, and Poisson’ ratio and density of metal are 0.44 and 19,300 kg/m<sup>3</sup>, respectively. The results are shown in Table 19. From Table 19, it is obvious that results agree with the practical range of material parameters and values reported in MEMSnet<sup>®</sup> (2016).

## Conclusion

In this chapter, methods are presented to measure Young’s modulus and residual stress of individual layer for MEMS composite films. Based on multilayer double-clamped beams and cantilevers, theoretical models for extracting material properties by resonant frequencies and pull-in voltage are analyzed. Material properties for the composite films can be obtained by a set of double-layer double-clamped beams and cantilevers with different lengths and widths. Both initially flat and buckled beams

have also been considered. The validity of proposed methods has been confirmed by the finite element analysis and experiments. Methods mentioned above are suitable for monitoring the MEMS fabrication process.

---

## References

- Bouwstra S, Geijselaers B (1991) On the resonance frequencies of microbridges. In: Proceedings of the 6th international conference on solid-state sensors and actuators (TRANSDUCERS'91), San Francisco, CA, IEEE, New York 2:1141–1144
- Chan EK, Garikipati K, Dutton RW (1999) Characterization of contact electromechanics through capacitance-voltage measurements and simulations. *J Microelectromech Syst* 8(2):208–217
- Chu LL, Que L, Gianchandani YB (2002) Measurements of material properties using differential capacitive strain sensors. *J Microelectromech Syst* 11(5):489–498
- De Coster J, Lofrano M, Jansen R, Rottenberg X, Severi S, Borremans J, Van der Plas G, Donnay S, Tilmans H, De Wolf I (2011) A novel test method for simultaneous measurement of thermal conductivity, CTE, residual stress and Young's modulus of suspended thin films using a laser doppler vibrometer. In: 2011 16th international solid-state sensors, actuators and microsystems conference. IEEE, pp 1701–1704
- Gabbay LD (1998) Computer aided macromodeling for MEMS (PhD Thesis), Massachusetts Institute of Technology, USA.
- Gere JM, Goodno BJ (2009) Mechanics of materials. Cengage learning Inc., Independence
- Gupta RK (1998) Electrostatic pull-in test structure design for in-situ mechanical property measurements of microelectromechanical systems (PhD Thesis), Massachusetts Institute of Technology, Cambridge, USA
- Henrych J (1981) The dynamics of arches and frames, vol 2. Elsevier Science Ltd, Amsterdam
- Hua R, Qing an H, Meng N, Weihua L (2004) A novel anchor for microbeam with perfect fixed-end boundary conditions. *Chinese J Semicond-Chinese Edition* 25(6):707–710
- Ilic B, Krylov S, Craighead H (2010) Young's modulus and density measurements of thin atomic layer deposited films using resonant nanomechanics. *J Appl Phys* 108(4):044317
- Kobrinisky MJ, Deutsch ER, Senturia SD (2000) Effect of support compliance and residual stress on the shape of doubly supported surface-micromachined beams. *J Microelectromech Syst* 9(3):361–369
- Li L, Gomes J, Brown G, Uttamchandani D, Pan W, Weiland D, Begbie M, Lowrie C, Desmulliez MP (2009) Simultaneous determination of the Young's modulus and Poisson's ratio in micro/nano materials. *J Micromech Microeng* 19(12):125027
- Li X, Ono T, Wang Y, Esashi M (2003) Ultrathin single-crystalline-silicon cantilever resonators: fabrication technology and significant specimen size effect on Young's modulus. *Appl Phys Lett* 83(15):3081–3083
- Marshall JC, Herman DL, Vernier PT, DeVoe DL, Gaitan M (2007) Young's modulus measurements in standard IC CMOS processes using MEMS test structures. *IEEE Electron Device Lett* 28(11):960–963
- MEMSCAP® (2016) PolyMUMPs reference material. [www.memscap.com/products/mumps/poly-mumps/reference-material](http://www.memscap.com/products/mumps/poly-mumps/reference-material). Accessed 07 Nov 2016
- MEMSnet® (2016) Material: polysilicon, film. [www.memsnets.org/material/polysiliconfilm](http://www.memsnets.org/material/polysiliconfilm). Accessed 07 Nov 2016
- Najafi K, Suzuki K (1989) A novel technique and structure for the measurement of intrinsic stress and Young's modulus of thin films. In: Micro electro mechanical systems, 1989, proceedings, an investigation of micro structures, sensors, actuators, machines and robots. IEEE, pp 96–97
- Namazu T, Fujii T, Takahashi M, Tanaka M, Inoue S (2013) A simple experimental technique for measuring the Poisson's ratio of microstructures. *J Microelectromech Syst* 22(3):625–636
- Nayfeh AH, Kreider W, Anderson T (1995) Investigation of natural frequencies and mode shapes of buckled beams. *AIAA J* 33(6):1121–1126

- Nemirovsky Y, Bochobza-Degani O (2001) A methodology and model for the pull-in parameters of electrostatic actuators. *J Microelectromech Syst* 10(4):601–615
- Nicu L, Bergaud C (1999) Experimental and theoretical investigations on nonlinear resonances of composite buckled microbridges. *J Appl Phys* 86(10):5835–5840
- Nie M, Huang Q-A, Li W, Rong H (2005) An in-situ technique for measuring Young's modulus and residual stress of each layer for multi-layer film. In: *The 13th international conference on solid-state sensors, actuators and microsystems, 2005. Digest of technical papers. TRANSDUCERS'05. IEEE*, pp 836–839
- O'Mahony C, Hill M, Duane R, Mathewson A (2003) Analysis of electromechanical boundary effects on the pull-in of micromachined fixed-fixed beams. *J Micromech Microeng* 13(4):S75
- Osterberg PM, Senturia SD (1997) M-TEST: a test chip for MEMS material property measurement using electrostatically actuated test structures. *J Microelectromech Syst* 6(2):107–118
- Pamidighantam S, Puers R, Baert K, Tilmans HA (2002) Pull-in voltage analysis of electrostatically actuated beam structures with fixed-fixed and fixed-free end conditions. *J Micromech Microeng* 12(4):458
- Petersen KE (1978) Dynamic micromechanics on silicon: techniques and devices. *IEEE Trans Electron Devices* 25(10):1241–1250
- Petersen KE, Guarnieri C (1979) Young's modulus measurements of thin films using micromechanics. *J Appl Phys* 50(11):6761–6766
- Rong H, Huang Q-A, Nie M, Li W (2004) An analytical model for pull-in voltage of clamped-clamped multilayer beams. *Sensors Actuators A Phys* 116(1):15–21
- Schweitz J-Å (1992) Mechanical characterization of thin films by micromechanical techniques. *MRS Bull* 17(07):34–45
- Shen F, Lu P, O'Shea S, Lee K, Ng T (2001) Thermal effects on coated resonant microcantilevers. *Sensors Actuators A Phys* 95(1):17–23
- Song J, Huang Q-A, Li M, Tang J-Y (2009) Effect of die-bonding process on MEMS device performance: system-level modeling and experimental verification. *J Microelectromech Syst* 18(2):274–286
- Sun C, Zhou Z-F, Li W-H, Huang Q-A (2014) A simple method for extracting material parameters of multilayered MEMS structures using resonance frequency measurements. *J Micromech Microeng* 24(7):075014
- Sun J-y, Hu J-l, Zheng Z-l, He X-t, H-h G (2011) A practical method for simultaneous determination of Poisson's ratio and Young's modulus of elasticity of thin films. *J Mech Sci Technol* 25(12):3165–3171
- Thomas RN, Guldberg J, Nathanson H, Malmberg P (1975) The mirror-matrix tube: a novel light valve for projection displays. *IEEE Trans Electron Devices* 22(9):765–775
- Tilmans HA, Elwenspoek M, Fluitman JH (1992) Micro resonant force gauges. *Sensors Actuators A Phys* 30(1–2):35–53
- Tilmans HA, Legtenberg R (1994) Electrostatically driven vacuum-encapsulated polysilicon resonators: part II. Theory and performance. *Sensors Actuators A Phys* 45(1):67–84
- Timoshenko SP, Woinowsky-Krieger S (1959) *Theory of plates and shells*. McGraw-hill, New York
- Van Drieënhuizen B, Goosen J, French P, Wolfenbuttel R (1993) Comparison of techniques for measuring both compressive and tensile stress in thin films. *Sensors Actuators A Phys* 37:756–765
- Wang S, Crary S, Najafi K (1992) Electronic determination of the modulus of elasticity and intrinsic stress of thin films using capacitive bridges. In: *MRS proceedings*. Cambridge University Press, p 203
- Weaver W Jr, Timoshenko SP, Young DH (1990) *Vibration problems in engineering*. John Wiley & Sons, New York
- WebElements™ (2016) .Gold: physical properties [www.webelements.com/gold/physics.html](http://www.webelements.com/gold/physics.html). Accessed 07 Nov 2016
- Zou Q, Li Z, Liu L (1995) New methods for measuring mechanical properties of thin films in micromachining: beam pull-in voltage (V PI) method and long beam deflection (LBD) method. *Sensors Actuators A Phys* 48(2):137–143





# Online Test Microstructures of the Thermophysical Properties of MEMS Conducting Films

Hai-Yun Liu, Zai-Fa Zhou, and Qing-An Huang

## Contents

Introduction .....	238
Online Test Microstructure of Thermal Conductivity .....	242
Fundamentals of Heat Transfer .....	242
Test Microstructure and Theoretical Model .....	244
Simulation and Verification .....	251
Experiment .....	253
Online Test Microstructure of Thermal Conductivity and Thermal Diffusivity .....	256
Transient-State Thermal Analysis of a Double-Clamped Beam .....	256
Thermal Resistors .....	259
Test Microstructure and Theoretical Model .....	260
Simulation and Verification .....	262
Experiment .....	264
Online Test Microstructure of the Coefficient of Thermal Expansion by Rotating Technique .....	266
Micro-rotating Structures .....	267
Test Microstructure and Theoretical Model .....	270
Simulation and Verification .....	274
Experiment .....	281
Online Test Microstructure of the Coefficient of Thermal Expansion by a Pull-In Approach .....	291
Pull-In Analysis of a Double-Clamped Beam .....	291
Test Microstructure and Theoretical Model .....	295
Simulation and Verification .....	296
Experiment .....	297
References .....	300

---

H.-Y. Liu (✉)

Department of Electronics and Information Engineering, Hohai University, Nanjing, Jiangsu, China  
e-mail: [haiyun\\_liu@hhu.edu.cn](mailto:haiyun_liu@hhu.edu.cn)

Z.-F. Zhou · Q.-A. Huang

Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, Jiangsu, China

e-mail: [zfzhou@seu.edu.cn](mailto:zfzhou@seu.edu.cn); [hqa@seu.edu.cn](mailto:hqa@seu.edu.cn)

---

**Abstract**

Thermophysical properties of MEMS materials, such as thermal conductivity, thermal diffusivity, and coefficient of thermal expansion (CTE), are one of the most important properties in MEMS technology. Steady-state thermal response and transient-state thermal response of MEMS devices depend on the thermal conductivity and the thermal diffusivity of device materials. Thermally driven microstructures, on the other hand, exploit the thermal expansion effect for their operation. It is necessary to characterize the thermophysical properties of MEMS materials for the design of MEMS devices.

This chapter will present online test microstructures and measurement methods for the thermophysical properties of MEMS conducting beams. The background of the work is reviewed in section “Introduction.” In section “Online Test Microstructure of Thermal Conductivity,” test microstructures for thermal conductivity based on steady-state thermal analysis are developed. Section “Online Test Microstructure of Thermal Conductivity and Thermal Diffusivity” is dedicated to discussing transient-state thermal analysis and proposing a test microstructure for both thermal conductivity and thermal diffusivity. In sections “Online Test Microstructure of the Coefficient of Thermal Expansion by Rotating Technique” and “Online Test Microstructure of the Coefficient of Thermal Expansion by a Pull-In Approach,” the coefficient of thermal expansion is extracted by micro-rotating structures and double-clamped beams, respectively. The former takes advantage of thermal actuation, while the latter makes use of the electrostatic pull-in approach. All the test microstructures proposed in sections. “Online Test Microstructure of Thermal Conductivity,” “Online Test Microstructure of Thermal Conductivity and Thermal Diffusivity” and “Online Test Microstructure of the Coefficient of Thermal Expansion by Rotating Technique” are stimulated electrically and measured electrically. They can find applications in MEMS fabrication process line to provide direct quality control and obtain the data needed by MEMS designers.

---

**Keywords**

Online test · Test structure · Thermal conductivity · Thermal diffusivity · The coefficient of thermal expansion · MEMS conducting thin films · Steady-state thermal analysis · Transient-state thermal analysis · Thermal actuation · Pull-in approach

---

**Introduction**

Thermophysical properties of MEMS materials typically include thermal conductivity, thermal diffusivity, coefficient of thermal expansion (CTE), etc. Thermal conductivity indicates the ability of a material to conduct heat. It is defined as the quantity of heat that passes in unit time through a unit area of a plate when its opposite faces differ in temperature by one kelvin. Thermal diffusivity measures the rate of transfer of heat of

a material from the hot side to the cold side. It is the central parameter for the time-dependent thermal response of MEMS microstructures (Brand and Fedder 2005). Thermal conductivity and thermal diffusivity are not only useful for analysis of thermal behavior of MEMS devices. The thermal transfer process depends on thermophysical properties as well as other certain physical variables. Many thermal sensors make use of this to sensing other physical variables, including distance, acceleration, flow speed, etc. (Liu 2006). The Coefficient of thermal expansion describes how the size of an object changes with a change in temperature. The performance of many thermally driven MEMS devices relies on thermal expansion effect of device materials (Ogando et al. 2012). However, although thermal expansion effect is useful in some MEMS devices, it can also be detrimental to others (Mag-isa et al. 2013). The mismatch of CTEs between attached films may introduce undesired residual stresses and deformations which lead to shift of performance or even failure of MEMS devices. To fully understand the thermal behavior of microstructures, optimize device performances, and improve the reliability of MEMS devices, thermophysical properties of MEMS materials should be characterized.

MEMS materials are mostly thin films that have small dimensions. Material properties of thin films including thermophysical properties are usually different from those of the bulk materials. Besides, the classical macroscopic characterization techniques are not suitable for micrometer samples, since the dimensions of thickness of thin films are too small to meet the sophisticated preparation requirements of traditional methods. Therefore, the characterization of materials for MEMS is a challenging task.

Steady-state methods have been used to measure the thermal conductivity of thin films. The measurements of the thermal conductivity of doped polysilicon thin films produced in complementary metal oxide semiconductor (CMOS) integrated circuit (IC) technology with one post-processing bulk silicon micromachining step have been developed (Arx et al. 2000; McConnell et al. 2001; Jansen and Obermeier 1996; Roncaglia et al. 2007). These test structures are generally composed of both a heater and a thermocouple (Wang et al. 2009) (or thermistor (Stojanovic et al. 2007)). The heater is heated by a dc current and provides a thermal profile for the sample film, and the thermocouple catches the temperature value at a given point. The thermal conductivity is then extracted according to an obtained equation of heat conduction. To reduce the error due to heat radiation and convection, the structure is heated to a temperature value at which heat radiation can be ignored, and it is kept at vacuum circumstances where convective heat losses through air could be neglected. A thermal van der Pauw test structure for measuring the thermal conductivity of thin films was proposed by Paul et al. (Hafizovic and Paul 2002; Paul et al. 2000). It consists of a Greek cross structure suspended on four arms above a micromachined cavity and makes use of a thermal analogy of the van der Pauw method. The measurement experiment was also conducted in vacuum circumstance. A polysilicon bridge used to measure the thermal conductivity of thin films was previously reported by Tai et al., where the bridge had a narrow lightly doped region at its center, while other regions were heavily doped (Tai et al. 1988). In this bridge, the determination of the thermal conductivity depends on measurements of

current–voltage ( $I$ – $V$ ) characteristics of the polysilicon bridges under two conditions: high vacuum and electrically insulating silicon oil.

Transient-state methods typically measure either the time response of the material of interest to a heating pulse or the thermal response to a sinusoidal heating current (Ziang and Grigoropoulos 1995). A double-end mounted beam structure with single- or multiplayer thin film membranes to measure the thermal conductivity and diffusivity was presented by Irace et al., which deploys polysilicon heating resistors and four polysilicon–aluminum thermocouples used as thermal sensors (Irace and Sarro 1999). The thermal conductivity is obtained by steady-state method, while the thermal diffusivity is extracted by applying a voltage pulse to the heater and analyzing the transient evolution of the temperature at different locations on the film. The measurements have to be performed under vacuum to minimize the influence of heat convection. Transient-state method, such as the  $3\omega$  method, can also be used to measure the thermal conductivity of thin films (Cahill 1990). A method for simultaneous determination of the thermal conductivity and the heat capacity of freestanding thin films based on  $3\omega$  method was introduced by Goodson et al. (Jain and Goodson 2008). A sinusoidal heating current at a frequency of  $\omega$  is applied on the metal heater patterned on top of the freestanding thin film. The resultant Joule heating and hence the temperature and the temperature-dependent electrical resistance of the heater fluctuate at  $2\omega$ , resulting in a  $3\omega$  voltage component that can be detected with a lock-in amplifier. Thermal conductivity and thermal diffusivity (or heat capacity) can be determined from the average heater temperature rise and the amplitude of the temperature oscillation at  $2\omega$ , both of which can be obtained indirectly by measuring the related voltage harmonics.

In addition to the methods mentioned above, a variety of measurement techniques based on transient-state analysis have been developed to determine thermal diffusivity (or heat capacity) of thin film materials. Von Arx et al. have reported a microstructure to test the heat capacity of the CMOS IC layer sandwiches. In this method, the structure was heated with a sinusoidal current with an angular frequency  $\omega$ . The Seebeck coefficient of gate polysilicon against the lower CMOS metal was used to monitor the propagation of heatwaves along the structure. The thermal parameters were obtained by analyzing the dissipating static or dynamic heating powers in the heating resistor and simultaneously monitoring the resulting temperature changes (Arx et al. 1997, 1998, 1995). Mastrangelo and Muller extracted the thermal diffusivity of low pressure chemical vapor deposition (LPCVD) polysilicon films using microbridge structures, and the measurements have also to be carried out in a high vacuum chamber (Mastrangelo and Muller 1988). Morikawa and Hashmoto utilized the means of Fourier transform thermal analysis to determine simultaneously the thermal diffusivity and the heat capacity as function of frequency and temperature. In the Fourier analysis, the harmonics were observed up to the 42nd order (Morikawa and Hashmoto 1998).

There are several available techniques used to measure the CTE of the thin films. Murarka et al. and Retajczyk et al. have deposited the thin film on two different substrates and optically measured the change in the curvature of wafer induced by the deposited thin film to determine the CTE of the thin film (Murarka and Retajczyk

1983; Retajczyk and Sinha 1980). In this method, the CTE of the thin film is calculated with the known Young's modulus, Poisson's ratio of the deposited thin film, and the CTE of the substrates. Fang et al. and Tada et al. have utilized the out-of-plane deflection of cantilevers caused by the temperature change to determine the CTE of the thin films (Fang et al. 1999; Tada et al. 2000; Cheng et al. 2015). The deflection is measured by using the optical interferometric technique. The CTE of the thin films can be obtained from an analytical expression that relates the angular deflection to the CTE. Pocratsky et al. have extracted the CTE with a single substrate using out-of-plane deflection measurements of freestanding fixed-fixed beams versus temperature (Pocratsky and Boer 2014). This simple approach determines the film CTE independent of other material properties and only depends on knowledge of the substrate CTE. Pan has reported a structure composed of a pair of cantilever beams with different lengths connected by a short tip beam to measure the CTE of the thin film (Pan 2002). As the temperature changes, deflection of the microstructure is induced by different expansion or contraction between the two cantilever beams, resulting in a large displacement of the tip beam. The amount of the displacement is measured with an optical microscope, and the CTE of the thin film can be calculated by using this piece of information. Chae et al. have adopted microgauge structure which is heated by applying a current into the structure and magnifies the tiny displacement caused by thermal expansion, to test the CTE of the thin film (Chae et al. 1999). The structure is placed in a vacuum chamber, and the displacement is measured by an optical microscope. Wang et al. have exploited tensile experiments at different temperatures to measure the CTE of the thin film (Wang et al. 2005). The measurements are performed by a special testing device. Coster et al. have proposed a method that uses a laser Doppler vibrometer (LDV) for extracting the CTE of the thin film in vacuum (Coster et al. 2011). With the knowledge of the CTE of the substrate and the measurement of the resonance frequencies with the LDV's laser pointed at the beam at different temperatures, the CTE of the thin film can be obtained.

Most extant extracting methods for the measurement of the thermophysical properties of thin films have to be carried out in vacuum environment, which render the test structures useless upon packaging and unqualified for use in routine wafer-level measurements. Many methods for determining thermal conductivity and thermal diffusivity require additional processing steps to fabricate the thermocouple (or thermistor), which may cause a change of the material properties. Techniques previously reported to determine the thin film CTE have usually made use of an optical or visual readout. Some methods use expensive instruments or special testing devices. Some methods require knowledge of other physical properties of the thin film specimen or the substrate. Although useful in a laboratory setting, the operations of such methods are inconvenient for high-volume manufacturing, which prevents these methods from meeting the requirement for online measurement. Constitutive properties of the thin films, including thermal conductivity, thermal diffusivity, and CTE, may vary significantly depending on the precise process sequence. Besides, the experimental values of MEMS material properties are relatively scarce and scattered, which cannot be used for MEMS

design and fabrication. Hence, it is important for MEMS designers to have access to test data from appropriate test microstructures that document the constitutive properties, and these data must come from the actual process line where the devices are to be manufactured. It is therefore hoped that test methods can be operated in atmosphere and can be applied at the wafer level using ordinary wafer-probe test equipment. An electrical readout better conforms to standard integrated circuit (IC) test equipment and procedures. Moreover, an electrical readout permits the test microstructure to be co-fabricated or co-packaged with other devices, potentially giving rise to many conceivable applications. Therefore, the test methods, which exploit electrical readout and can be carried out in free air, are much more preferred by foundries. Parameters extracted in such methods could be directly used in quality control.

Since the surface manufacturing process of MEMS is one of the most popular MEMS manufacturing process and polysilicon thin film is a major constituent in many MEMS devices, polysilicon thin film fabricated by surface manufacturing process is chosen as the investigation object in this chapter.

---

## **Online Test Microstructure of Thermal Conductivity**

The online test demands that the test microstructures undergo the same processing and steps as other devices fabricated and the extracting methods are based on electric stimulation and electric measurement. By using ordinary wafer-probe test equipment and operating the measurements in free air, the test microstructures can be applied in MEMS fabrication process line.

The purpose of this section is to propose online test microstructures for the thermal conductivity of surface micromachined polysilicon thin films by a novel steady-state approach. To fulfill online test requirements, the structures are heated electrically. Heat dissipation by convection, radiation, and heat transfer through the air gap and into the substrate is considered in an electrothermal model. The determination for the thermal conductivity of polysilicon thin films is derived from the steady-state thermal response of the test microstructures.

## **Fundamentals of Heat Transfer**

Heat transfer occurs whenever a temperature difference is present in a material. The fundamentals of heat transfer not only explain how heat energy may be transferred but also predict the rate at which the exchange will take place under certain specified conditions. There are three possible mechanisms for heat to move from one point to another: thermal conduction, thermal convection, and thermal radiation.

### **Thermal Conduction**

Thermal conduction is the transfer of heat by microscopic collisions of particles and movement of electrons within a body. From a macroscopic point of view, thermal

conduction denotes an energy transfer from the high-temperature region to the low-temperature region when a temperature gradient exists in a body. The relationship between the transferred energy due to thermal conduction and the temperature gradient can be described as (Holman 1997)

$$q = -kA \frac{dT}{dx} \quad (1)$$

where  $q$  is the heat transfer rate,  $dT/dx$  denotes the temperature gradient in the direction of the heat flow,  $k$  represents the thermal conductivity of the material, and  $A$  stands for the cross-sectional area. The minus sign indicates that the heat must flow downhill on the temperature scale, which is in accordance with the second law of thermodynamics.

### Thermal Convection

Thermal convection is the transfer of heat from one place to another by the movement of fluids. According to Newton's law of cooling, the heat transfer rate induced by thermal convection may be written as (Holman 1997)

$$q = \beta A(T_w - T_f) \quad (2)$$

where  $T_w$  and  $T_f$  are the temperature of the object's surface and the fluid, respectively, and  $\beta$  is the convection heat transfer coefficient.

There are two types of convective heat transfer: natural convection and forced convection. Opposed to forced convection, which is the transfer of heat to a body of moving fluid, natural convection is the transfer of heat from a surface into a stationary body of fluid. The convective heat transfer involved in this chapter is natural convection.

### Thermal Radiation

Thermal radiation is electromagnetic radiation generated by the thermal motion of charged particles in matter. It occurs through a vacuum or any transparent medium. Ideally, a thermal radiator emits energy at a rate proportional to the fourth power of the absolute temperature of the body. This relationship is called Stefan–Boltzmann law of thermal radiation, which is (Holman 1997)

$$q_{\text{emitted}} = \delta A T^4 \quad (3)$$

where  $\delta = 5.67 \times 10^{-8} \text{ W} \cdot \text{m}^2 \cdot \text{K}^{-4}$  is the Stefan–Boltzmann constant.

If there is a heat transfer surface at temperature  $T_1$  which is completely enclosed by a much larger surface at temperature  $T_2$ , the net radiant exchange in this case can be estimated as

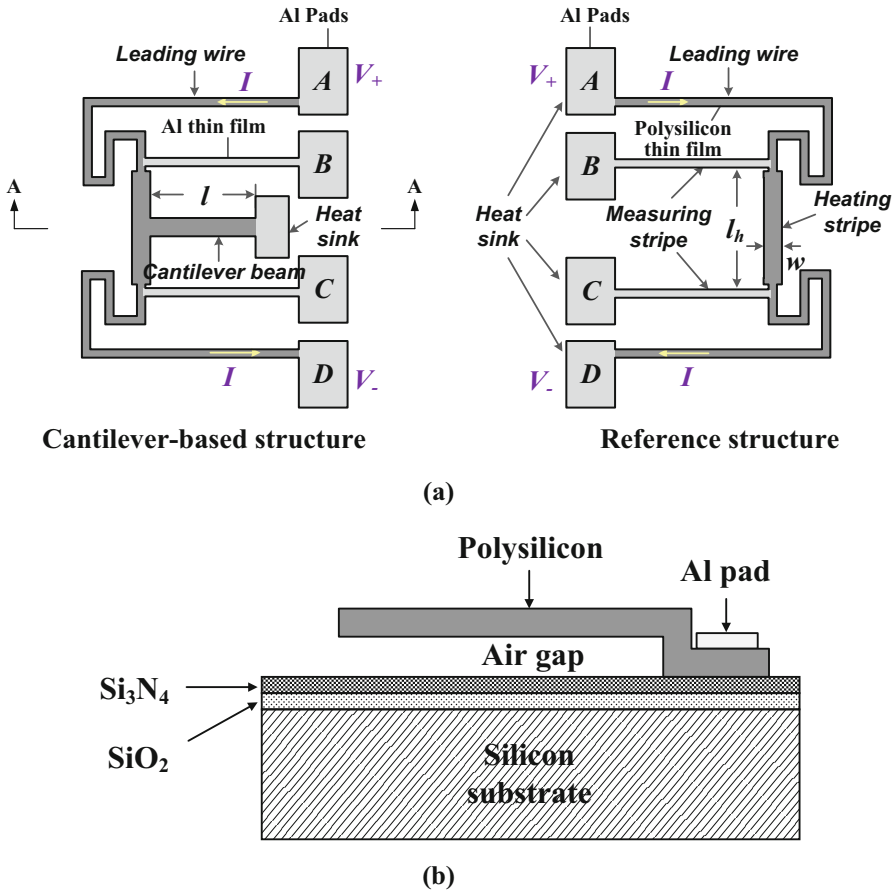
$$q = \varepsilon^* \delta A (T_1^4 - T_2^4) \quad (4)$$

where  $\varepsilon^*$  represents the radiation emissivity.

## Test Microstructure and Theoretical Model

### II-Shaped Test Microstructure

The  $\Pi$ -shaped test microstructure for thermal conductivity of surface micro-machined polysilicon thin films consists of a reference structure and a cantilever-based structure (Xu and Huang 2006; Xu et al. 2002), as shown in Fig. 1. Both reference structure and cantilever-based structure are made of polysilicon thin film and contain three suspended components: one heating stripe, two measuring stripes, and two leading wires of the heating stripe. The physical dimensions of these components in the two structures are identical. Metal thin films are deposited on the measuring stripes that will greatly increase the measurement accuracy of the voltage potential across the central heating stripes. The heating stripe is designed to be wide so that the temperature across the stripe is uniform. The difference between



**Fig. 1** Schematic representation of the  $\Pi$ -shaped test microstructure. (a) Top view. (b) Cross section view



the reference structure and the cantilever-based structure is that there is a polysilicon cantilever beam connected between the heating stripe and an anchor in the cantilever-based structure. During the measurement, a voltage is applied between A and D, while the current passing from A to D is tested and the voltage drop across the heating stripe is measured through B and C.

To extract the thermal conduction properties of polysilicon thin films, the electrothermal model is developed here for the test microstructures. The bulk silicon substrate is assumed to be a heat sink with a constant temperature  $T_0$ , i.e., environment temperature (see Fig. 1). And all the anchors (A, B, C, and D in Fig. 1) are assumed to be  $T_0$  because the area of the anchors is so big and they will be connected to the metal probe during the test. The heat transfer of the polysilicon heating stripe is illustrated in Fig. 2. According to the Stefan–Boltzmann law, the effective radiation losses can be linearized if the temperature is not high (Volkein and Baltes 1992), and it shows that the radiation effects can even be ignored since routine operation of the beam is with 300 to 800 K (Geisberger et al. 2003). Since the size in the length direction of the heating stripe is much larger than that of its width direction and thickness direction, the electrothermal analysis of the heating stripe is generally simplified for one dimension (Lin and Chiao 1996). Considering the effects of heat exchange factors such as thermal conduction, thermal convection, thermal radiation, and the heat loss through the air gap into the substrate, a one-dimensional heat balance equation for the element of length  $dx$  in polysilicon heating stripe is shown as follows (Lin and Chiao 1996; Huang and Lee 1999):

$$-k_p wh \left[ \frac{dT}{dx} \right]_x + J^2 \rho wh dx = -k_p wh \left[ \frac{dT}{dx} \right]_{x+dx} + \beta(2h+w) dx [T(x) - T_0] + 4\epsilon^* \delta(2h+w) dx T_0^3 [T(x) - T_0] + \frac{S}{R_T} w dx [T(x) - T_0] \quad (5)$$

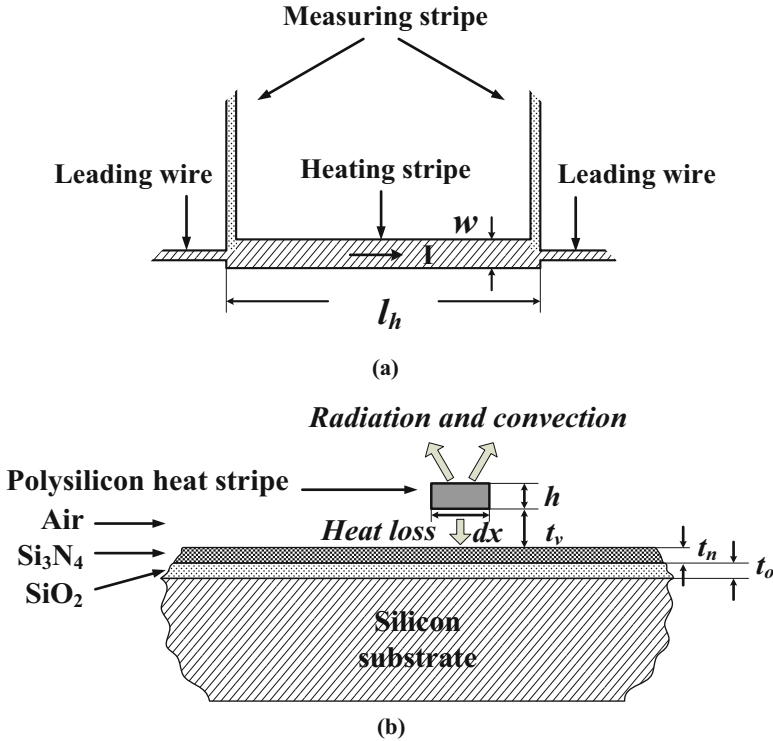
with

$$S = \frac{h}{w} \left( \frac{2t_v}{h} + 1 \right) + 1 \quad (6)$$

$$R_T = \frac{t_v}{k_v} + \frac{t_n}{k_n} + \frac{t_o}{k_o} \quad (7)$$

where  $w$  and  $h$  are the width and the thickness of the polysilicon heating stripe, respectively;  $T$  is the temperature in the heating stripe;  $k_p$  and  $\rho$  are the thermal conductivity and the electrical resistivity of the polysilicon thin film, respectively;  $t$  represents time;  $J$  denotes the current density;  $S$  is the shape factor; and  $R_T$  stands for the equivalence thermal resistance between the polysilicon heating stripe and the substrate.  $t_v$ ,  $t_n$ , and  $t_o$  and  $k_v$ ,  $k_n$ , and  $k_o$  represent the thickness and the thermal conductivity of the air gap, the  $\text{Si}_3\text{N}_4$  layer, and the  $\text{SiO}_2$  layer, respectively.

On the left-hand side of Eq. 5, the first term is the energy conducted into the left face of the element, and the second term represents the heat generated within the



**Fig. 2** Scheme of the heat transfer path in the polysilicon heating stripe. (a) Joule heat generated by electric current. (b) Thermal convection, thermal radiation, and the heat loss through the air gap into the substrate

element. On the right-hand side, the first term is the energy conducted out of the right face of the element, the second term and the third term describe the thermal convection and the thermal radiation from the top and side surface of the element into the surrounding environment, respectively, and the fourth term denotes the heat loss through the air gap into the substrate.

Define a parameter  $G_A$  to represent the heat transfer coefficient of the polysilicon thin film per unit length due to convection, radiation, and heat loss through the air gap into the substrate; then, it yields

$$G_A = (2h + w)\beta + 4(2h + w)\varepsilon^* \delta T_0^3 + \frac{S}{R_T} w \tag{8}$$

The heat balance condition of the reference structure is

$$P_R = G_R \Delta T_R \tag{9}$$

where  $P_R$  and  $G_R$  are the heating power and the total heat transfer coefficient of the reference structure, respectively.  $\Delta T_R = T_R - T_0$ , where  $T_R$  is the temperature of the heating stripe for the reference structure.

No thermal current passes the heating stripe since there are no temperature gradients along the heating stripe. Therefore, the heat balance for the heating strip can be written as

$$\frac{J^2 \rho w h}{\Delta T_R} = (2h + w)\beta + 4(2h + w)\varepsilon^* \delta T_0^3 + \frac{S}{R_T} w \quad (10)$$

That is,

$$G_A = \frac{V_{Rh} I_R}{\Delta T_R l_h} = \frac{P_{Rh}}{\Delta T_R l_h} \quad (11)$$

where  $l_h$  is the length of the heating stripe and  $V_{Rh}$  and  $I_R$  are the measured voltage and current through the heating stripe, respectively. Thus  $P_{Rh} = V_{Rh} \cdot I_R$  is the heating power for the heating stripe.

The heat balance condition of the cantilever-based structure is given by

$$P_c = P'_R + P_B = G_R \Delta T_c + G_B \Delta T_c \quad (12)$$

where  $P_c$  and  $P_B$  are the total heating power for the cantilever-based structure and the heat dissipation power of the polysilicon cantilever beam, respectively;  $P'_R$  denotes the heating power for the heating stripe, the leading wires of the heating stripe, and the measuring stripes of the cantilever-based structure; and  $G_B$  stands for the heat transfer coefficient of the polysilicon cantilever beam.  $\Delta T_c = T_c - T_0$ , where  $T_c$  is the temperature of the heating stripe for the cantilever-based structure.

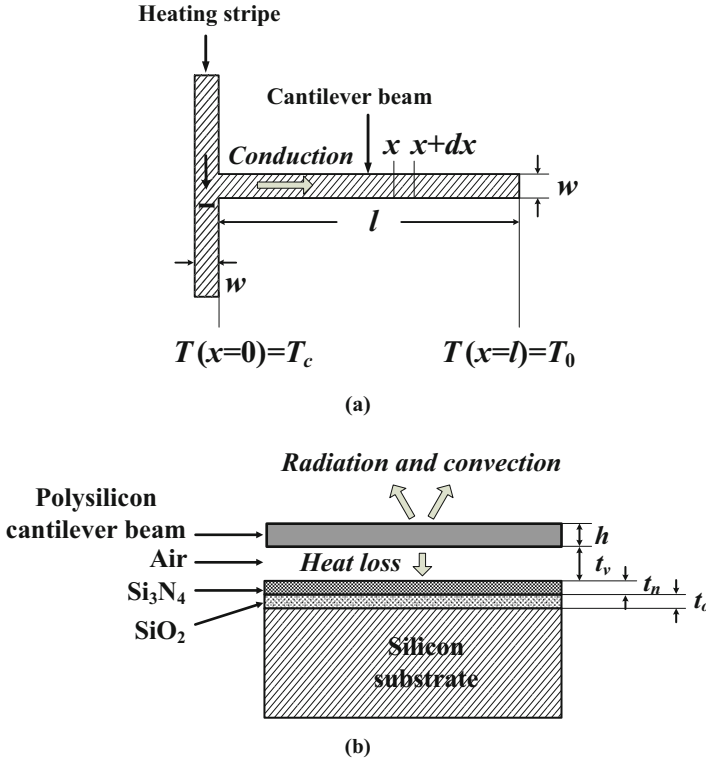
Equation 5 is also suitable for the cantilever beam. The heat transfer for the cantilever beam is illustrated in Fig. 3. Since there is no electric heating current passing through the polysilicon cantilever beam of the cantilever-based structure, the heat balance equation for the polysilicon cantilever beam becomes

$$\frac{d^2 T(x)}{dx^2} - \frac{G_A}{k_p w h} (T(x) - T_0) = 0 \quad (13)$$

where  $T(x)$  is the temperature distribution of the polysilicon cantilever beam. The boundary condition for Eq. 13 is

$$\begin{cases} T(0) = T_c \\ T(l) = T_0 \end{cases} \quad (14)$$

where  $l$  is the length of the polysilicon cantilever beam. By solving Eq. 13, the temperature distribution of the polysilicon cantilever beam can be calculated as



**Fig. 3** Scheme of the heat transfer path in the polysilicon cantilever beam. **(a)** Thermal conduction. **(b)** Thermal convection, thermal radiation, and the heat loss through the air gap into the substrate

$$T(x) = T_0 - \frac{(T_c - T_0) \cdot e^{-l\sqrt{\frac{G_A}{k_pwh}}}}{e^{l\sqrt{\frac{G_A}{k_pwh}}} - e^{-l\sqrt{\frac{G_A}{k_pwh}}}} \cdot e^{x\sqrt{\frac{G_A}{k_pwh}}} + \frac{(T_c - T_0) \cdot e^{l\sqrt{\frac{G_A}{k_pwh}}}}{e^{l\sqrt{\frac{G_A}{k_pwh}}} - e^{-l\sqrt{\frac{G_A}{k_pwh}}}} \cdot e^{-x\sqrt{\frac{G_A}{k_pwh}}} \quad (15)$$

According to Fourier's Law, the heat dissipation power ( $P_B$ ) of the polysilicon cantilever beam can be described as

$$P_B = -k_pwh \left. \frac{dT}{dx} \right|_{x=0} \quad (16)$$

Substitute Eqs. 9, 11, 12, and 15 into Eq. 16; the thermal conductivity of the polysilicon cantilever beam is derived as

$$k_p = \left[ \left( \frac{P_c}{\Delta T_c} - \frac{P_R}{\Delta T_R} \right) - \frac{P_{Rh}}{2\Delta T_R} \cdot \frac{l_B}{l_h} \right] \times \frac{l_B}{wh} \quad (17)$$

It is clear from Eq. 17 that only the thermal conductivity of the polysilicon thin film appears, whereas the natural convection coefficient, the effective emissivity, the Stefan–Boltzmann constant, the shape factor of the beam, and the equivalence thermal resistance of the air layer,  $\text{Si}_3\text{N}_4$  layer, and  $\text{SiO}_2$  layer between the beam and substrate have all been canceled out.

### T-Shaped Test Microstructure

The T-shaped test microstructure for thermal conductivity of surface micro-machined polysilicon thin films consists of a reference structure and two cantilever-based structures, as displayed in Fig. 4. The shape of the reference structure and the cantilever-based structure in T-shaped test microstructure is identical to that of the  $\Pi$ -shaped test microstructure. The cantilever-based structure contained an additional cantilever beam comparing to the reference structure, and the lengths of the cantilever beams in the two cantilever-based structures are different. Apart from that, all three structures, i.e., the reference structure and the two cantilever-based structures, have the same dimension. The heat balance conditions of the reference structure and the two cantilever-based structures (as shown in Fig. 5) can be expressed, respectively, as follows:

$$P_R = G_R \Delta T_R \quad (18)$$

$$P_{c1} = G_R \Delta T_{c1} + G_{B1} \Delta T_{c1} \quad (19)$$

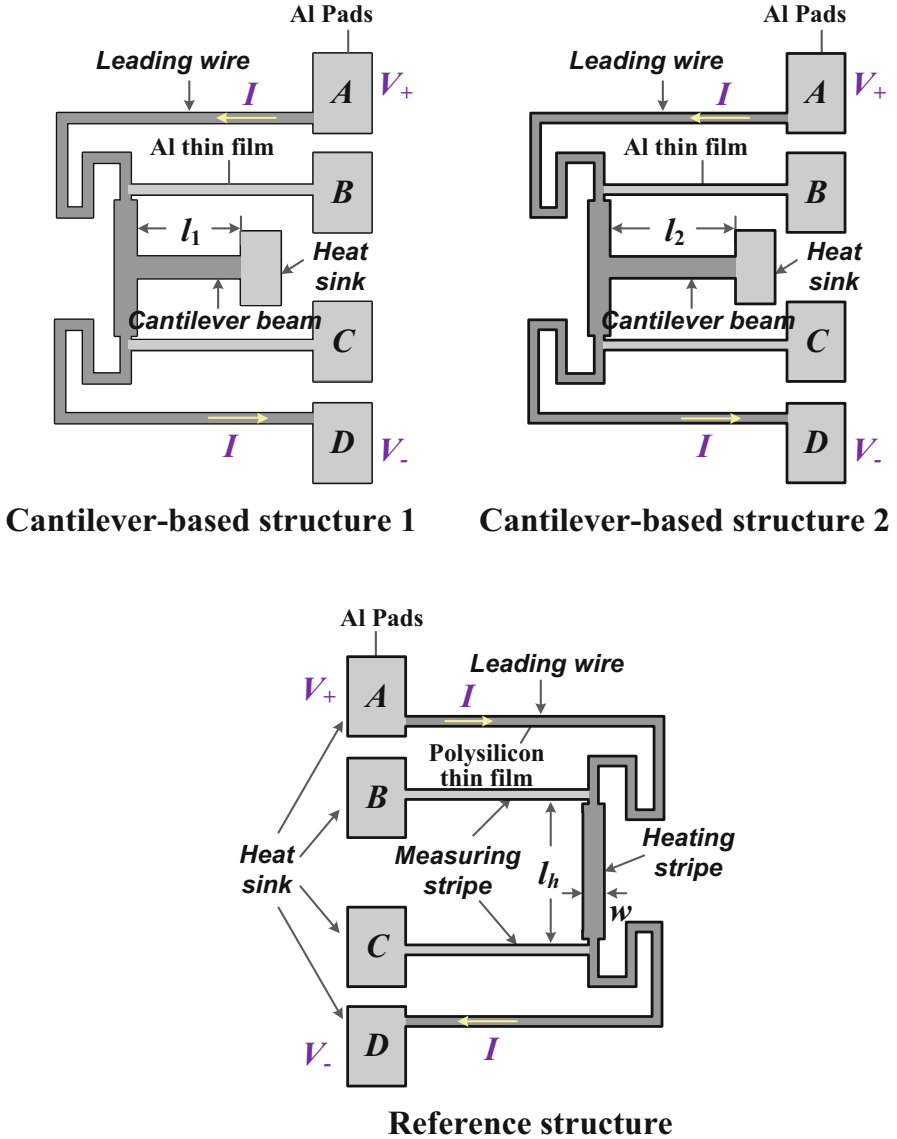
$$P_{c2} = G_R \Delta T_{c2} + G_{B2} \Delta T_{c2} \quad (20)$$

where  $P_{c1}$  and  $P_{c2}$  are the total heating power for the cantilever-based structures, respectively;  $\Delta T_{c1} = T_{c1} - T_0$  and  $\Delta T_{c2} = T_{c2} - T_0$ , where  $T_{c1}$  and  $T_{c2}$  represent the temperature of the heating stripes for the cantilever-based structures, respectively;  $G_{B1}$  and  $G_{B2}$  denote the heat transfer coefficients of the polysilicon cantilever beams for the cantilever-based structures, respectively; and  $P_R$  and  $G_R$  have the same definition as before.

Substitute Eq. 5 into Eq. 16, and combine Eqs. 18, 19, and 20; heat transfer coefficients of the polysilicon cantilever beams for the cantilever-based structures can be estimated, respectively, as

$$\begin{aligned} G_{B1} &= \left( \frac{P_{c1}}{\Delta T_{c1}} - \frac{P_R}{\Delta T_R} \right) \\ &= k_p h \frac{w}{l_1} + \beta(2h + w)l_1 + 4(2h + w)\varepsilon^* \delta l_1 T_0^3 + \frac{S}{R_T} w l_1 \end{aligned} \quad (21)$$

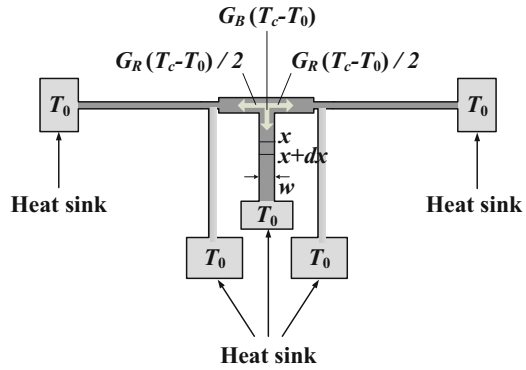
$$\begin{aligned} G_{B2} &= \left( \frac{P_{c2}}{\Delta T_{c2}} - \frac{P_R}{\Delta T_R} \right) \\ &= k_p h \frac{w}{l_2} + \beta(2h + w)l_2 + 4(2h + w)\varepsilon^* \delta l_2 T_0^3 + \frac{S}{R_T} w l_2 \end{aligned} \quad (22)$$



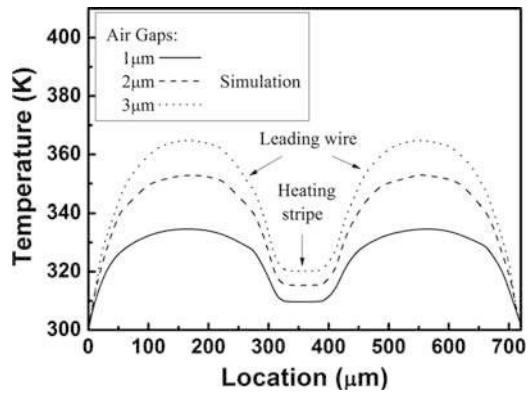
**Fig. 4** Schematic representation of the T-shaped test microstructure

where  $l_1$  and  $l_2$  are the lengths of the cantilever beams for the cantilever-based structures, respectively. Combining Eqs. 21 and 22, the thermal conductivity of the polysilicon thin film is then found by

**Fig. 5** Schematic representation of the heat balance for the cantilever-based structure



**Fig. 6** Temperature profile for the leading wire and heating stripe (from A to D). The air gap is 1, 2, and 3 μm from bottom in the figure, respectively



$$k_p = \left[ \left( \frac{P_{c1}}{\Delta T_{c1}} - \frac{P_r}{\Delta T_r} \right) \frac{1}{l_1} - \left( \frac{P_{c2}}{\Delta T_{c2}} - \frac{P_R}{\Delta T_R} \right) \frac{1}{l_2} \right] \cdot \frac{1}{wh} \cdot \left( \frac{1}{l_1^2} - \frac{1}{l_2^2} \right)^{-1} \quad (23)$$

**Simulation and Verification**

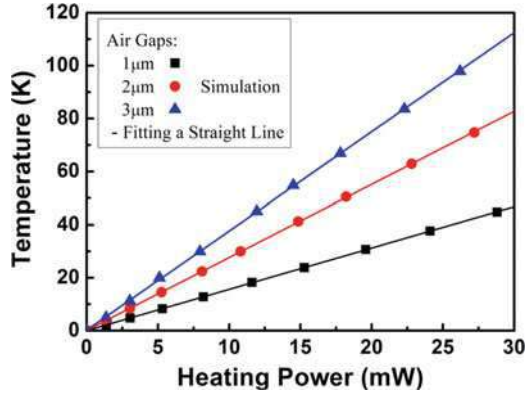
To confirm the electrothermal analysis, numerical simulations were made by ANSYS. The temperature profile of the heating stripe and the leading wire (heating polysilicon from A to D) for the reference structure is demonstrated in Fig. 6. In the simulation, an applied voltage between A and D is 5 V. The geometric dimensions are given in Table 1. The material parameters used for analysis is from Huang and Lee (1999).

It is clear from Fig. 6 that the temperature is uniform along the heating stripe because the measuring stripe with aluminum films is connected to two ends of the heating stripe, and the heating stripe is wide.

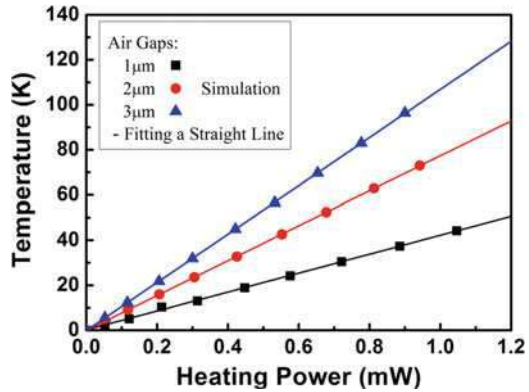
**Table 1** Geometric dimensions for test microstructure (Xu and Huang 2006)

Dimension	Length ( $\mu\text{m}$ )	Width ( $\mu\text{m}$ )	Thickness ( $\mu\text{m}$ )
Heating stripe	80	15	2
Measuring stripe	200	10	2
Cantilever beam 1	150	10	2
Cantilever beam 2	200	10	2
Leading wire	320 ( $\times 2$ )	5	2

**Fig. 7** Simulation results: relationship between the heating power ( $P_R$ ) and the temperature of the heating stripe ( $T_R$ ) for the reference structure



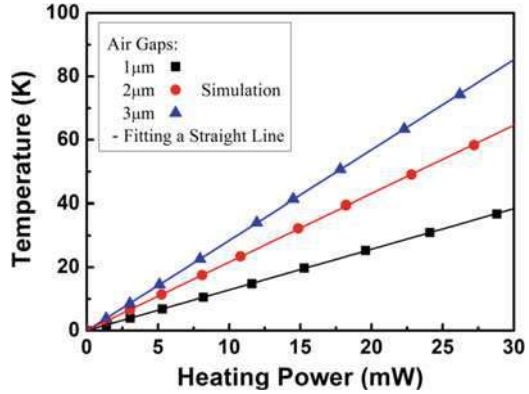
**Fig. 8** Simulation results: relationship between the thermal power consumption ( $P_{Rh}$ ) and the temperature of the heating stripe ( $T_R$ ) for the reference structure



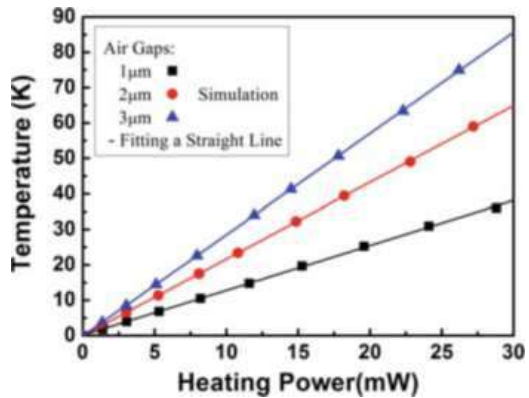
Figures 7, 8, 9, and 10 display the simulation results of the structure temperature as a function of heating power or power consumption. These relationships have already been explained in Eqs. 9, 11–12, and 18, 19, 20. Figures 7 and 8 are the simulation results of heat transfer for the reference structure, which can be used to calculate the heat transfer coefficient  $G_A$ , while Figs. 9 and 10 show the simulation results of heat transfer for the cantilever-based structures. For  $\Pi$ -shaped test microstructure, the thermal conductivity of polysilicon thin film can be estimated from



**Fig. 9** Simulation results: relationship between the heating power ( $P_{c1}$ ) and the temperature of the heating stripe ( $T_{c1}$ ) for the cantilever-based structure 1



**Fig. 10** Simulation results: relationship between the heating power ( $P_{c2}$ ) and the temperature of the heating stripe ( $T_{c2}$ ) for the cantilever-based structure 2



Figs. 7, 8, and 9 by substituting simulation results into Eq. 17. For T-shaped test microstructure, the thermal conductivity of polysilicon thin film can be obtained by Eq. 23 with simulation results depicted in Figs. 7, 9, and 10. The deviation of the calculation values of the thermal conductivity of polysilicon thin film is less than 5%, with respect to the set value for simulation.

### Experiment

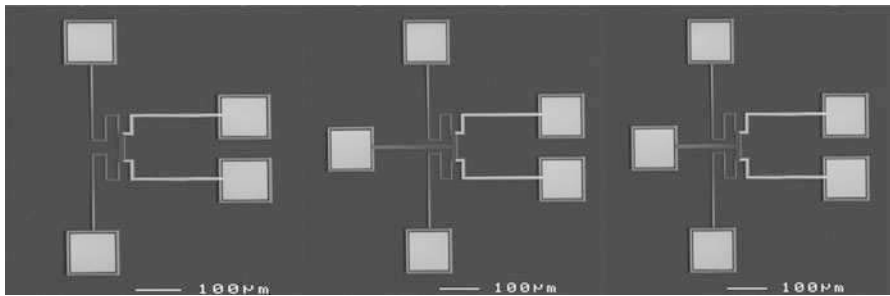
It is noted from Eqs. 17 and 23 that only heat power and the temperature change of the heating stripe are needed to extract the thermal conductivity of polysilicon thin film by using the proposed novel steady-state method. The heating power can be obtained by measuring the related current and voltage, and the temperature change may be determined using the measurement of polysilicon resistance (Leon et al. 2004). All measurements can be carried out in free air, while vacuum testing circumstance is not required, and only the four-probe arrangement, the digital

oscilloscope, and the semiconductor characterization system are used since only I–V measurements are needed. Therefore, the proposed test microstructures and test methods conform to the online test requirements.

The samples were fabricated by standard surface micromachining techniques. Figure 11 is the scanning electron microscopy (SEM) pictures of a reference structure and two cantilever-based structures.

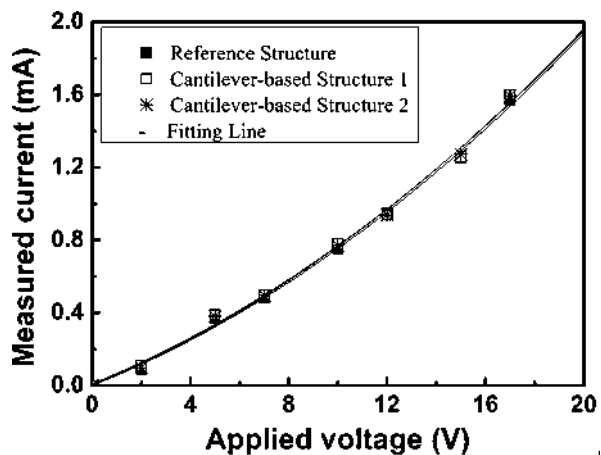
Voltages were applied between pads A and D. Joule heating will heat the reference structure and the cantilever-based structure, as illustrated in Figs. 1 and 4. The applied voltages and corresponding passing current were recorded. The voltage potential across the two central heating stripes was then be measured on pads B and C using a separate circuit.

Figures 12 and 13 show the applied voltage between A and D as a function of measured current and measured voltage across the heating stripe as a function of the

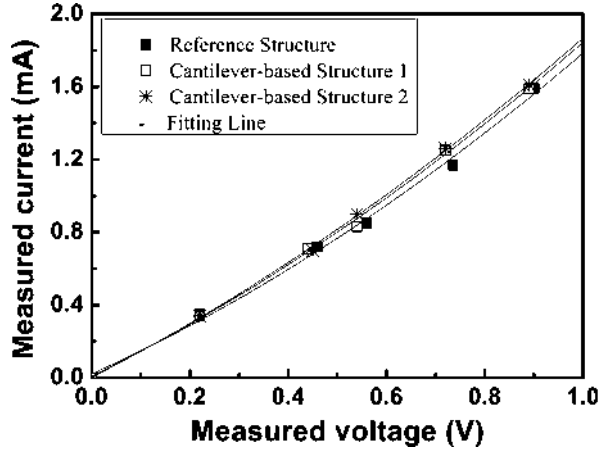


**Fig. 11** SEM photographs of the II-shaped and T-shaped test microstructures for the on-line test of thermal conductivity

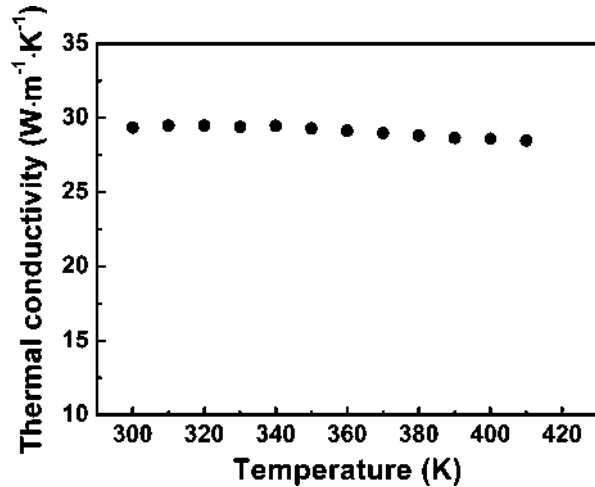
**Fig. 12** Applied voltage versus measured current



**Fig. 13** Measured voltage versus measured current

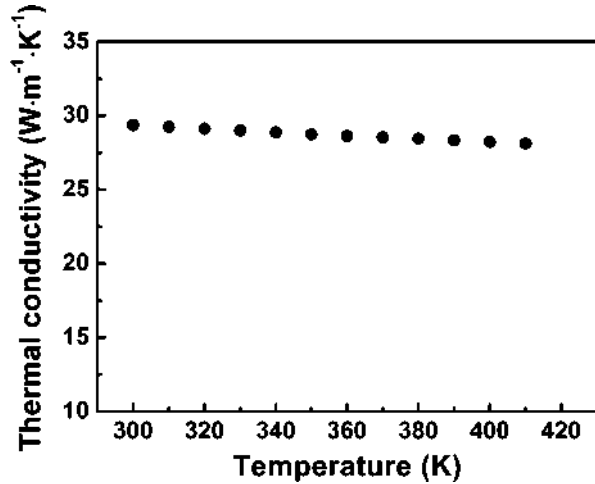


**Fig. 14** Thermal conductivity as a function of temperature in the heating stripe (*II-shaped test microstructure*)



measured current, respectively. From these figures, the heating powers  $P_R$ ,  $P_{c1}$  (or  $P_c$  for  $\Pi$ -shaped test microstructure), and  $P_{c2}$ , the heat consumption  $P_{Rh}$ , and the same temperature changes  $\Delta T_R$ ,  $\Delta T_{c1}$  (or  $\Delta T_c$  for  $\Pi$ -shaped test microstructure), and  $\Delta T_{c2}$  deduced by resistance temperature coefficient can be obtained, respectively. The thermal conductivity of polysilicon thin films can be then calculated using Eqs. 17 and 23, respectively. The thermal conductivity obtained from  $\Pi$ -shaped test microstructure and T-shaped test microstructure as a function of temperature in the heating stripe is demonstrated in Figs. 14 and 15, respectively. The average measured value of thermal conductivity of polysilicon thin film was estimated to be  $28.85 \pm 0.65 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$  with temperature ranging from 300 to 410 K.

**Fig. 15** Thermal conductivity as a function of temperature in the heating stripe (*T-shaped test microstructure*)



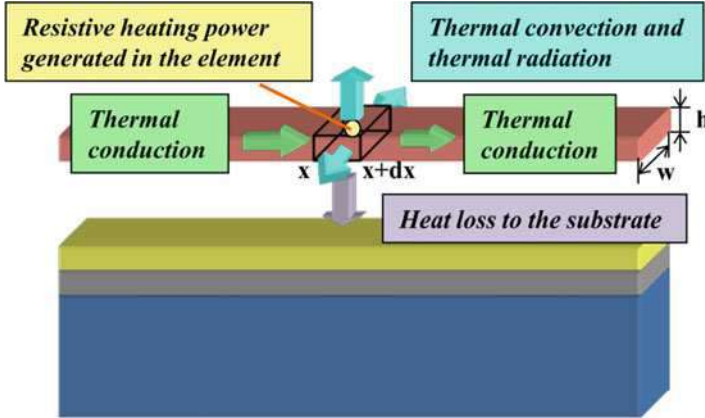
## Online Test Microstructure of Thermal Conductivity and Thermal Diffusivity

Thermal diffusivity is the central parameter for the time-dependent thermal response of MEMS devices. To seek out a solution for the characterization of thermal diffusivity, the transient-state thermal analysis of the double-clamped beam is discussed in this section. A test microstructure based on transient response of the double-clamped beam is then developed to determine both thermal diffusivity and thermal conductivity of the polysilicon beam. Different from the conventional transient-state method, the test microstructure presented here does not require vacuum environment for the measurement.

### Transient-State Thermal Analysis of a Double-Clamped Beam

Joule heating occurs as a direct electric current is applied through a double-clamped beam. Part of the heat turns into internal energy; the rest will transfer due to the presence of the temperature gradient. As illustrated in Fig. 16, heat transfer in the double-clamped beam includes thermal conduction, thermal convection, thermal radiation, and the heat loss through the air gap into the substrate.

Take a surface micromachined polysilicon double-clamped beam as an example, suppose the length of the beam is  $l$ , the width of the beam is  $w$ , and the thickness of the beam is  $h$ , where  $l$  is much larger than  $w$  and  $h$ . Under this condition, the temperature gradient only exists in the direction of beam length, and the electrothermal analysis of the double-clamped beam is generally simplified for one dimension (Lin and Chiao 1996). For the element of length  $dx$ , the energy balance is written as



**Fig. 16** Scheme of the heat transfer in the double-clamped beam

$$\begin{aligned}
 -k_pwh \left[ \frac{dT}{dx} \right]_x + J^2 \rho wh dx &= -k_pwh \left[ \frac{dT}{dx} \right]_{x+dx} + C \rho_m wh dx \frac{\partial T}{\partial t} \\
 &+ \beta(w + 2h) dx (T - T_0) \\
 &+ \varepsilon^* \delta(w + 2h) dx (T^4 - T_0^4) + S dx w \frac{T - T_0}{R_T}
 \end{aligned} \tag{24}$$

where  $T$  and  $T_0$  are the temperature in the double-clamped beam and substrate, respectively, and  $\rho_m$  and  $C$  denote density and heat capacity of polysilicon, respectively. Other parameters in this equation have been defined in section “[Online Test Microstructure of Thermal Conductivity](#).” In comparison with Eq. 5, one additional term,  $C \rho_m wh dx \partial T / \partial t$ , is included in Eq. 24. This term describes the change in internal energy.

According to the Stefan–Boltzmann law, if the average temperature of the beam does not exceed 600 °C, the effective radiation losses can be linearized or even ignored (Chae et al. 1999). A second-order partial differential heat equation can be derived as (Huang and Lee 1999)

$$\frac{\partial^2 T}{\partial x^2} = \frac{1}{a_p} \frac{\partial T}{\partial t} + \gamma \left( T - T_0 - \frac{J^2 \rho_0}{k_p \gamma} \right) \tag{25}$$

with

$$\gamma = \frac{\beta(2h + w) + 4\varepsilon^* \delta(2h + w)T_0^3 + \frac{S}{R_T} w}{k_p wh} - \frac{J^2 \rho_0 \xi}{k_p} \tag{26}$$

$$a_p = \frac{k_p}{\rho_m C} \tag{27}$$

where  $\rho_0$  is the initial electrical resistivity of the polysilicon thin film at temperature  $T_0$ ,  $a_p$  represents thermal diffusivity of polysilicon thin film, and  $\xi$  denotes the linear temperature coefficient of resistance of polysilicon thin film. In order to derive the steady-state solution and transient-state solution, it is assumed that

$$\begin{cases} T(x, 0) = T_0 \\ T(0, t) = T_0 \\ T(l, t) = T_0 \end{cases} \quad (28)$$

By solving Eq. 28, one obtains the temperature distribution of the beam as follows:

$$T(x, t) = \Phi(x) + \sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi}{l}x\right) \exp\left(-\frac{t}{l^2/a_p(\gamma l^2 + n^2\pi^2)}\right) \quad (29)$$

with

$$\Phi(x) = T_0 + \frac{J^2\rho_0}{k\gamma} \left[ 1 - \frac{\cosh(\sqrt{\gamma}(x - l/2))}{\cosh(\sqrt{\gamma}l/2)} \right] \quad (30)$$

$$C_n = \frac{2}{l} \int_0^l [T_{\infty} - \Phi(x)] \sin\left(\frac{n\pi}{l}x\right) dx \quad (31)$$

where  $\Phi(x)$  indicates the temperature distribution of the beam under steady-state conditions.  $C_n$  has the following characteristics:

$$C_n = \begin{cases} 0 & \text{if } n \text{ is even} \\ \left(-\frac{2}{n\pi} + \frac{2n\pi}{\gamma l^2 + n^2\pi^2}\right) \frac{2J^2\rho_0}{k\gamma} & \text{if } n \text{ is odd} \end{cases} \quad (32)$$

It is noted that only odd terms are present in the summation. At sufficiently long time, the contribution of the high-order eigenfunctions,  $C_n$  ( $n \geq 2$ ), to the sum becomes negligible. Therefore,  $\sum C_n \approx C_1$ . Equation 29 is rewritten as

$$\begin{aligned} T(x, t) = T_0 + \frac{J^2\rho_0}{k_p\gamma} & \left\{ 1 - \frac{\cosh\left[\sqrt{\gamma}\left(x - \frac{l}{2}\right)\right]}{\cosh\left(\frac{\sqrt{\gamma}l}{2}\right)} \right\} \\ & - \frac{4J^2\rho_0 l^2}{k_p\pi(\gamma l^2 + \pi^2)} \sin\left(\frac{\pi}{l}x\right) \exp\left(-\frac{t}{\tau}\right) \end{aligned} \quad (33)$$

with the time constant

$$\tau = \frac{l^2}{a_p(\gamma l^2 + \pi^2)} \quad (34)$$

The transient average temperature of the beam can be expressed as a function of time:

$$\begin{aligned} T_{\text{ave}}(t) &= \frac{1}{l} \int_0^l T(x, t) dx \\ &= T_0 + \frac{J^2 \rho_0}{k_p \gamma} - \frac{2J^2 \rho_0}{k_p \gamma l \sqrt{\gamma}} \cdot \tanh\left(\frac{\sqrt{\gamma} l}{2}\right) - \frac{8J^2 \rho_0 l^2}{k_p \pi^2 (\gamma l^2 + \pi^2)} \cdot \exp\left(-\frac{t}{\tau}\right) \end{aligned} \quad (35)$$

When  $t \rightarrow \infty$ , the steady-state average beam temperature is

$$\bar{T}_{\infty} = T_0 + \frac{J^2 \rho_0}{k_p \gamma} - \frac{2J^2 \rho_0}{k_p \gamma l \sqrt{\gamma}} \cdot \tanh\left(\frac{\sqrt{\gamma} l}{2}\right) \quad (36)$$

Thus, the average beam temperature change at steady-state is

$$\Delta \bar{T} = \frac{J^2 \rho_0}{k_p \gamma} - \frac{2J^2 \rho_0}{k_p \gamma l \sqrt{\gamma}} \cdot \tanh\left(\frac{\sqrt{\gamma} l}{2}\right) \quad (37)$$

## Thermal Resistors

Thermal resistor is an electrical resistor with appreciable temperature sensitivity. The doped polysilicon is often used as a temperature-sensitive resistive element. The electrical resistivity of polysilicon thermal resistor changes with temperature. Usually, we assume that the electrical resistivity of polysilicon has a linear temperature coefficient:

$$\rho = \rho_0[1 + \xi(T - T_0)] \quad (38)$$

where  $\rho$  and  $\rho_0$  stand for the electrical resistivity of the polysilicon thermal resistor at temperature  $T$  and  $T_0$ , respectively, and  $\xi$  is the linear temperature coefficient of resistance (TCR), whose value depends on the concentration of doping (Schafer et al. 1989).

The TCR of polysilicon thermal resistor can be obtained by simply measuring the electrical resistance of the resistor under different temperatures provided by a temperature-controlled heating plate. The electrical resistance of the polysilicon thermal resistor at temperature  $T$  is

$$R = R_0[1 + \xi(T - T_0)] \quad (39)$$

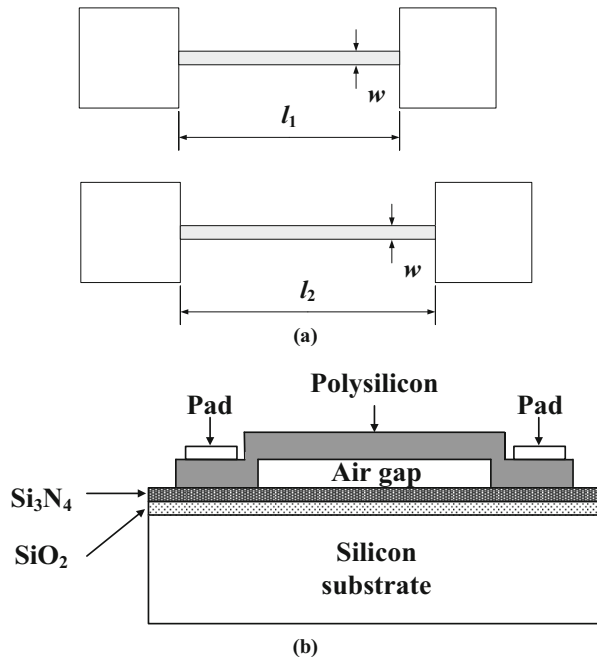
where  $R_0$  is the electrical resistance of the polysilicon thermal resistor at temperature  $T_0$ . The current and voltage used to interrogate the resistance value of a thermal resistor may introduce heat to it, which may change the temperature and the resistance value of the thermal resistor. This phenomenon is called self-heating (Liu 2006). The temperature and hence resistance value of a resistor under self-heating condition is dependent on the rate of heat input and dissipation through conduction convection and radiation.

## Test Microstructure and Theoretical Model

As shown in Fig. 17, the online test microstructure for thermal conductivity and thermal diffusivity of surface micromachined polysilicon thin films consists of two double-clamped beams with the same width ( $w$ ) and thickness ( $h$ ) but different lengths ( $l_1$  and  $l_2$ ) (Huang et al. 2006). The beam is heated by applying a direct current between two anchors. The transient average temperature of the beam can be calculated from Eq. 35. The electrical resistivity of polysilicon is assumed to have a linear temperature coefficient. The transient electrical resistance of the polysilicon beam may be expressed as below:

$$\bar{R}(t) = R_0 [1 + \xi(\bar{T}(t) - T_0)] = R_\infty + D \exp\left(-\frac{t}{\tau}\right) \quad (40)$$

**Fig. 17** Schematic illustrations of the online test microstructure for measuring the thermal conductivity and thermal diffusivity of polysilicon thin film. **(a)** Top view. **(b)** Cross section view





with

$$D = -\frac{8\xi J^2 R_0^2 whl}{k_p \pi^2 (\gamma l^2 + \pi^2)} \quad (41)$$

$$R_\infty = R_0 + \frac{J^2 R_0^2 \xi wh}{k_p \gamma l} - \frac{2J^2 R_0^2 \xi wh}{k_p \gamma l^2 \sqrt{\gamma}} \cdot \tanh\left(\frac{\sqrt{\gamma} l}{2}\right) \quad (42)$$

where  $R_0$  is the initial electrical resistance of the polysilicon beam before it is heated and  $R_\infty$  is the steady-state electrical resistance of the polysilicon beam after the current is applied. Electrical resistivity  $\rho_0 = R_0 wh/l$  can be calculated when  $R_0$  is measured.

By rearranging Eq. 40 and using the Log function on both sides of Eq. 40, the following equation can be found:

$$\begin{cases} \ln[R_\infty - \bar{R}(t)] = \ln(-D) - \frac{t}{\tau} & \text{when } \xi > 0 \\ \ln[\bar{R}(t) - R_\infty] = \ln(D) - \frac{t}{\tau} & \text{when } \xi < 0 \end{cases} \quad (43)$$

When  $\ln((R_\infty - R(t))/R_0)$  is plotted as the Y-axis and time  $t$  as the X-axis, the slope of the linear fitting curve is  $-1/\tau$ . By measuring the resistance as a function of time  $R(t)$ , the parameter  $\tau$  can be determined from the linear fitting curve. It is noted from Eq. 26 that the thermal diffusivity of polysilicon  $a_p$  and the parameter  $\gamma$  are independent of the beam length  $l$ . By measuring electrical resistance characteristics of two polysilicon double-clamped beams with the same thickness and width but different lengths  $l_1$  and  $l_2$ ,  $\gamma$  is given by

$$\gamma = \frac{\pi^2 \left( \frac{\tau_1}{l_1^2} - \frac{\tau_2}{l_2^2} \right)}{\tau_2 - \tau_1} \quad (44)$$

where  $\tau_1$  and  $\tau_2$  are the time constants of the two beams heated by the same current.

By substituting the value of  $\gamma$  into Eq. 42, thermal conductivity of polysilicon thin film can be obtained:

$$k_p = \frac{J^2 R_0^2 \xi \left[ l \sqrt{\gamma} - 2 \tanh\left(\frac{\sqrt{\gamma} l}{2}\right) \right]}{\gamma wh l^2 \sqrt{\gamma} (R_\infty - R_0)} \quad (45)$$

Thermal diffusivity of polysilicon thin film may be estimated from Eq. 34:

$$a_p = \frac{\left( \frac{1}{\tau_1} - \frac{1}{\tau_2} \right)}{\pi^2 \left( \frac{1}{l_1^2} - \frac{1}{l_2^2} \right)} \quad (46)$$

### Simulation and Verification

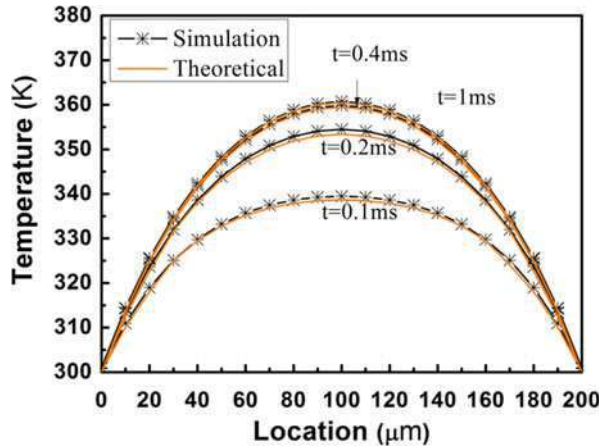
The finite element software ANSYS is used here to perform the electrothermal analysis of the double-clamped beam. The parameters for the simulation are listed in Table 2.

Figures 18 and 19 demonstrate the transient and steady temperature distribution of a 200  $\mu\text{m}$  long double-clamped beam along the beam length direction, respectively. Comparing the finite element method (FEM) simulation with the theoretical model, it is observed that the theoretical model agrees well with the FEM simulation. Figures 20 and 21 compare the transient average temperature and the transient electrical resistance obtained by theoretical model and FEM simulation. The simulation results are identical with the analytical solution. It is shown that the transient average temperature and the transient electric resistance curves first rise rapidly and

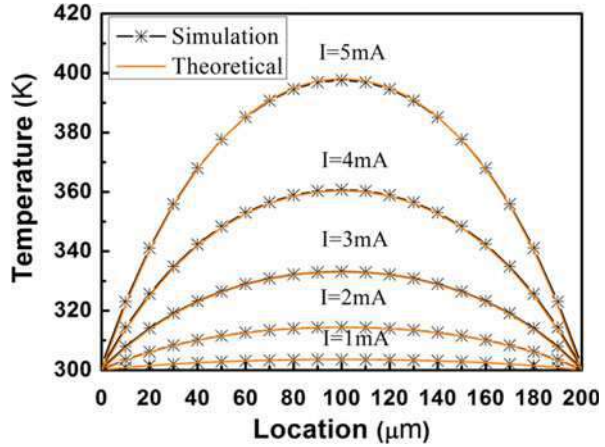
**Table 2** Parameters for FEM simulations

Parameter	Value	Parameter	Value
$l$	200 or 250 $\mu\text{m}$	$w$	10 $\mu\text{m}$
$h$	2 $\mu\text{m}$	$k_p$	$30 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
$t_v$	2 $\mu\text{m}$	$k_v$	$0.026 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
$t_n$	0.6 $\mu\text{m}$	$k_n$	$2.25 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
$t_s$	0.5 $\mu\text{m}$	$k_s$	$30 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
$t_o$	0.3 $\mu\text{m}$	$k_o$	$1.4 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
$\rho_0$	$2.11 \times 10^{-5} \Omega\cdot\text{m}$	$\rho_m$	$2.3 \times 10^3 \text{ kg}\cdot\text{m}^{-3}$
$\xi$	$1.25 \times 10^{-3} \text{ K}^{-1}$	$C$	$0.7 \times 10^3 \text{ J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$
$\beta$	$10 \text{ W}\cdot\text{m}^{-2}\cdot\text{K}^{-1}$	$\epsilon^*$	0.6
$T_0$	300 K	$I$	4 mA
$\alpha_T$	$2.6 \times 10^{-6} \text{ K}^{-1}$	$a_p$	$1.86 \times 10^{-5} \text{ m}^2\cdot\text{s}^{-1}$

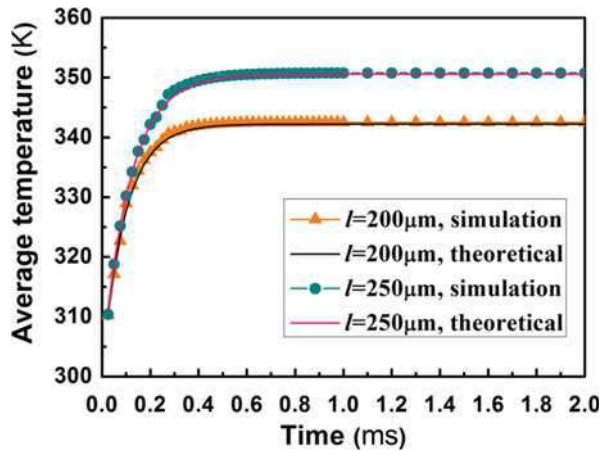
**Fig. 18** Comparison of the temperature distribution of a 200  $\mu\text{m}$  long double-clamped beam under transient-state conditions obtained by theoretical model and FEM simulation



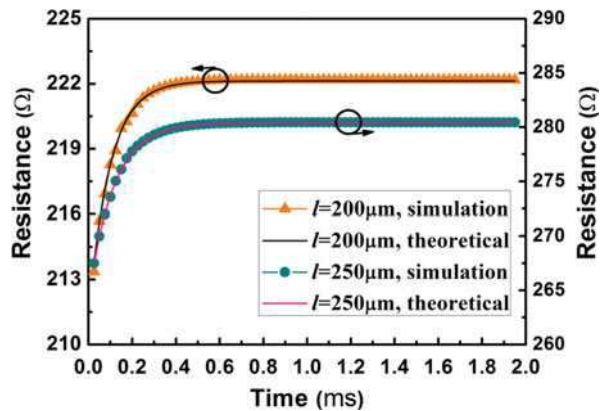
**Fig. 19** Comparison of the temperature distribution of a 200  $\mu\text{m}$  long double-clamped beam under steady-state conditions obtained by theoretical model and FEM simulation



**Fig. 20** Comparison of the transient average temperature of two double-clamped beams obtained by theoretical model and FEM simulation



**Fig. 21** Comparison of the transient electrical resistance of two double-clamped beams obtained by the theoretical model and the FEM simulation



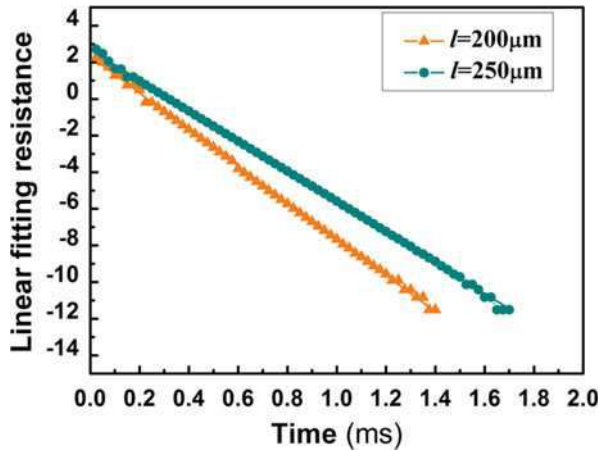
then slow down and tend to reach a constant value. These comparisons demonstrate that Eqs. 30, 35, 36, and 40 are valid.

By substituting simulation results into Eq. 43, two linear fitting curves are obtained and depicted in Fig. 22. The parameter  $\gamma$  and thermal diffusivity are then calculated from Eqs. 44 and 46 with the information displayed in Fig. 22. Substituting the value of  $\gamma$  into Eq. 45, thermal conductivity of the double-clamped beam is determined. The deviation of the calculation values of the thermal conductivity and thermal diffusivity of polysilicon thin film are less than 5%, with respect to the set value for simulation listed in Table 2.

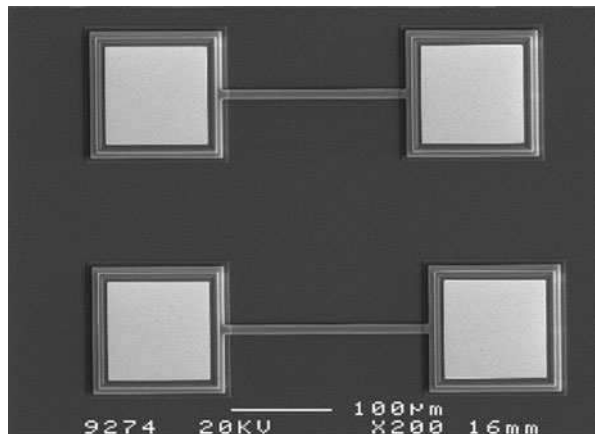
## Experiment

Figure 23 shows the SEM pictures of the test microstructure fabricated by surface micromachining process. The two different beams have 15  $\mu\text{m}$  width, 2  $\mu\text{m}$  thick-

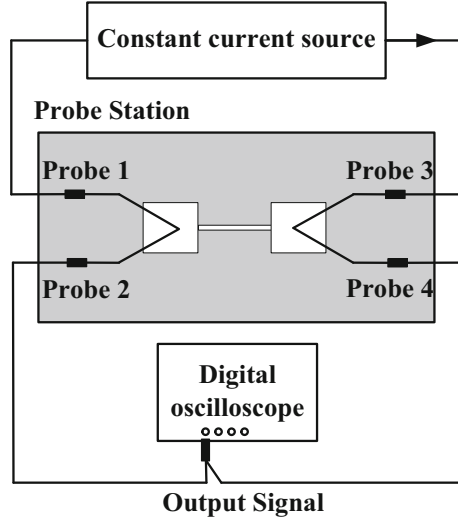
**Fig. 22** Linear fitting resistance of two double-clamped beams



**Fig. 23** SEM photographs of the double clamped beams for the on-line test of thermal conductivity and thermal diffusivity



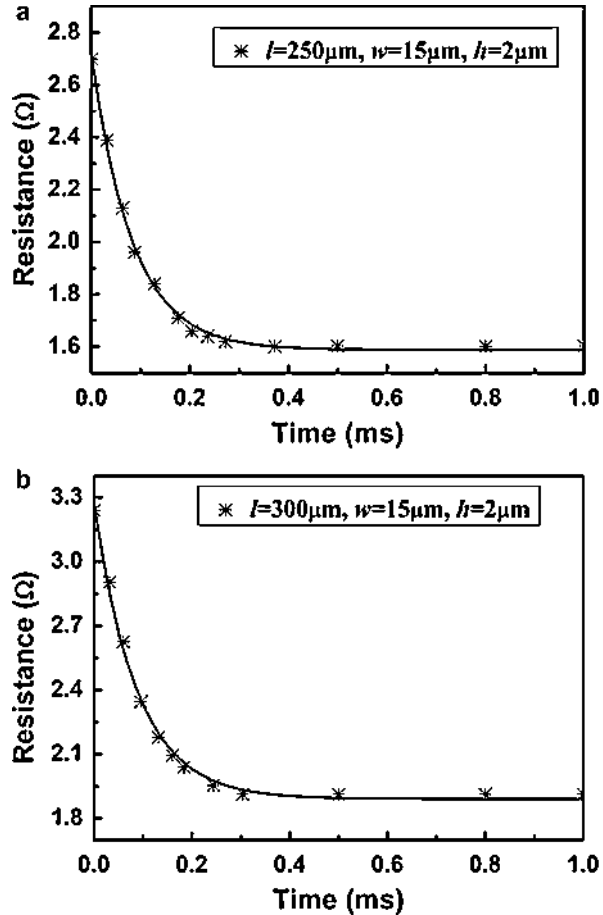
**Fig. 24** The measurement setup for the on-line test of thermal conductivity and thermal diffusivity



ness, and 250 and 300  $\mu\text{m}$  lengths, respectively. The measurement setup is illustrated in Fig. 24. The heating current is applied by a constant current source, while the resistance characteristic of the test microstructure is stored by a digital oscilloscope. All the measurements can be directly performed on the four-point probe station in free air. The transient electrical resistance of two samples is demonstrated in Fig. 25, and their linear fitting curves are displayed in Fig. 26. The slope of the curve in Fig. 26 is the negative reciprocal of the thermal time constant. It is shown that the test microstructures have negative temperature coefficients of resistance. The thermal conductivity can be extracted according to Eq. 45, and the thermal diffusivity is obtained by Eq. 46. Because of the temperature difference between the anchor and the center of the beam, the measured value of thermal conductivity and thermal diffusivity is actually the average value along the beam. The value of the thermal conductivity and thermal diffusivity of polysilicon thin films was measured to be  $31.32 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$  and  $0.165 \text{ cm}^2 \text{ s}^{-1}$ , respectively.

In our electrothermal model, some approximations are made. In Eq. 29, the higher-order terms  $\sum Cn$  are neglected, contributing an error less than 3%. Another source of error comes from the applied current. If the current is small, the Joule heat is so small that the thermal stress and average temperature change could not be measured. If the current is too large, the polysilicon beam may be overheated and destroyed. Therefore, the applied current should be controlled properly. Besides, the radiation losses are simplified to be linear with temperature differences, and the material parameters are assumed to remain constant during the heating process. Thus we suggest that the applicable range of the temperature is from room temperature up to 400 K. When considering the errors caused by dimensions for the beams, the measurement error of the thermal conductivity and thermal diffusivity are both estimated to be less than 10%.

**Fig. 25** Transient electrical resistance of test microstructures after heating by a current of 2 mA

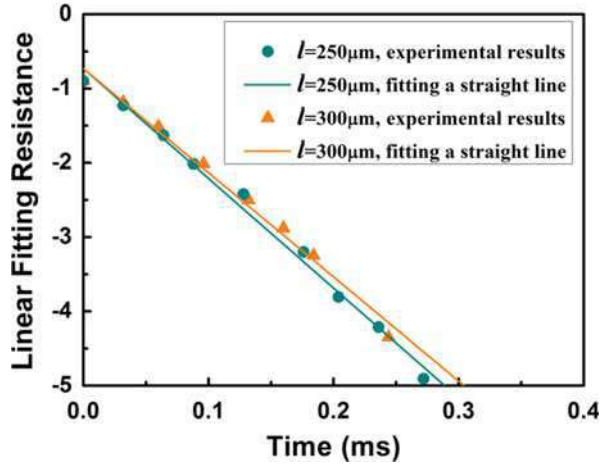


### Online Test Microstructure of the Coefficient of Thermal Expansion by Rotating Technique

Thermal expansion is an important thermomechanical effect, which denotes the tendency of matter to change in size as it is heated (Senturia 2001). The rate of the fractional change in size with temperature is described by the coefficient of thermal expansion (CTE). Several types of coefficients have been developed: volumetric, area, and linear. For conducting beams, the change along beam length, and hence linear expansion, is only considered, since the beam length is much larger than beam width and beam thickness. Linear coefficient of thermal coefficient is defined as

$$\alpha_T = \frac{1}{l} \frac{\Delta l}{\Delta T} \quad (47)$$

**Fig. 26** Linear fitting curves of measured transient electrical resistances



where  $\Delta l$  represents the change of beam length due to temperature variation  $\Delta T$ . It is clear from Eq. 47 that if  $\Delta l$  and  $\Delta T$  are obtained, the coefficient of thermal expansion is then determined.

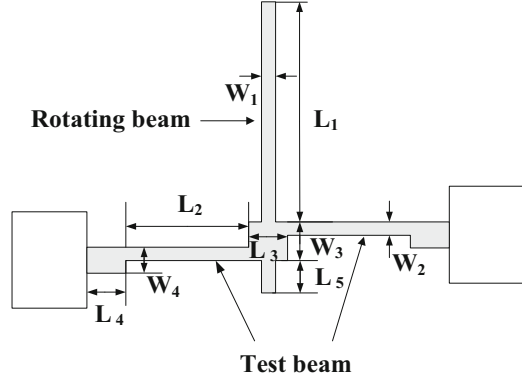
The idea presented in this section is to utilize the deflection of micro-rotating structures caused by thermal expansion. The structures are heated electrically and exhibit an in-plane deflection. The lateral displacement is designed to be a constant value, while the average temperature of the test microstructure is obtained from the measurement of resistance of the test microstructure. The coefficient of thermal expansion is then extracted. Instead of an optical or visual readout, which is generally used in previously reported methods, an electrical readout, which is better compatible with standard IC test equipment, is adopted. In the experiments, current–voltage measurements are only required and all measurements can be operated in atmosphere. The proposed extracting method is simple, fast, and inexpensive and has a potential in applications of online monitoring surface micro-machining process lines.

## Micro-rotating Structures

### Double-Test-Beam Micro-rotating Structure

The schematic view of the double-test-beam micro-rotating structure is illustrated in Fig. 27. The clear parts are movable structures and the dark parts are anchors. The structure consists of one rotating beam and two test beams. The test beams are connected off-center to the opposite sides of the rotating beam. As shown in Fig. 27,  $L_1$  is the length of the top side of the rotating beam;  $L_5$  is the length of the bottom side of the rotating beam;  $W_1$  is the width of the rotating beam;  $L_2$  and  $W_2$  are the length and width of the test beam, respectively;  $L_3$  and  $W_3$  are the length and width of the connecting part between test beams, respectively; and  $L_4$  and  $W_4$  are the length and

**Fig. 27** Schematic diagram of the double-test-beam micro-rotating structure



width of the connecting part between the test beam and the anchor, respectively. When forcing a current through two test beams, the double-test-beam micro-rotating structure is heated with Joule heat, and the thermal expansion of the test beams occurs. Since the length of the test beam is much larger than its width and thickness, thermal expansion mainly occurs along the length direction of the test beams. The thermal strain is

$$\varepsilon_T = \alpha_T (T_{\text{avr}} - T_0) \quad (48)$$

where  $\varepsilon_T$  is the thermal strain along the length direction of the test beam,  $\alpha_T$  is the CTE,  $T_{\text{avr}}$  is the average temperature of the test beam when the current is applied, and  $T_0$  is the initial temperature of the test beam. It is obvious that  $\Delta T = T_{\text{avr}} - T_0$  is the temperature change of the test beam.

The thermal expansion of the test beam due to the thermal strain causes the movement of the rotating beam. The rotating beam magnifies the tiny expansion of the test beam, and a large displacement is generated at the tip of the rotating beam. For a given thermal strain, the expansion of the test beam is directly proportional to the length of the test beam. A large expansion leads to a large moment acting on the rotating beam and results in a large displacement. The rotation center of the rotating beam is located at the center of the connecting part between test beams. Therefore, for a given rotation angle, the lateral displacement of the tip of the rotating beam is directly in proportion to the ratio of the distance between the rotating beam tip and the rotation center ( $L_1 + W_3/2$ ) over the vertical distance between the rotation center and the horizontal center line of the test beam ( $W_3/2 - W_2/2$ ). These relationships can be represented as

$$\delta = \frac{\varepsilon_T \cdot L_2 \cdot (L_1 + W_3/2)}{(W_3/2 - W_2/2)} \quad (49)$$

where  $\delta$  is the lateral displacement of the rotating beam tip.

The moment applied by the test beam to the rotating beam is related to the released force generated from the thermal stress of the test beam. The test beam



cannot transfer all the thermal stress to the rotating beam because the free end of the test beam partially releases the thermal stress in the test beam. Therefore, Eq. 49 is modified as

$$\delta = \frac{\epsilon_T \cdot L_2 \cdot (L_1 + W_3/2) \cdot D}{(W_3/2 - W_2/2)} \tag{50}$$

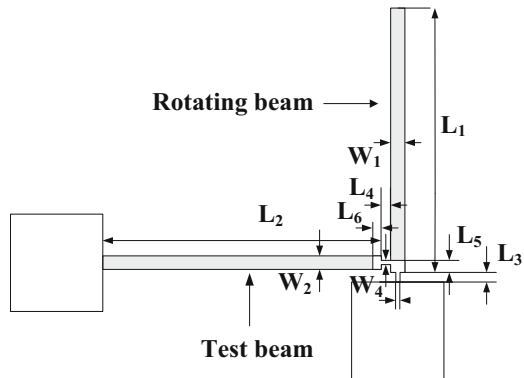
where  $D$  ( $0 \leq D \leq 1$ ) is a correction factor, which will be investigated and discussed in section “Simulation and Verification.”

**Single-Test-Beam Micro-rotating Structure**

The schematic illustration of single-test-beam micro-rotating structure is shown in Fig. 28. The structure is composed of a rotating beam and a test beam.  $L_1$  is the length of the rotating beam;  $W_1$  is the width of the rotating beam;  $L_2$  and  $W_2$  are the length and width of the test beam, respectively;  $L_3$  and  $W_3$  are the length and width of the neck between the rotating beam and the anchor, respectively; and  $L_4$  and  $W_4$  are the length and width of the neck between the test beam and the rotating beam, respectively. The bottom part of the rotating beam between necks with a length of  $L_5$ , the right part of the test beam with a length of  $L_6$ , and all necks are covered by a metal layer so that the Joule heat will mainly occur in the left part of the test beam with a length of  $L_2-L_6$ . The structure is heated by applying a current through two anchors, causing the thermal expansion of the test beam. The thermal strain is a function of CTE and temperature change of the test beam, as demonstrated in Eq. 48. It is noted that  $T_0$  and  $T_{avr}$  for single-test-beam micro-rotating structure are the average temperature of the effective test beam, the single-layer part of the test beam with a length of  $L_2-L_6$ , before and after the structure is heated.

The rotating beam deflects due to the thermal expansion of the test beam and magnifies the expansion. The rotation center of the rotating beam is located at the center of the bottom edge of the neck between the rotating beam and the anchor. Similar to the double-test-beam micro-rotating structure, the relationship between the lateral displacement at the tip of the rotating beam and the thermal strain along the test beam is derived as

**Fig. 28** Schematic diagram of the single-test-beam micro-rotating structure



$$\delta = \frac{\varepsilon_T \cdot (L_2 - L_6) \cdot (L_1 + L_3) \cdot D}{(L_3 + L_5 - W_4/2)} \quad (51)$$

where  $D$  ( $0 \leq D \leq 1$ ) is a correction factor induced to improve the analytical model.

## Test Microstructure and Theoretical Model

### H-Shaped Test Microstructure

Figure 29 depicts the H-shaped test microstructure for online test of the CTE of polysilicon thin film, which consists of a double-test-beam micro-rotating structure and two contact electrodes (Liu et al. 2012a, 2014). As shown in Fig. 29a, the initial gap between the tip of the rotating beam and the contact electrode is designed to be a constant value, i.e., 2  $\mu\text{m}$ . The test beams are electrically heated, and the expansion of the test beams causes the rotating beam to rotate counterclockwise. When the tip of the rotating beam touches the left contact electrode, the displacement of the rotating beam ( $\delta$ ) is equal to the initial gap between the tip and the left contact electrode ( $g$ ). The CTE can be expressed from Eqs. 48 and 50 as

$$\alpha_T = \frac{g \cdot (W_3/2 - W_2/2)}{(T_{\text{avr}} - T_0) \cdot (L_1 + W_3/2) \cdot L_2 \cdot D} \quad (52)$$

It is clear that if the structure geometry and the correction factor are known and the average temperature of the test beam when the contact occurs is measured, the CTE can be extracted.

Assuming  $T_0$  is equal to room temperature, i.e., 300 K,  $T_{\text{avr}}$  can be extracted by using the measured polysilicon resistance, which is

$$T_{\text{avr}} = T_0 + \frac{R_{iT} - R_{i0}}{R_{i0} \cdot \xi} \quad (53)$$

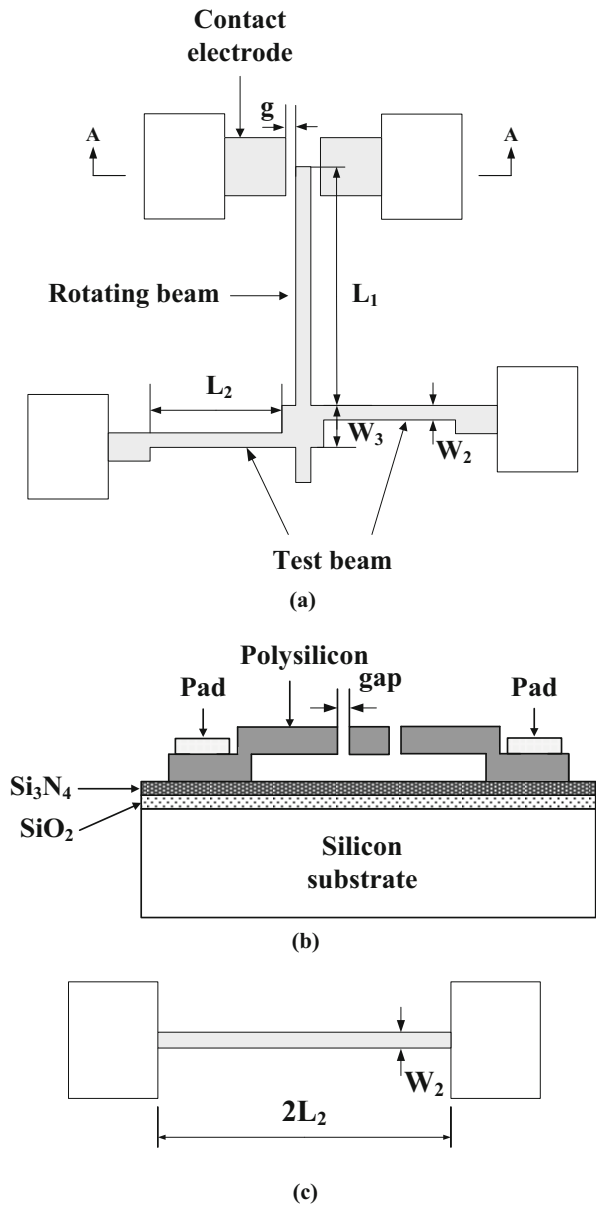
where  $R_{iT}$  is the total electrical resistance of the two test beams after heated and  $R_{i0}$  is the total initial electrical resistance of two test beams. A microbridge with a length of  $2L_2$  and a width of  $W_2$  was designed as a compensation structure to measure  $R_{i0}$ , as displayed in Fig. 29c. Compared with the two test beams, the resistance change of the connecting parts (shown in Fig. 29a) due to the Joule heating is negligible. Therefore, the resistance change between pad A and pad B due to the Joule heating is approximate to the resistance change of two test beams. The temperature change of the test beams are estimated as

$$T_{\text{avr}} = T_0 + \frac{R_{ABT} - R_{AB0}}{R_{i0} \cdot \xi} \quad (54)$$

where  $R_{ABT}$  is the resistance between pad A and pad B after heating and  $R_{AB0}$  is its initial resistance.

H-shaped test microstructure for the extraction of CTE makes use of the rotating technique. The location of the contact electrodes is tactically arranged so that the deflection of the test microstructure is a known value instead of using optical techniques to measure the deflection reported in traditional methods. As a result,

**Fig. 29** Schematic diagram of the H-shaped test microstructure. (a) Top view, (b) cross section view, and (c) compensation structure

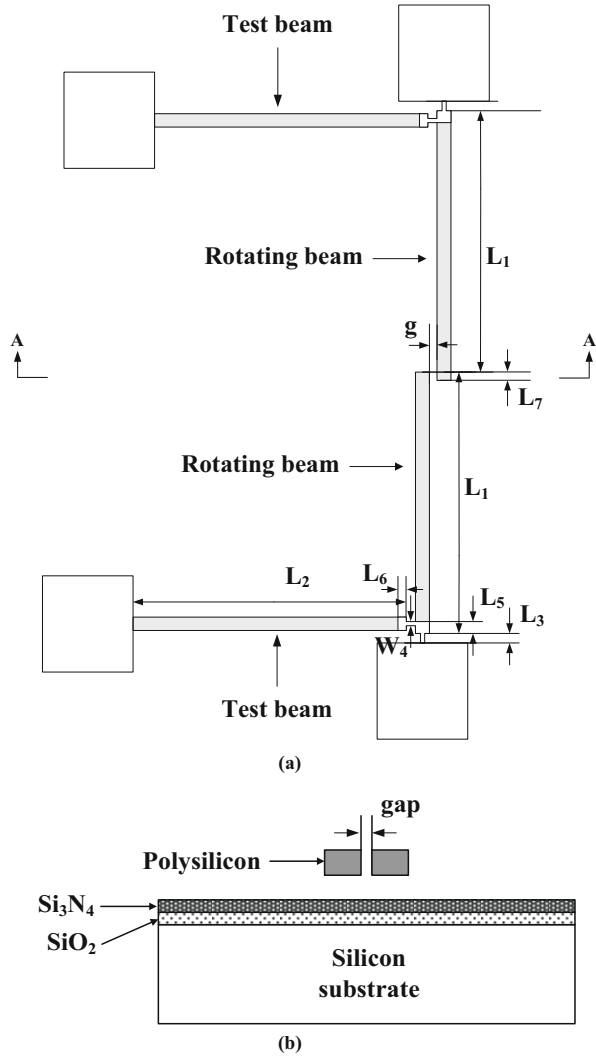


only I–V measurements with ordinary IC test equipment are needed, which conforms to online test requirements. However, the deflection induced by initial residual stress is not considered in this structure. This approximation may lead to a significant error if the residual stress is large enough. Therefore, H-shaped test microstructure is only suited for materials with small initial residual stresses.

**II-Shaped Test Microstructure**

II-shaped test microstructure for online test of the CTE of polysilicon thin film consists of two single-test-beam micro-rotating structures that are symmetrically placed (Liu et al. 2013). As shown in Fig. 30, the rotating beams point toward each

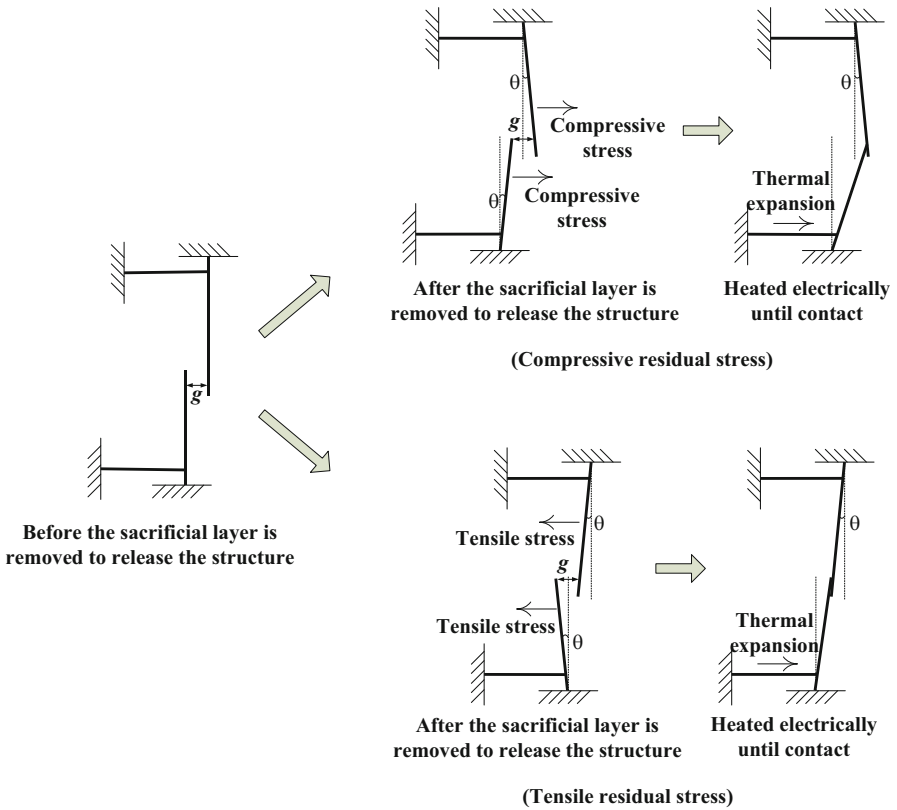
**Fig. 30** Schematic diagram of the II-shaped test microstructure. (a) Top view. (b) Cross section view



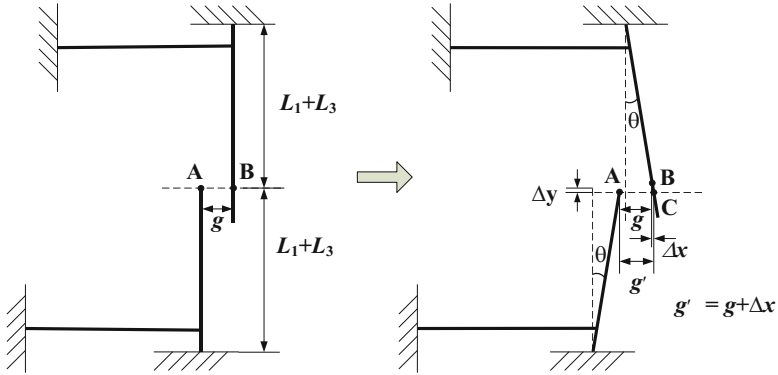
other, and both the test beams are connected on the left side to the rotating beam by a neck. The sizes of the top and bottom micro-rotating structures are the same except the length of the rotating beam. The rotating beam length of the bottom structure is  $L_1$ , while the rotating beam length of the top structure is  $L_1 + L_7$  in which  $L_7$  is the overlap length of these two rotating beams.

As illustrated in Fig. 31, the initial gap between the tips of the rotating beams is  $g$ . When the sacrificial layer is etched away, the test beams of two structures are free to extend or contract as a result of the compressive or tensile residual stress, and the rotating beams of two structures will rotate by the same angle in the same direction. As displayed in Fig. 32, the tip gap will increase by  $\Delta x$  after the rotation, where  $\Delta x = 2\Delta y \times \tan\theta = 2(L_1 + L_3) \times (\tan\theta - \sin\theta)$ . Since the rotation angle  $\theta$  is very small, the approximation,  $\tan\theta \approx \sin\theta$ , is assumed. Therefore,  $\Delta x$  is neglectable, and the gap between the tips of the rotating beams can be regarded as unchanged.

The  $\Pi$ -shaped test microstructure is heated by applying direct voltage current through two anchors of the bottom structure, and the rotating beam of the bottom



**Fig. 31** Schematic diagram of the mechanism of the  $\Pi$ -shaped test microstructure for the determination of CTE



**Fig. 32** Schematic diagram of the tip gap of the Π-shaped test microstructure when the residual stress is compressive

structure rotates clockwise. When the tip of the rotating beam of the bottom structure touches the tip of the rotating beam of the top structure, the tip displacement of the rotating beam ( $\delta$ ) is equal to the initial gap between the tips ( $g$ ). The CTE can be derived from Eqs. 48 and 51:

$$\alpha_T = \frac{g \cdot (L_3 + L_5 - W_4/2)}{(T_{avr} - T_0) \cdot (L_2 - L_6) \cdot (L_1 + L_3) \cdot D} \tag{55}$$

The temperature change of the effective test beam is extracted by

$$T_{avr} = T_0 + \frac{R_{iT} - R_{i0}}{R_{i0} \cdot \xi} \tag{56}$$

where  $R_{iT}$  is the electrical resistance of the effective test beam after heated and  $R_{i0}$  is the initial electrical resistance of the effective test beam. Since the areas of  $L_5$ ,  $L_6$ , and two necks are covered by metal, the resistance between two pads of the single-test-beam micro-rotating structure is approximate to the resistance of the effective test beam.

For Π-shaped test microstructure, the influence of the initial residual stress is canceled out by strategically placing the top and the bottom single-test-beam micro-rotating structures. The Π-shaped test microstructure is therefore suitable for materials with considerable initial residual stresses.

### Simulation and Verification

In order to verify the analytical model and determine the correction factor, ANSYS is used to perform the numerical analysis on both H-shaped test microstructure and Π-shaped test microstructure.

### H-Shaped Test Microstructure

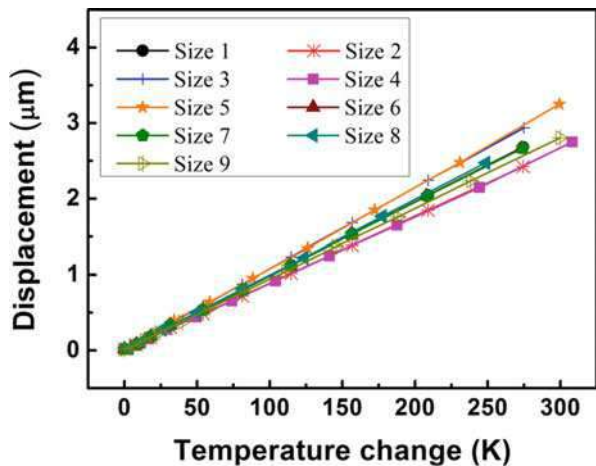
The dimensions of the test microstructures for simulations are listed in Table 3. Figure 33 displays simulation results of the tip displacement of the rotating beam as a function of the temperature change of the test beam. For given geometry parameters, the tip displacement increases linearly with the temperature change. This relationship is expressed in Eq. 50. As demonstrated in Eq. 52, the line slope in Fig. 33 is in proportion to the CTE,  $\alpha_T$ . Comparing the simulated results of structures with different dimensions in Fig. 33, it is observed that a larger  $L_2$ , or a larger  $L_1 + W_3/2$ , or a smaller  $W_3/2 - W_2/2$  creates a larger tip displacement at the same temperature change, while the variation of  $W_1$  has little influence on the tip displacement. This verifies that the theoretical model is basically correct.

As mentioned in section “Micro-rotating Structures,” a correction factor  $D$  ( $0 \leq D \leq 1$ ) is introduced into the theoretical model. In extreme cases, a correction factor of 1 is equivalent to totally releasing the thermal stress to the rotating beam, while with a correction factor of 0, all thermal stress remains on the test beam. In

**Table 3** Dimensions of the test microstructures for FEM simulation (in  $\mu\text{m}$ ) (Liu et al. 2012a, 2014)

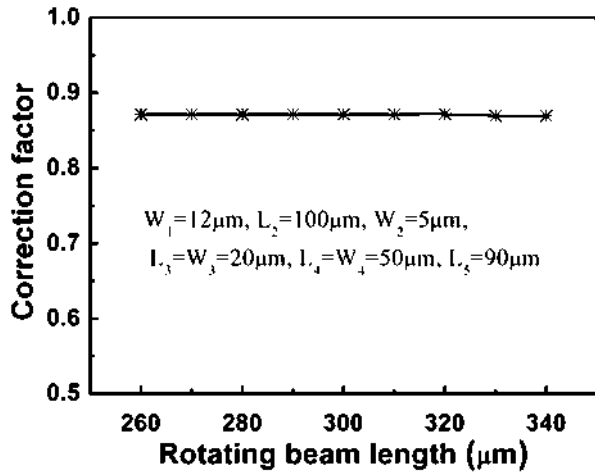
	$L_1$	$W_1$	$L_2$	$W_2$	$L_3$	$W_3$	$L_4$	$W_4$	$L_5$
Size 1	300	12	100	5	20	20	50	50	90
Size 2	270	12	100	5	20	20	50	50	90
Size 3	330	12	100	5	20	20	50	50	90
Size 4	300	12	90	5	20	20	50	50	90
Size 5	300	12	110	5	20	20	50	50	90
Size 6	300	10	100	5	20	20	50	50	90
Size 7	300	14	100	5	20	20	50	50	90
Size 8	300	12	100	8	20	20	50	50	90
Size 9	300	12	100	6	20	20	50	50	90

**Fig. 33** Simulation results of the tip displacement of the rotating beam as a function of the temperature change of the test beam

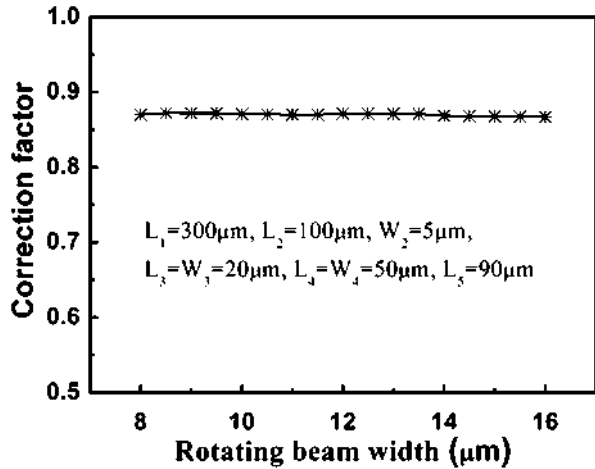


other words, a correction factor near 0 indicates more “correction” than a correction factor near 1. It is observed from the simulation results that the correction factor is related to geometry parameters and independent of the applied current if the maximum temperature of the test microstructure is lower than 680 K. This is because material physical properties change at high temperature and the micro-rotating structure shows nonlinear behavior at high temperature. Figures 34, 35, and 36 demonstrate the simulation relations of the correction factor  $D$  with geometrical parameters. Figure 34 shows the relationship between  $D$  and the dimensions of the rotating beam. The change in the correction factor is within 0.22% and 0.61% for  $260 \mu\text{m} \leq L_1 \leq 340 \mu\text{m}$  and  $8 \mu\text{m} \leq W_1 \leq 16 \mu\text{m}$ , respectively. Therefore, the

**Fig. 34** Simulation relations of the correction factor  $D$  with dimensions of the rotating beam. (a) Length of the top side of the rotating beam ( $L_1$ ) and (b) width of the rotating beam ( $W_1$ )



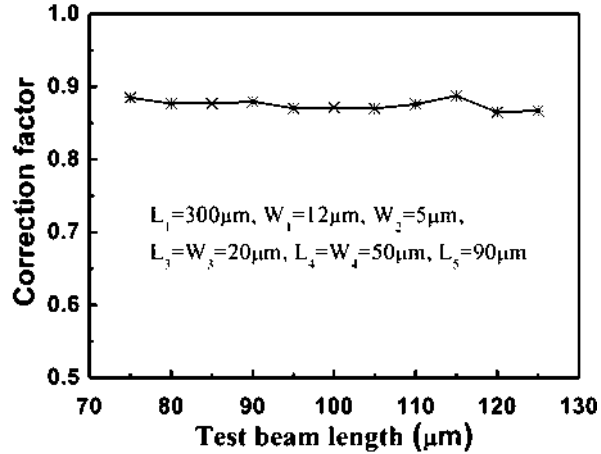
(a)



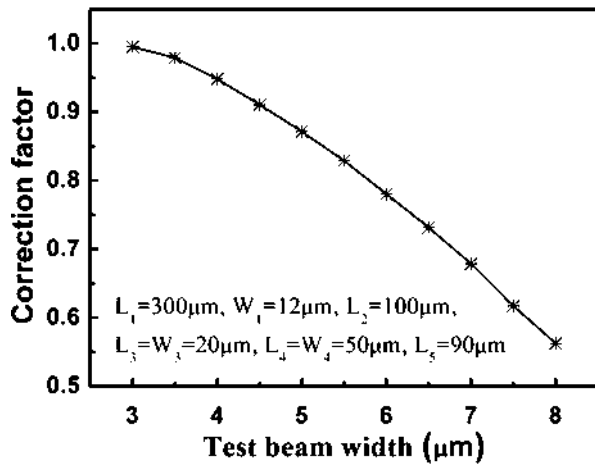
(b)



**Fig. 35** Simulation relations of the correction factor  $D$  with dimensions of the test beam. (a) Length of the test beam ( $L_2$ ) and (b) width of the test beam ( $W_2$ )



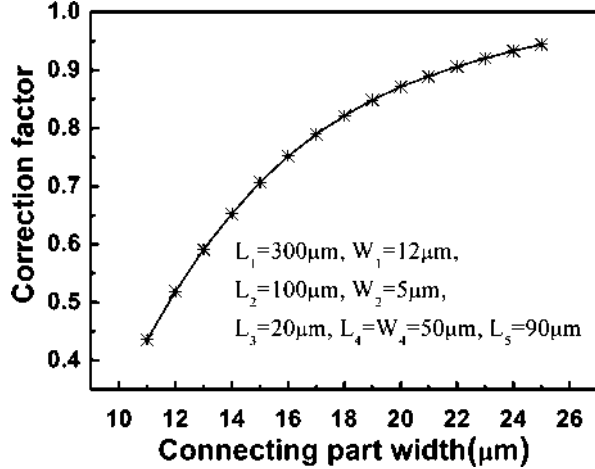
(a)



(b)

correction factor can be regarded as a constant for certain values of the geometry of the test beams and connecting parts. As can be seen in Fig. 35a, the correction factor does not change much as  $L_2$  varies. Hence, it can be assumed that the correction factor is independent of  $L_2$  for  $75 \mu\text{m} \leq L_2 \leq 125 \mu\text{m}$ . As shown in Figs. 35b and 6, the correction factor increases significantly as the width of the connecting part between test beams ( $W_3$ ) increases or the width of the test beam ( $W_2$ ) decreases. For a fixed  $W_2$  of  $5 \mu\text{m}$ ,  $D$  varies from 0.436 to 0.944 as  $W_3$  changes from 11 to  $25 \mu\text{m}$ , while for a given  $W_3$  of  $20 \mu\text{m}$ ,  $D$  decreases from 0.995 to 0.563 with  $W_2$  increasing from 3 to  $8 \mu\text{m}$ . The separation between horizontal center lines of two test beams can be expressed as  $W_3/2 - W_2/2$ . As mentioned before, the test beam partially transfers the thermal stress to the rotating beam. Stress remaining on the test beam decreases as the center separation increases. So more “correction” occurs

**Fig. 36** Simulation relations of the correction factor  $D$  with the width of the connecting part between test beams



at smaller center separation. This is the reason why increasing  $W_3$  or decreasing  $W_2$  makes the correction factor nearer to 1. Another reason for the significant influence of  $W_2$  on the correction factor is concerning the nonuniformity of the extension along the width direction at the free end of the test beam. Due to the presence and constraint of the rotating beam, the upper edge and the lower edge of the free end of the test beam cannot change their dimensions exactly the same. The wider the test beam is, the larger the constraint is and, therefore, the smaller the correction factor is.

The correction factor is estimated from the least-square fitting analysis based on a series of simulation data, and an empirical expression for the correction factor can be approximately written as

$$D \approx -0.191182508 \times e^{1.7392 \times 10^5 \times W_2} - 3.438927897 \times e^{-1.6393 \times 10^5 \times W_3} + 1.456823105 \tag{57}$$

### II-Shaped Test Microstructure

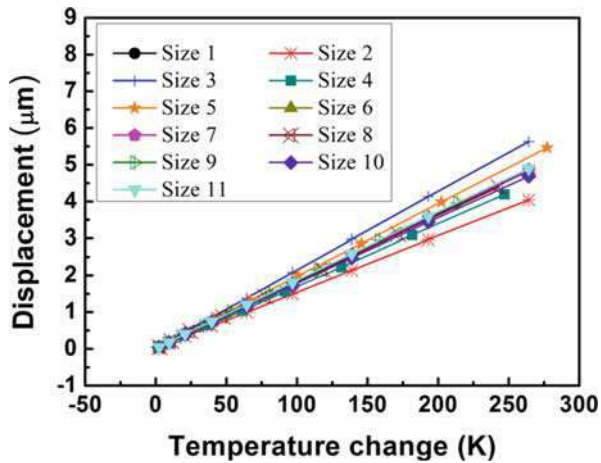
The dimensions of the test microstructures for simulations are listed in Table 4. Figure 37 shows simulation results of the tip displacement of the rotating beam as a function of the temperature change of the effective test beam. For given dimensions, the tip displacement of the rotating beam increases linearly with the temperature change of the effective test beam. This relationship is expressed in Eq. 51. Comparing the simulation results of structures with different dimensions in Fig. 37, it is observed that a larger  $L_2-L_6$ , or a larger  $L_1+L_3$ , or a smaller  $L_3+L_5-W_4/2$  creates a larger tip displacement of the rotating beam at the same temperature change of the test beam, while the variation of  $W_1$  or  $W_2$  has little influence on the tip displacement. This verifies that the analytical model is basically correct.

For the same reason as before, the correction factor  $D$  ( $0 \leq D \leq 1$ ) is derived from the simulation results to improve the analytical model. A correction factor near 0 indicates more “correction” than a correction factor near 1. If the temperature of

**Table 4** Dimensions of the test microstructures for FEM simulations (in  $\mu\text{m}$ ) (Liu et al. 2013)

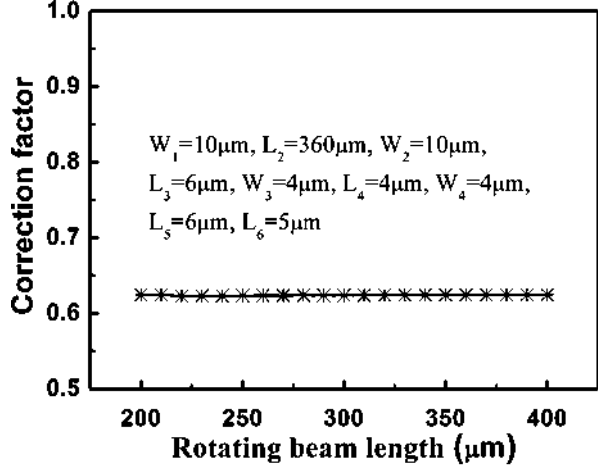
	$L_1$	$W_1$	$L_2$	$W_2$	$L_3$	$W_3$	$L_4$	$W_4$	$L_5$	$L_6$
Size 1	300	10	360	10	6	4	4	4	6	5
Size 2	250	10	360	10	6	4	4	4	6	5
Size 3	350	10	360	10	6	4	4	4	6	5
Size 4	300	10	320	10	6	4	4	4	6	5
Size 5	300	10	400	10	6	4	4	4	6	5
Size 6	300	9	360	10	6	4	4	4	6	5
Size 7	300	11	360	10	6	4	4	4	6	5
Size 8	300	10	360	9	6	4	4	4	6	5
Size 9	300	10	360	11	6	4	4	4	6	5
Size 10	300	10	360	10	6	4	4	4	5	5
Size 11	300	10	360	10	6	4	4	4	7	5

**Fig. 37** Simulation results of the tip displacement of the rotating beam as a function of the temperature change of the test beam

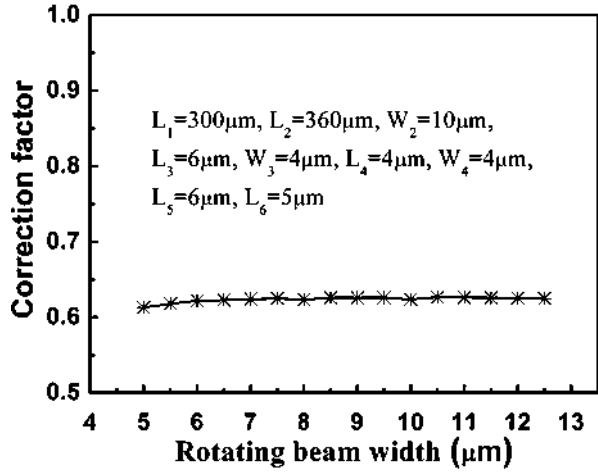


the test microstructure is below 850 K, the correction factor is only related to dimensions and independent of the applied current. Figures 38, 39, and 40 display the simulation relations of the correction factor  $D$  with geometrical parameters. Figure 38 shows the relationship between the correction factor and the dimensions of the rotating beam. When the length of the rotating beam ( $L_1$ ) varies from 200 to 400  $\mu\text{m}$ , or the width of the rotating beam ( $W_1$ ) increases from 5 to 12.5  $\mu\text{m}$ , the correction factor does not change much. Therefore, the correction factor can be regarded as a constant for  $200 \mu\text{m} \leq L_1 \leq 400 \mu\text{m}$  or  $5 \mu\text{m} \leq W_1 \leq 12.5 \mu\text{m}$ , for certain values of the geometry of the test beam and the necks. Figure 39 demonstrates the effect of the test beam sizes, length  $L_2$  and width  $W_2$ , on the correction factor  $D$ . The test beam would totally release the residual stress without the presence of the rotating beam. In that case, the test beam would extend only in the direction of beam length as we assumed in sections “[Test Microstructure and Theoretical Model](#)” and “[Simulation and Verification](#).” However, since the test beam is connected to the rotating beam by a neck, the

**Fig. 38** Simulation relations of the correction factor  $D$  with dimensions of the rotating beam. (a) Length of the rotating beam ( $L_1$ ) and (b) width of the rotating beam ( $W_1$ )



(a)

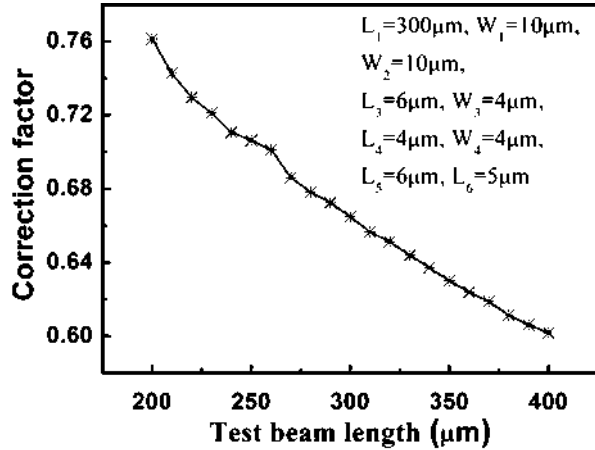


(b)

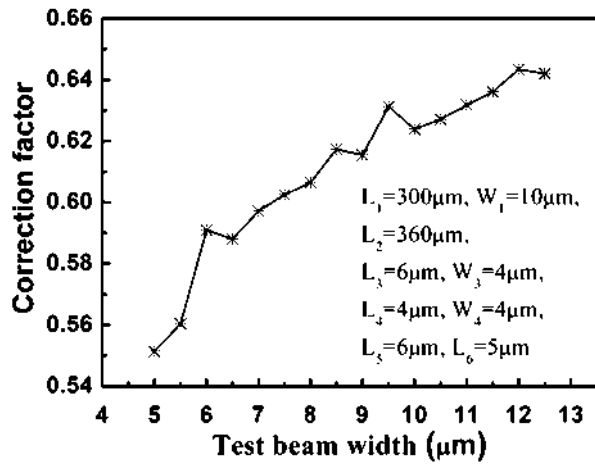
test beam may bend in the direction of beam width. The wider or shorter the test beam is, the less the test beam will bend. This is the reason why the correction factor increases as the width of the test beam ( $W_2$ ) increases or the length of the test beam ( $L_2$ ) decreases. As shown in Fig. 8, for a fixed  $L_3$  of  $6\mu\text{m}$  and a fixed  $W_4$  of  $4\mu\text{m}$ ,  $D$  varies from 0.457 to 0.860 as  $L_5$  changes from  $4\mu\text{m}$  to  $10.5\mu\text{m}$ . The separation between the horizontal center line of the test beam and the rotation center can be expressed as  $L_3 + L_5 - W_4/2$ . As mentioned above, the test beam partially transfers the thermal stress to the rotating beam. Stress remaining on the test beam decreases as the separation increases. Therefore, increasing  $L_5$  makes the correction factor nearer to 1.

In order to apply the FEM results conveniently, a series of simulation data of the correction factor are listed in Table 5, where all the values were calculated for

**Fig. 39** Simulation relations of the correction factor  $D$  with dimensions of the test beam. (a) Length of the test beam ( $L_2$ ) and (b) width of the test beam ( $W_2$ )



(a)



(b)

$L_3 = 6 \mu\text{m}$ ,  $W_3 = 4 \mu\text{m}$ ,  $L_4 = 4 \mu\text{m}$ ,  $W_4 = 4 \mu\text{m}$ , and  $L_6 = 5 \mu\text{m}$ . It should be pointed out that the correction factor given in Table 5 is not valid at high temperature levels because the microstructure shows nonlinear behavior.

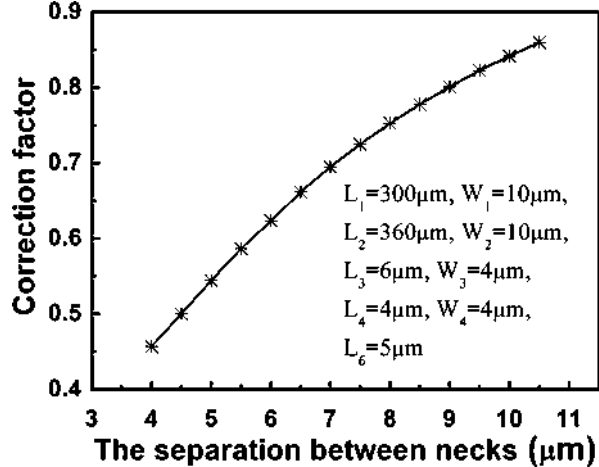
## Experiment

### H-Shaped Test Microstructure

#### Structure Design

In order to provide guidelines for the test microstructure design, a set of simulations with different dimensions and heating currents are performed via ANSYS. The

**Fig. 40** Simulation relations of the correction factor  $D$  with the separation between necks ( $L_5$ )



connecting part between test beams is designed to reduce strain distribution after deformation and produce larger measurement range compared with conventional micro-rotating structures (Zhang et al. 1998; Liu et al. 2004, 2006), while the connecting part between the test beam and the anchor decreases the heat loss from the test beam to the anchor. Figure 41 is the average temperature of the test beam changing with the input heating power. It is observed that with the existence of beam-anchor connecting part, temperature of the test beam is generally higher than temperature of the test beam which directly attached to the anchor. The size of the beam-anchor connecting part, on the other hand, does not significantly affect the average temperature of the test beam at same heating power. Since the thermal expansion mainly occurs along the test beams and the rotating beam acts as an important role of magnifying the expansion, we focus on the geometry of the test beams and the rotating beam for the structure design. For convenience,  $L_3$  and  $W_3$  are both set to be  $20\ \mu\text{m}$ ,  $L_4$  and  $W_4$  are both set to be  $50\ \mu\text{m}$ , and  $L_5$  is set to be  $90\ \mu\text{m}$ . The relationships between the beam size and the tip displacement of the rotating beam are shown in Fig. 42. It is observed that the tip displacement is not significantly influenced by the change of the width of the rotating beam when applying the same heating current, while larger tip displacement occurs by increasing the length of the test beams, or increasing the length of the rotating beam, or decreasing the width of the test beams at the same heating current. As a result, double-test-beam micro-rotating structure with a long rotating beam and long and narrow test beams is utilized to obtain large tip displacement.

To avoid overheating or destroying the test microstructure, the temperature of the test beam should be limited, that is, the applied current should be limited properly. The relationship between the applied current and the maximum temperature of the test beam is illustrated in Fig. 43. According to the electrothermal theory demonstrated in sections “[Online Test Microstructure of Thermal Conductivity](#)” and “[Online Test Microstructure of Thermal Conductivity and Thermal Diffusivity](#),”

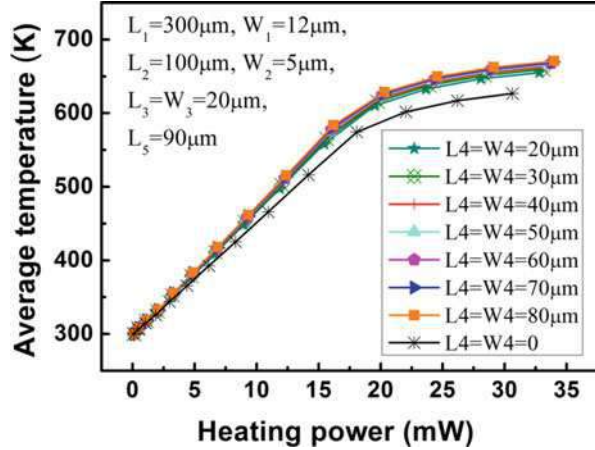
**Table 5** Correction factor  $D$  with different beam sizes ( $L_3 = 6 \mu\text{m}$ ,  $W_3 = 4 \mu\text{m}$ ,  $L_4 = 4 \mu\text{m}$ ,  $W_4 = 4 \mu\text{m}$ ,  $L_6 = 5 \mu\text{m}$ ) (Liu et al. 2013)

No.	$L_1(\mu\text{m})$	$W_1(\mu\text{m})$	$L_2(\mu\text{m})$	$W_2(\mu\text{m})$	$L_5(\mu\text{m})$	$D$
1	260	10	360	10	6	0.6235
2	280	10	360	10	6	0.6237
3	300	10	360	10	6	0.6239
4	320	10	360	10	6	0.6240
5	340	10	360	10	6	0.6242
6	360	10	360	10	6	0.6243
7	300	10	300	10	6	0.6649
8	300	10	320	10	6	0.6512
9	300	10	340	10	6	0.6371
10	300	10	360	10	6	0.6239
11	300	10	380	10	6	0.6113
12	300	10	400	10	6	0.6018
13	300	7	360	10	6	0.6242
14	300	8	360	10	6	0.6236
15	300	9	360	10	6	0.6259
16	300	10	360	10	6	0.6239
17	300	11	360	10	6	0.6263
18	300	12	360	10	6	0.6254
19	300	10	360	7	6	0.5972
20	300	10	360	8	6	0.6065
21	300	10	360	9	6	0.6154
22	300	10	360	10	6	0.6239
23	300	10	360	11	6	0.6318
24	300	10	360	12	6	0.6434
25	300	10	360	10	4	0.4568
26	300	10	360	10	5	0.5442
27	300	10	360	10	6	0.6239
28	300	10	360	10	7	0.6949
29	300	10	360	10	8	0.7528
30	300	10	360	10	9	0.8009

$L_1$  is varied from No. 1 to No. 6,  $L_2$  is varied from No. 7 to No. 12,  $W_1$  is varied from No. 13 to No. 18,  $W_2$  is varied from No. 19 to No. 24,  $L_5$  is varied from No. 25 to No. 30

the temperature is related to the square of the applied current density. Thus, the curve in Fig. 43 is not linear. The dimensions of the test microstructures in Fig. 43 are listed in Table 4. Both the maximum temperature of the test beam and the tip displacement of the rotating beam should be considered for the design of the test microstructure. For example, a double-test-beam micro-rotating structure with a 330  $\mu\text{m}$  long, 12  $\mu\text{m}$  wide rotating beam and two 100  $\mu\text{m}$  long, 5  $\mu\text{m}$  wide test beams produces a 2.24  $\mu\text{m}$  tip displacement at a heating current of 4.5 mA. The maximum temperature of the test beam is limited at 545 K at the same heating current, which is reasonable for structure design.

**Fig. 41** Simulation results: relationship between the average temperature of the test beam and the input heating power



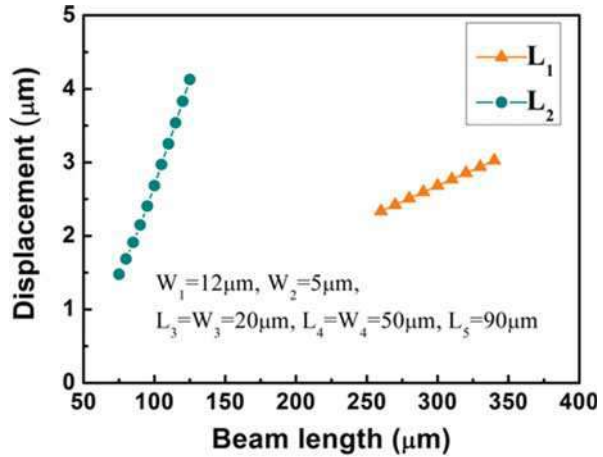
### Experiment Results

The SEM photographs of the surface micromachined H-shaped test microstructures are shown in Fig. 44. The measurement method is presented as follows. First, the resistance of the microbridge and the resistance between pad A and pad B are measured at atmospheric temperature. Then, a current sweep is applied between pad A and pad B, while the resistance between pad B and pad C is monitored. The contact is detected as a sudden step in the measured resistance. When the contact occurs, the resistance between pad A and pad B is recorded. As shown in Fig. 45, a semiconductor characterization system and a probe station were utilized to perform experiments mentioned above. When the average temperature of the test beam is calculated from Eq. 54 by using the experiment results, the CTE for the polysilicon thin films is then extracted from Eq. 52. For the convenience of structure design and CTE measurements, six H-shaped test microstructures with different sizes were fabricated. Experimental results of different specimens are shown in Fig. 46 and Table 6. The average measured value of the CTE of the polysilicon thin film is  $(2.76 \pm 0.09) \times 10^{-6} \text{ K}^{-1}$  with temperature ranging from 450 to 500 K.

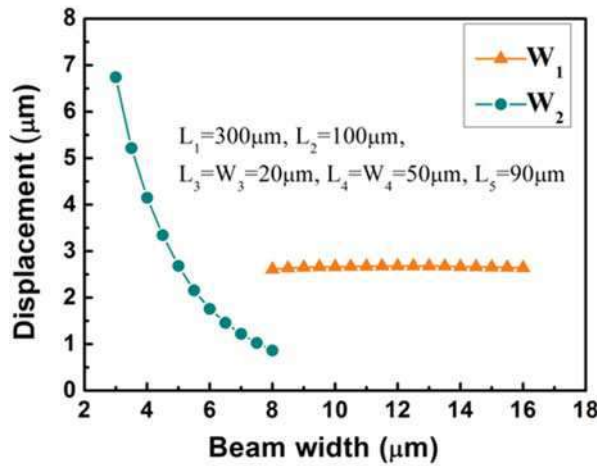
Some factors may influence the results and limit the applications of the test microstructure. First, one potential source of error in calculating the CTE is the geometrical deviation of the test microstructure. A first-order error analysis can be carried out by examining Eqs. 50 and 52. The errors for the tip displacement and the CTE induced by dimension errors are both less than 5%. Second, the deflection induced by residual stress is not considered in this test, which may lead to a significant error if the residual stress is large enough. However, it is observed from the SEM photographs that the residual stress in our experiment is too small to read from the vernier gauge (as shown in Fig. 47). So it is valid to assume that the initial gap between the rotating beam tip and the contact electrode remains constant after the structure is detached from the underlying sacrificial layer. Third, since the



**Fig. 42** Simulation results: relationship between the beam size and the tip displacement of the rotating beam (the heating current is 5 mA).  
 (a) Beam length and  
 (b) beam width



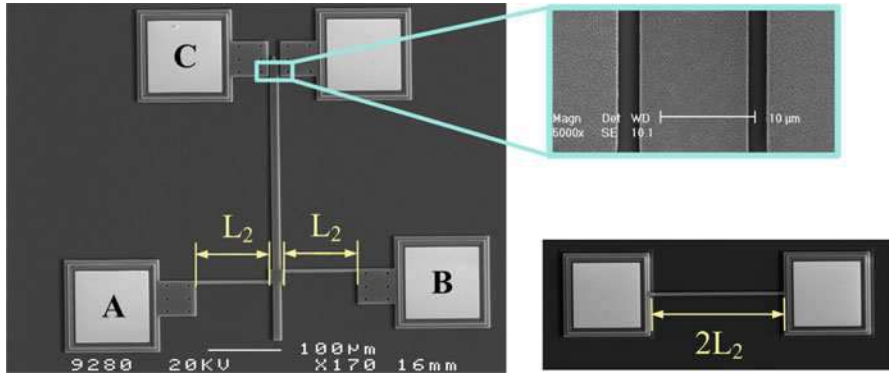
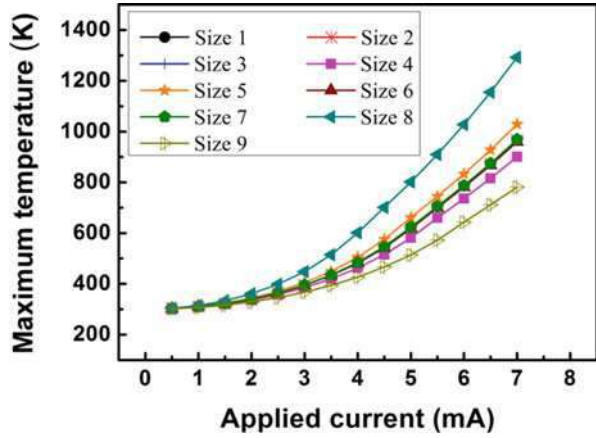
(a)



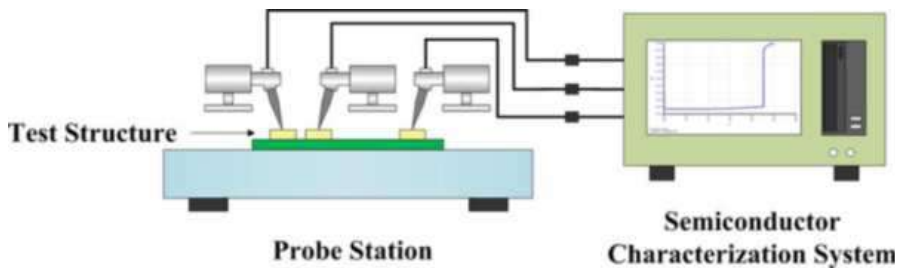
(b)

double-test-beam micro-rotating structure is able to achieve a tip displacement larger than or smaller than  $2 \mu\text{m}$ , the CTE for polysilicon thin films at higher or lower temperature can be extracted by increasing or decreasing the gap between the rotating beam and the contact electrode. However, at high temperatures, Eq. 57 is no longer valid because we verified that the correction factor is independent of temperature only from room temperature up to 680 K, while on the other hand, the gap cannot be designed less than  $2 \mu\text{m}$  in our experiment due to the limit of the fabrication technology. Moreover, a non-negligible buckling effect could occur along the test beam due to a large thermal stress at high temperatures. The narrow test beam may even be overheated and destroyed. Thus, it is estimated that the applicable range of the temperature is approximately from 450 to 600 K.

**Fig. 43** Simulation results: relationship between the applied current and the maximum temperature of the test beam

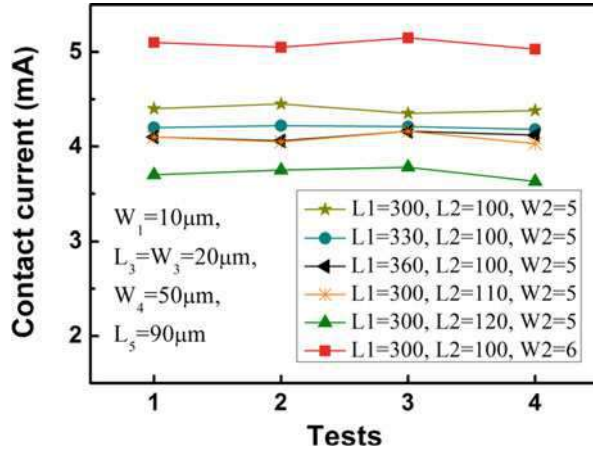


**Fig. 44** SEM photographs of the H-shaped test microstructure for the on-line test of CTE



**Fig. 45** The measurement setup for the on-line test of CTE by using rotating technique

**Fig. 46** Experimental results of contact current measured with different beam sizes (the contact current is the applied current when the contact occurs)



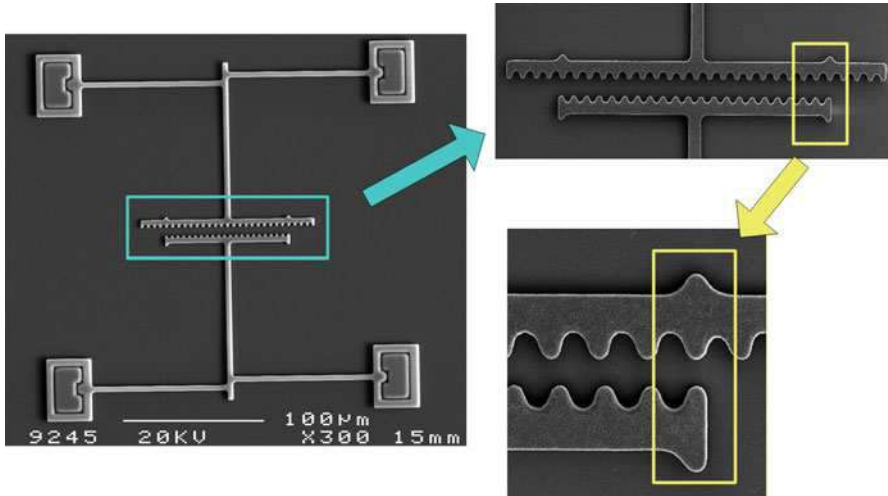
**Table 6** The results of measurements ( $L_3 = 20 \mu\text{m}$ ,  $W_3 = 20 \mu\text{m}$ ,  $L_4 = 50 \mu\text{m}$ ,  $W_4 = 50 \mu\text{m}$ ,  $L_5 = 90 \mu\text{m}$ ) (Liu et al. 2014)

$L_1$ ( $\mu\text{m}$ )	$L_2$ ( $\mu\text{m}$ )	$W_1$ ( $\mu\text{m}$ )	$W_2$ ( $\mu\text{m}$ )	$I_{\text{contact}}$ (mA)	$T_{\text{avr}}$ (K)	$\alpha$ ( $\times 10^{-6} \text{K}^{-1}$ )
300	100	10	5	4.4	494.04	2.86
330	100	10	5	4.2	487.89	2.70
360	100	10	5	4.1	463.69	2.84
300	110	10	5	4.1	487.26	2.70
300	120	10	5	3.7	475.14	2.64
300	100	10	6	5.1	506.16	2.79

## II-Shaped Test Microstructure

### Structure Design

A set of simulations with different dimensions and heating currents were performed by ANSYS to provide guidelines for structure design. For convenience, we focus on the geometry of the test beam and the rotating beam, and  $L_3$  and  $W_3$  are set to be 6 and 4  $\mu\text{m}$ , respectively,  $L_4$  and  $W_4$  are both set to be 4  $\mu\text{m}$ , and  $L_6$  is set to be 5  $\mu\text{m}$ . The relationships between the beam size and the tip displacement of the rotating beam are illustrated in Fig. 48. It is observed that the tip displacement of the rotating beam is significantly influenced by the length of the rotating beam and the dimension of the test beam. In order to achieve a large tip displacement, the test microstructure should have a long rotating beam and a long and narrow test beam. The effect of  $L_5$  on the tip displacement is more complex. The decrease of  $L_5$  raises the displacement magnification of the rotating beam, while the test beam transfers more stress to the rotating beam with a larger  $L_5$ . In Fig. 48, for  $4 \mu\text{m} \leq L_5 \leq 7 \mu\text{m}$ , the effect of  $L_5$  on stress transfer is greater than that on displacement magnification, whereas for  $7 \mu\text{m} \leq L_5 \leq 10.5 \mu\text{m}$ , it is the other way around. Thus, 7  $\mu\text{m}$  is the optimal value



**Fig. 47** SEM photographs of a structure for measuring the residual stress (*the structure is fabricated along with the test microstructure for the CTE on the same die*)

of L5 for  $L_1 = 300 \mu\text{m}$ ,  $W_1 = 10 \mu\text{m}$ ,  $L_2 = 360 \mu\text{m}$ ,  $W_2 = 10 \mu\text{m}$ ,  $L_3 = 6 \mu\text{m}$ ,  $W_3 = 4 \mu\text{m}$ ,  $L_4 = 4 \mu\text{m}$ ,  $W_4 = 4 \mu\text{m}$ , and  $L_6 = 5 \mu\text{m}$ .

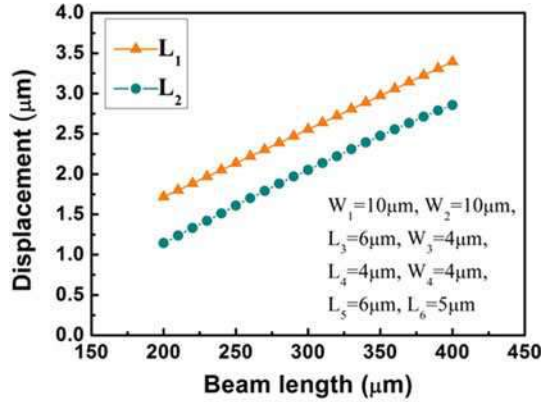
The maximum temperature of the test beam should be limited to avoid destruction of the test microstructure. Thus the applied current should be limited properly. The relationship between the applied current and the maximum temperature of the test beam is displayed in Fig. 49, and the dimensions used in Fig. 49 are listed in Table 4.

Both the maximum temperature of the test beam and the tip displacement of the rotating beam should be considered for the design of the test microstructure. For instance, a single-test-beam micro-rotating structure with a  $300 \mu\text{m}$  long,  $10 \mu\text{m}$  wide rotating beam and a  $300 \mu\text{m}$  long,  $10 \mu\text{m}$  wide test beam produces a  $2.84 \mu\text{m}$  tip displacement at a heating current of  $4 \text{ mA}$ . The maximum temperature of the test beam is limited at  $540 \text{ K}$  at the same heating current, which is reasonable for structure design.

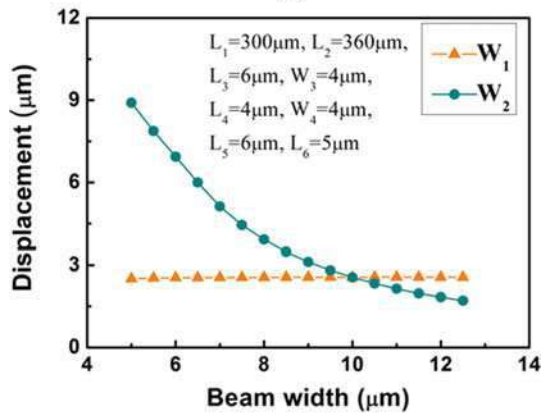
### Experiment Results

The SEM photographs of the surface micromachined  $\Pi$ -shaped test microstructure are shown in Fig. 50. The measurement method is presented as follows. First, the resistance between pad A and pad B is measured at atmospheric temperature. Then, a current sweep is applied between pad A and pad B, while the resistance between pad B and pad C is monitored. The contact is detected as a sudden step in the measured resistance. When the contact occurs, the resistance between pad A and pad B is recorded. The experiment apparatus is the same as that shown in Fig. 45. When the average temperature of the effective test beam is calculated from Eq. 56 by using the experiment results, the CTE of the polysilicon thin film is then extracted from Eq. 55.

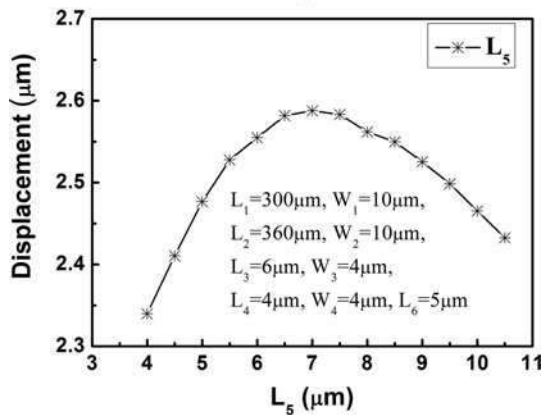
**Fig. 48** Simulation results: relationship between the beam size and the tip displacement of the rotating beam (the heating current is 3.5 mA). (a)  $L_1$  and  $L_2$ , (b)  $W_1$  and  $W_2$ , and (c)  $L_5$



(a)

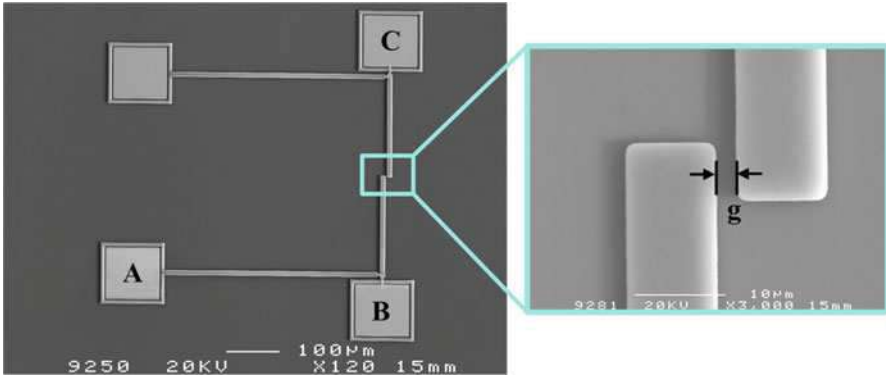
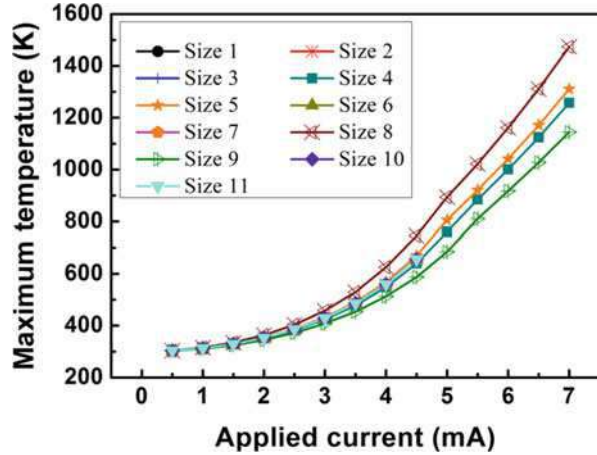


(b)



(c)

**Fig. 49** Simulation results: relationship between the applied current and the maximum temperature of the test beam

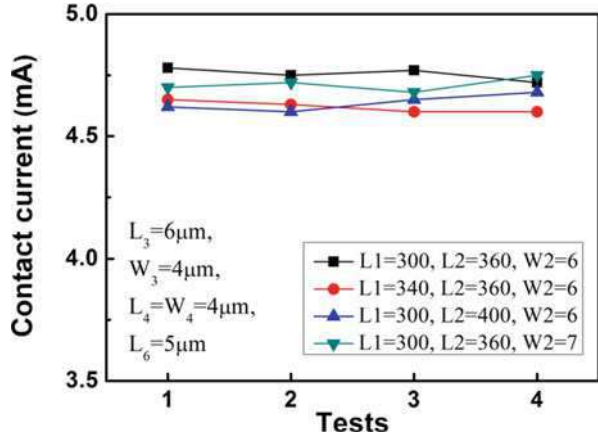


**Fig. 50** SEM photographs of the  $\Pi$ -shaped test microstructure for the on-line test of CTE

Experimental results of different specimens are shown in Fig. 51 and Table 7. The average measured value of thermal expansion coefficient of the polysilicon thin film is  $(2.61 \pm 0.04) \times 10^{-6} \text{ K}^{-1}$  with temperature ranging from 400 to 420 K.

Some factors may influence the results and limit the applications of the test microstructure. First, the error for the CTE induced by dimension errors is less than 5.0%. Second, the correction factor is no longer independent of temperature if the average temperature of the effective test beam is higher than 850 K. Besides, a considerable buckling effect would occur along the test beam at high temperatures. The narrow test beam may even be overheated and destroyed. Thus, it is estimated that the applicable range of the temperature for TEC measurement is approximately from 400 to 650 K. Accordingly, the gap between the tips of the rotating beams should be carefully designed to coordinate with the applicable experiment temperature.

**Fig. 51** Experimental results of contact current measured with different beam sizes (*the contact current is the applied current when the contact occurs*)



**Table 7** The results of measurements ( $L_3 = 6 \mu\text{m}$ ,  $W_3 = 4 \mu\text{m}$ ,  $L_4 = 4 \mu\text{m}$ ,  $W_4 = 4 \mu\text{m}$ ,  $L_6 = 5 \mu\text{m}$ ) (Liu et al. 2013)

$L_1$ ( $\mu\text{m}$ )	$W_1$ ( $\mu\text{m}$ )	$L_2$ ( $\mu\text{m}$ )	$W_2$ ( $\mu\text{m}$ )	$L_5$ ( $\mu\text{m}$ )	$T_1$ (K)	$\alpha$ ( $\times 10^{-6}\text{K}^{-1}$ )
300	10	360	10	6	414.06	2.59
340	10	360	10	6	400.05	2.61
300	10	400	10	6	403.55	2.66
300	10	360	10	7	413.05	2.58

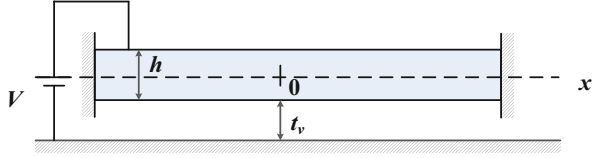
## Online Test Microstructure of the Coefficient of Thermal Expansion by a Pull-In Approach

Traditional techniques to measure the CTE of the thin film generally made use of the displacement or deflection from the test microstructure measured by optical techniques. It is inconvenient to use those techniques in the process line. In this section, a novel electrical test microstructure is proposed for online determination of the CTE of polysilicon thin film. Instead of using displacement or deflection, the electrothermal analysis of the test microstructure is utilized, and the pull-in method is extended to the characterization of the CTE of the thin film. In this structure, the polysilicon beams are heated by direct current, and the transient electrical resistance and the pull-in voltage are measured. The test microstructure is fabricated by surface micro-machined devices without additional technology, and it is characterized under the natural operation environment.

### Pull-In Analysis of a Double-Clamped Beam

Electrostatic pull-in is a well-known sharp instability in the behavior of a voltage-controlled parallel-plate electrostatic actuator (Osterberg and Senturia 1997). For a

**Fig. 52** Schematic diagram of the deformation of a double-clamped beam under electrostatic force



double-clamped beam shown in Fig. 52, the beam is assumed to be a polysilicon beam. Length of the beam is  $l$ , width is  $b$ , thickness is  $h$ , Young’s modulus is  $E$ , Poisson ratio is  $\nu$ , residual stress along length of the beam is  $\sigma$ , and  $t_v$  is the initial gap between the beam and the fixed polysilicon plate used as the bottom electrode. When a DC voltage  $V$  is applied between the beam and bottom electrode, the beam will bend due to the attractive electrostatic force.

When the beam bends due to the electrostatic force, a trial function in the form of a cosine is selected (Senturia 2001):

$$w(x) = \frac{c}{2} \left( 1 + \cos \frac{2\pi x}{l} \right) \tag{58}$$

where  $c$  is an unknown parameter to be determined. It is actually the central deflection of the beam.

Since the two ends of the beam are both clamped, axial stretch is yielded after the beam bent. The total strain  $\varepsilon_T$  must include bending strain  $\varepsilon_b$ , stretching strain  $\varepsilon_s$ , and the strain  $\sigma/\tilde{E}$  due to the residual stress  $\sigma$ :

$$\varepsilon_T = \varepsilon_b + \varepsilon_s + \frac{\sigma}{\tilde{E}} = -z \frac{d^2 w}{dx^2} + \frac{1}{l} \int_{-\frac{l}{2}}^{\frac{l}{2}} \frac{1}{2} \left( \frac{dw}{dx} \right)^2 dx + \frac{\sigma}{\tilde{E}} \tag{59}$$

The total elastic energy is given by (Zou et al. 1995)

$$\begin{aligned} W_{\text{elastic}} &= \frac{\tilde{E}bh}{2} \int_{-\frac{l}{2}}^{\frac{l}{2}} \varepsilon_T^2 dx \\ &= c^2 \left( \frac{\tilde{E}b}{3} \cdot \frac{\pi^4}{l^3} \cdot \frac{h^3}{4} + \frac{\pi^2}{4l} b\sigma h \right) + c^4 \cdot \frac{\tilde{E}b\pi^4}{32l^3} h + \frac{b}{2} \cdot \frac{\sigma^2}{\tilde{E}} h \end{aligned} \tag{60}$$

where  $\tilde{E}$  denotes the effective Young’s Modulus. For a narrow double-clamped beam ( $b < 5h$ ),  $\tilde{E} = E$ , while for a wide beam ( $b \geq 5h$ ),  $\tilde{E} = E/(1 - \nu^2)$ .

Neglect fringing field effect; the capacitance formed by the part of the beam between  $x \sim x + dx$  and the bottom electrode is

$$dC = \frac{b}{\frac{t_v - w(x)}{\varepsilon_0 dx}} = \frac{b\varepsilon_0 dx}{t_v - w(x)} \tag{61}$$



where  $\varepsilon_0$  is the permittivity of vacuum.

The electrostatic co-energy can be written as (Gabbay 1998)

$$\begin{aligned} W_e^* &= \frac{V^2}{2} \varepsilon_0 b \int_{-\frac{l}{2}}^{\frac{l}{2}} \frac{dx}{t_v - w(x)} \\ &= \frac{V^2}{2} \varepsilon_0 b \frac{l}{\sqrt{t_v(t_v - c)}} \end{aligned} \quad (62)$$

The total energy  $W$  of the system includes elastic energy  $W_{\text{elastic}}$  and electrostatic energy  $W_e$ , that is,

$$W = W_{\text{elastic}} + W_e \quad (63)$$

According to the principle of virtual work, we can obtain that

$$\frac{\partial W}{\partial c} = \frac{\partial W_{\text{elastic}}}{\partial c} + \frac{\partial W_e}{\partial c} = 0 \quad (64)$$

For energy domains using co-energy, we shall make substitution

$$\frac{\partial W_e}{\partial c} = - \frac{\partial W_e^*}{\partial c} \quad (65)$$

Therefore,

$$\frac{\partial W}{\partial c} = \frac{\partial W_{\text{elastic}}}{\partial c} - \frac{\partial W_e^*}{\partial c} = 0 \quad (66)$$

Substituting Eqs. 61 and 62 into Eq. 66 yields (Rong et al. 2004)

$$Ac + Bc^3 = \frac{V^2}{4} \cdot \frac{\varepsilon_0}{\sqrt{t_v} \cdot (t_v - c)^{3/2}} \quad (67)$$

with

$$A = 2 \left[ \frac{\tilde{E}}{3} \cdot \left( \frac{\pi}{l} \right)^4 \cdot \frac{h^3}{4} + \frac{\pi^2}{4l^2} \cdot \sigma h \right] \quad (68)$$

$$B = \frac{\pi^4}{8l^4} \tilde{E} h \quad (69)$$

The beam will bend further with increasing applied voltage. There is a critical voltage at which the stability of the equilibrium will be lost. This is called the pull-in voltage, denoted  $V_{PI}$ . At pull-in, Eq. 67 can be written as

$$Ac_{PI} + Bc_{PI}^3 = \frac{V_{PI}^2}{4} \cdot \frac{\varepsilon_0}{\sqrt{t_v} \cdot (t_v - c_{PI})^{3/2}} \quad (70)$$

where  $c_{PI}$  is pull-in displacement of the center of the beam. As pull-in occurs, the following equation should also be satisfied besides (6.70)

$$\frac{\partial^2 W}{\partial c^2} = 0 \quad (71)$$

that is,

$$A + Bc_{PI}^2 = \frac{V_{PI}^2}{4} \cdot \frac{3}{2} \frac{\varepsilon_0}{\sqrt{t_v} \cdot (t_v - c_{PI})^{5/2}} \quad (72)$$

Combining Eqs. 70 and 72,  $V_{PI}$  and  $c_{PI}$  can be calculated as (Rong et al. 2004)

$$V_{PI} = \sqrt{\frac{4(Ac_{PI} + Bc_{PI}^3)\sqrt{t_v}(t_v - c_{PI})^{3/2}}{\varepsilon_0}} \quad (73)$$

$$c_{PI} = \frac{t_v}{9} \left( 36 \times \frac{A}{Bt_v^2} + 8 + D \right)^{1/3} - \frac{\left( 15 \times \frac{A}{Bt_v^2} - 4 \right) t_v}{9 \left( 36 \times \frac{A}{Bt_v^2} + 8 + D \right)^{1/3}} + \frac{2}{9} t_v \quad (74)$$

with

$$D = 3 \sqrt{375 \times \left( \frac{A}{Bt_v^2} \right)^3 - 156 \times \left( \frac{A}{Bt_v^2} \right)^2 + 144 \times \frac{A}{Bt_v^2}} \quad (75)$$

The analytical expression in Eq. 73 can be modified by considering fringing effect and additional fitting parameters, in which the additional fitting parameters are obtained by the least-square method. The final analytical model for pull-in voltage of the double-clamped beam is derived as (Nie et al. 2005, 2009)

$$V_{PI} = \eta \sqrt{\frac{4(Ac_{PI} + Bc_{PI}^3)\sqrt{t_v}(t_v - c_{PI})^{3/2}}{\varepsilon_0 \left( 1 + 0.42 \times \frac{t_v}{b} \right)}} \quad (76)$$

with

$$\eta = \frac{1}{1.1454(1 - \eta_2 e^{\eta_3})} \quad (77)$$

$$\eta_1 = \frac{1.0874 \times 10^{-2} \cdot t_v}{h} \tag{78}$$

$$\eta_2 = 1 - 0.8894 \cdot e^{\eta_1} \tag{79}$$

$$\eta_3 = -\frac{2.1064\sigma^{0.623}}{\tilde{E}h} 1.9252 \times 10^3 (l - 10^{-4} + 0.06724) \tag{80}$$

### Test Microstructure and Theoretical Model

The test microstructure consists of two double-clamped polysilicon beams and two bottom electrodes for each beam, respectively (Fig. 53) (Liu et al. 2012b). The bottom electrode is a polysilicon plate, the thickness of which is  $t_s$ . The length of the beams is  $l_1$  and  $l_2$ , respectively; the width and thickness of the beams are both  $w$  and  $h$ , respectively. The gap between the beam and the bottom electrode is  $t_v$ .

Since two ends of the beam are clamped, the strain induced by thermal expansion is offset due to axial strain when the beam temperature changes. This relationship can be described as

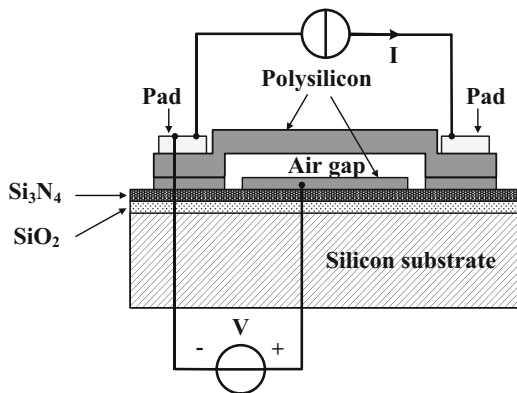
$$\sigma_{th} = -\tilde{E}\alpha_T\Delta\bar{T} \tag{81}$$

where  $\alpha_T$  is the CTE of the polysilicon beam and  $\sigma_{th}$  is the stress of the beam at the change of the average temperature  $\Delta\bar{T}$ .

The CTE of the beam can be expressed as

$$\alpha_T = -\frac{\sigma_{th}}{\tilde{E}\Delta\bar{T}} \tag{82}$$

**Fig. 53** Schematic diagram of a test microstructure for CTE determination



In order to determine the CTE, thermal stress, effective Young's modulus, and average temperature change of the beam are required to be obtained.

The double-clamped beam is heated by applying direct current between two anchors. As discussed in section "Online Test Microstructure of Thermal Conductivity and Thermal Diffusivity," the energy balance of an element of length  $dx$  in the double-clamped beam can be described as Eq. 24. It is noted that Eq. 7 is modified to be  $R_T = t_v/k_v + t_n/k_n + t_s k_s + t_o/k_o$  due to the existence of the bottom electrode, in which  $k_s$  is the thermal conductivity of the bottom electrode. The average temperature change of the double-clamped beam ( $\Delta\bar{T}$ ) can be obtained from Eq. 37, where  $\gamma$  and  $k_p$  are calculated from Eqs. 44 and 45, respectively.

As stated above, the polysilicon beam is heated by Joule heat, causing it to expand, while the thermal expansion of the beam is restricted due to the clamped-clamped boundary, resulting in the increase of stress along the beams. The measurement of effective Young's modulus and thermal stress requires the pull-in method. As demonstrated in Eqs. 68–69, and 74, 75, 76, 77, 78, 79, 80, pull-in voltage is derived as a function of effective Young's modulus, residual stress, and geometry of the test microstructure.

The change of the beam temperature is assumed to have negligible effect on effective Young's modulus. Thermal stress, which contributes to the change of residual stress, varies with the beam temperature. The method to extract the effective Young's modulus and residual stresses of the double-clamped beam is derived (Nie et al. 2009). According to Eq. 76, by measuring pull-in voltages of two beams with the same thickness and width but different length  $l_1$  and  $l_2$ , before and after heating, the following equations can be given:

$$\begin{cases} V_{PI1}(\tilde{E}, \sigma_0, l_1) = V_1 \\ V_{PI2}(\tilde{E}, \sigma_0, l_2) = V_2 \\ V_{PI3}(\tilde{E}, \sigma_1, l_1) = V_3 \\ V_{PI4}(\tilde{E}, \sigma_2, l_2) = V_4 \end{cases} \quad (83)$$

where  $\sigma_0$  is the initial residual stress of the beam before heating and  $\sigma_1$  and  $\sigma_2$  are the residual stress of the beam  $l_1$  and  $l_2$ , respectively, after heating.  $V_1$ ,  $V_2$ ,  $V_3$ , and  $V_4$  are the measured pull-in voltages. Solving these four equations, one obtains the residual stresses and effective Young's modulus before and after heating. Thermal stress is then determined.

Since thermal stress, effective Young's modulus, and average temperature change are obtained, CTE for polysilicon thin films can be extracted from Eq. 82.

## Simulation and Verification

The software ANSYS is used here to perform the electrothermal analysis of the test microstructure, and the software CoventorWare is used to perform the pull-in analysis. The parameters for the simulation are listed in Table 2.

The simulation verification is divided into three stages with respect to the average temperature change, the thermal stress and effective Young's modulus, and the CTE.

### **Average Temperature Change**

First, a steady-state thermal-electric conduction analysis is carried out by ANSYS to confirm Eq. 30. Then a transient-state simulation is conducted to verify Eq. 35. The parameter  $\gamma$  and thermal conductivity are calculated from Eqs. 44 and 46 with the simulation results. The temperature change is then obtained by Eq. 42. Comparing this simulation value of temperature change with its theoretical value calculated from Eqs. 26 and 42 with parameters in Table 2, the deviation is less than 3%. This stage verifies the electrothermal model of the double-clamped beam and the extraction method of average temperature change of the beam. Details of this stage have been discussed in section "[Simulation and Verification](#)."

### **Thermal Stress and Effective Young's Modulus**

For simulation comparison, the thermal stress and Young's modulus of the polysilicon beam are both initially set to be certain values. By substituting pull-in voltages extracted from CoventorWare simulation into Eq. 76, the thermal stress and effective Young's modulus are calculated. The deviation between the simulation value and set value of the thermal stress and Young's modulus are both less than 3%. This stage verifies the pull-in analysis and the extraction method of thermal stress and effective Young's modulus.

### **CTE**

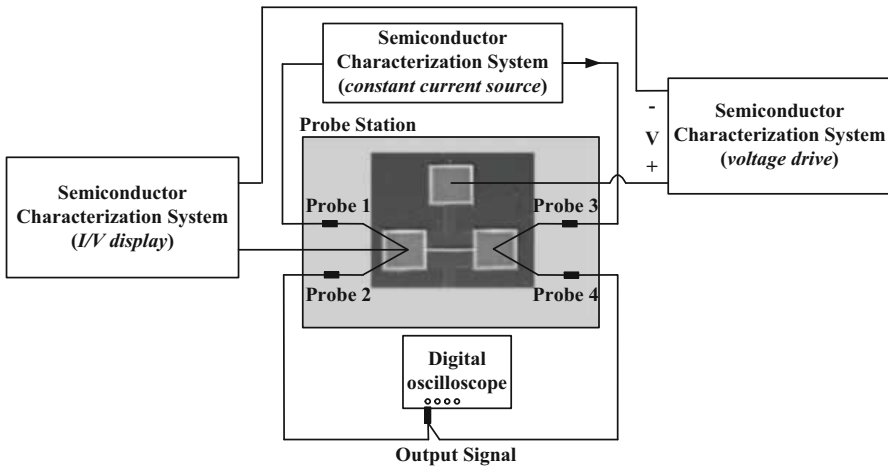
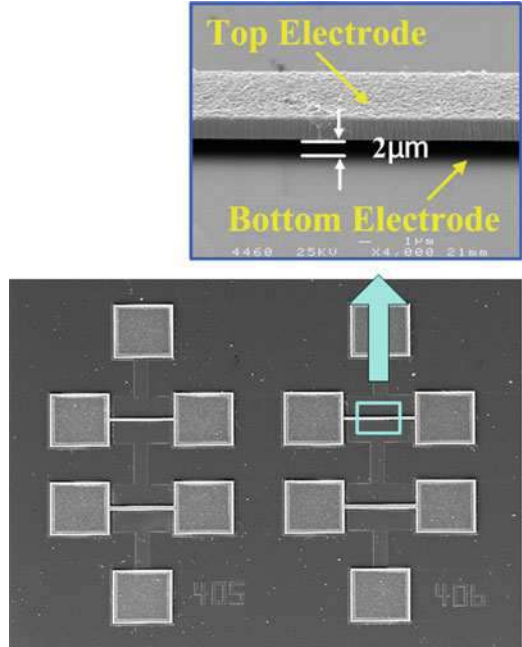
Thermal stress of the double-clamped beam heated by Joule heat can be obtained from steady-state thermal-electric conduction analysis performed by ANSYS. Substituting the simulation value of thermal stress, average temperature change (from stage 1), and set value of Young's modulus (from Table 2) into Eq. 82, CTE of the polysilicon thin film can be extracted. The deviation between the simulation value of CTE and its set value (from Table 2) is less than 3%.

## **Experiment**

The SEM photographs of the test microstructure are shown in Fig. 54. The experimental apparatus, as is shown in Fig. 55, consists of a digital oscilloscope and a semiconductor characterization system. The semiconductor characterization system provides voltage sweep for measuring the pull-in voltages and current bias for generating Joule heat, whereas the digital oscilloscope records electrical resistance characteristics of the test microstructures.

The measurement method is presented as follows. A voltage sweep between the beam and the bottom electrode is applied to measure the pull-in voltage of the beam before the heating process. The pull-in is detected on the I/V screen output interface of the semiconductor characterization system as a sudden step in the current that will increase from zero to several microamperes. Then the voltage sweep stops, and a

**Fig. 54** SEM picture of the double clamped beams for the on-line test of CTE



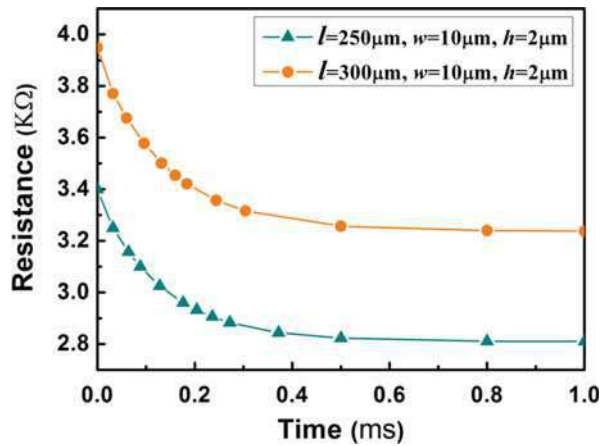
**Fig. 55** The measurement setup for the on-line test of CTE by using pull-in approach

current through the beam is applied to heat it up to thermal steady state. The transient resistance characteristic of the test microstructure during the heating process is captured and stored by the digital oscilloscope. When the beam is at thermal steady state, the pull-in voltage is measured again by the semiconductor characterization system. Performing these measurements on two beams with different lengths, the

**Table 8** The results of measurements (Liu et al. 2012b)

$I$ (mA)	$l$ ( $\mu\text{m}$ )	$w$ ( $\mu\text{m}$ )	$\Delta\bar{T}$ (K)	$V_{PI1}$ (V)	$V_{PI2}$ (V)	$\tilde{E}$ (GPa)	$\sigma_{th}$ (MPa)	$\alpha_T$ ( $\text{K}^{-1} \times 10^{-6}$ )
1	180	10	23.6	115	108	145	8.6	2.51
	200		24.6	95	88		9.0	2.52
1.5	250	10	52.6	68	55	143	18.5	2.46
	300		54.7	51	38		19.3	2.47
2	250	15	47.7	67	56	138	16.4	2.49
	300		50.7	50	39		17.5	2.50

**Fig. 56** Transient electrical resistance of test microstructures after heating by a current of 1.5 mA

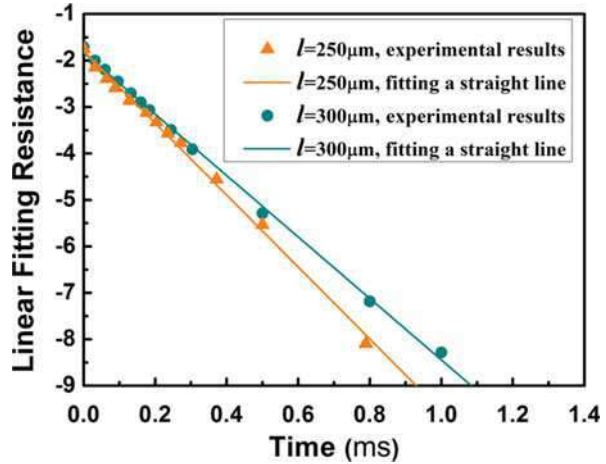


thermal stress and the Young’s modulus can be calculated from the pull-in voltages, while the change of the average temperature can be obtained by resistance characteristics. The CTE for the polysilicon thin films is then extracted.

The results of measurements are listed in Table 8. Three groups of test microstructures with different lengths and widths were tested. The average value of effective Young’s modulus calculated from the pull-in method is  $(142.3 \pm 4.0)$  GPa. The electrical resistance characteristics of two samples are displayed in Fig. 56, and their linear fitting curves are illustrated in Fig. 57. Measured average value of the CTE is  $(2.49 \pm 0.03) \times 10^{-6} \text{ K}^{-1}$  with temperature ranging from 300 to 350 K.

In our extraction model, some approximations and assumptions are made that may influence the accuracy of measurement results. One error source for the CTE is the geometrical deviation of the test microstructure. Main errors for the measurement result from the deviations of the average temperature change and material properties (the stress and effective Young’s modulus) according to Eq. 82. Based on our experiments, the deviation of the average temperature change is 4.6%, and the test error for the stress and Young’s modulus is 5%. When considering the errors caused by dimensions for the beams, the measurement error of the polysilicon CTE is estimated to be less than 15%. Another source of error comes from the applied

**Fig. 57** Linear fitting curves of measured transient electrical resistances



current (or temperature). According to section “[Experiment](#),” the applicable range of the temperature is suggested to be from room temperature up to 400 K. In addition, the air gap between the beam and substrate may vary with temperature due to buckling of the beam during the heating process. However, it is observed that the residual stresses of test microstructures before and after heating are both tensile in our experiment. Thus it is valid to assume that the shape and geometry of the beam remain constant after heating.

## References

- Arx M-V, Paul O, Baltes H (1995) Determination of the heat capacity of CMOS layers for optimal CMOS sensor design. *Sensors Actuators A Phys* 47(1–3):428–431
- Arx M-Von, Paul O, Baltes H (1997) Thermoelectric test structures to measure the heat capacity of CMOS layer sandwiches. In: *Proceedings of TRANSDUCERS’97*, pp 619–22
- Arx M-V, Paul O, Baltes H (1998) Test structures to measure the heat capacity of CMOS layer sandwiches. *IEEE Trans Semicond Manuf* 11(2):217–224
- Arx M-v, Paul O, Baltes H (2000) Process dependent thin film thermal conductivities for thermal CMOS MEMS. *J Microelectromech Syst* 9(1):136–145
- Brand O, Fedder GK (2005) *CMOS mems*. Wiley, Weinheim
- Cahill DG (1990) Thermal conductivity measurement from 30 to 750 K: the  $3\omega$  method. *Rev Sci Instrum* 61(2):802–808
- Chae JH, Lee JY, Kang SW (1999) Measurement of thermal expansion coefficient of poly-Si using microgauge sensors. *Sensors Actuators A Phys* 75(3):222–229
- Cheng C, Tsai M, Fang W (2015) Determining the thermal expansion coefficient of thin films for a CMOS MEMS process using test cantilevers. *J Micromech Microeng* 25(2):025014.1–025014.14
- deCoster J, Lofrano M, Jansen R, Rottenberg X, Severi S, Borremans J, VanderPlas G, Donnay S, Tilmans HAC (2011) A novel test method for simulations measurement of thermal conductivity, CTE, residual stress and Young’s modulus of suspended thin film using a laser Doppler vibrometer. In: *Proceedings of TRANSDUCERS’11*, pp 1701–1704



- Fang W, Tsai HC, Lo CY (1999) Determining thermal expansion coefficients of thin films using micromachined cantilevers. *Sensors Actuators A Phys* 77(1):21–27
- Gabbay LD (1998) Computer aided macromodeling for MEMS. Dissertation, Massachusetts Institute of Technology
- Geisberger AA, Sarkar N, Skidmore GD (2003) Electrothermal properties and modeling of polysilicon microthermal actuators. *J Microelectromech Syst* 12(4):513–523
- Hafizovic S, Paul O (2002) Temperature-dependent thermal conductivities of CMOS layers by micromachined thermal van der Pauw test structures. *Sensors Actuators A Phys* 97–98:246–252
- Holman JP (1997) Heat transfer. McGraw-Hill, New York
- Huang Q-A, Lee NKS (1999) Analysis and design of polysilicon thermal flexure actuator. *J Micromech Microeng* 9(1):64–70
- Huang Q-A, Xu G, Qi L, Li W (2006) A simple method for measuring the thermal diffusivity of surface micromachined polysilicon thin films. *J Micromech Microeng* 16(5):981–985
- Irace A, Sarro PM (1999) Measurement of thermal conductivity and diffusivity of single and multilayer membranes. *Sensors Actuators A Phys* 76(1–3):323–328
- Jain A, Goodson KE (2008) Measurement of the thermal conductivity and heat capacity of freestanding shape memory thin films using the  $3\omega$  method. *J Heat Transf* 130(10):102402.1–102402.7
- Jansen E, Obermeier E (1996) Thermal conductivity measurements on thin films based on micro-mechanical devices. *J Microelectromech Syst* 6(1):118–121
- Leon I, Amador R, Kohlhof K (2004) Evaluation of MUMPS polysilicon structures for thermal flow sensors. *Microelectron Reliab* 44(4):651–655
- Lin L, Chiao M (1996) Electrothermal responses of lineshape microstructures. *Sensors Actuators A Phys* 55(1):35–41
- Liu C (2006) Foundations of MEMS. Prentice-Hall, Upper Saddle River
- Liu Z, Huang Q-A, Li W (2004) Analysis of optimized micro-rotating-structure for MEMS. In: Proceedings of the 7th international conference on solid-state and integrated circuits Technology, pp 1747–1750
- Liu Z, Huang Q-A, Li W (2006) A new micro-rotating structure. *J Phys Conf Ser* 34(34):552–557
- Liu H, Li W, Yuan F, Jiang M, Huang QA (2012a) Micro-rotating structures for determining thermal expansion coefficients of polysilicon thin films. In: Proceedings of 2012 I.E. Sensors, pp 1596–1599
- Liu H, Zhou Z, Li W, Huang Q-A (2012b) An online test structure for the thermal expansion coefficient of surface micromachined polysilicon beams by a pull-in approach. *J Micromech Microeng* 22(5): 055017.1–055017.8
- Liu H, Li W, Zhou Z, Huang Q-A (2013) In situ test structures for the thermal expansion coefficient and residual stress of polysilicon thin films. *J Micromech Microeng* 23(7):075019.1–075019.9
- Liu H, Li W, Zhou Z, Huang Q-A (2014) In-situ determination of the coefficient of thermal expansion of polysilicon thin films using micro-rotating structures. *Thin Solid Films* 552(3):184–191
- Mag-isa AE, Kim S-M, Kim J-H, Lee H-J, Oh C-S (2013) Out-of-plane CTE measurement method for freestanding thin films. *Exp Mech* 53(6):1017–1024
- Mastrangelo CH, Muller RS (1988) Thermal diffusivity of heavily doped low pressure chemical vapor deposited polycrystalline silicon films. *Sensors Mater* 3:133–142
- McConnell AD, Uma S, Goodson KE (2001) Thermal conductivity of doped polysilicon layers. *J Microelectromech Syst* 10(3): 360–369
- Morikawa J, Hashimoto T (1998) Analysis of high-order harmonics of temperature wave for Fourier transform thermal analysis. *Japan. J Appl Phys* 37:1484–1487
- Murarka SP, Retajczyk TF (1983) Effect of phosphorus doping on stress in silicon and polycrystalline silicon. *J Appl Phys* 54(4): 2069–2072
- Nie M, Huang Q-A, Li W, Rong H (2005) An in-situ technique for measuring Young's modulus and residual stress of each layer for multi-layer film. In: Proceedings of TRANSDUCERS '05, pp 836–839

- Nie M, Huang Q-A, Li W (2009) Pull-in characterization of doubly-clamped composite beams. *Sensors Actuators A Phys* 151(2):118–126
- Ogando K, laForgia N, Zárate JJ, Pastoriza H (2012) Design and characterization of a fully compliant out-of-plane thermal actuator. *Sensors Actuators A Phys* 183:95–100
- Osterberg PM, Senturia SD (1997) M-test: a test chip for MEMS material property measurement using electrostatically actuated test structures. *J Microelectromech Syst* 6(2):107–118
- Pan CH (2002) A simple method for determining linear thermal expansion coefficients of thin films. *J Micromech Microeng* 12:548–555
- Paul O, Ruther P, Plattner L, Baltes H (2000) A thermal van der Pauw test structure. *IEEE Trans Semicond Manufact* 13(2):159–166
- Pocratsky RM, deBoer MP (2014) Determination of thin film coefficient of thermal expansion and residual strain from freestanding fixed–fixed beams. *J Vac Sci Technol B* 32(6):062001.1–062001.6
- Retajczyk TF, Sinha AK (1980) Elastic stiffness and thermal expansion coefficient of BN films. *Appl Phys Lett* 36(2):161–163
- Roncaglia A, Cozzani E, Mancarella F, Passini M, Cardinali GC, Severi M (2007) Influence of air heat exchange upon on-chip measurement of thermal conductivity using MEMS test structures. In: *Proceedings of TRANSDUCERS'07*, pp 615–618
- Rong H, Huang Q-A, Nie M, Li W (2004) An analytical model for pull-in voltage of clamped-clamped multilayer beams. *Sensors Actuators A Phys* 116(1):15–21
- Schafer H, Graeger V, Kobs R (1989) Temperature independent pressure sensors using polycrystalline silicon strain gauges. *Sensors Actuators* 17(3–4):521–527
- Senturia SD (2001) *Microsystem design*. Kluwer, New York
- Stojanovic N, Yun J, Washington EBK, Berg JM, Holtz MW, Temkin H (2007) Thin-film thermal conductivity measurement using microelectrothermal test structures and finite-element-model-based data analysis. *J Microelectromech Syst* 16(5):1269–1275
- Tada H, Kumpel AE, Lathrop RE, Slanina JB, Nieva P, Zavracky P, Miaoulis IN, Wong PY (2000) Thermal expansion coefficient of polycrystalline silicon and silicon dioxide thin films at high temperatures. *J Appl Phys* 87(9):4189–4193
- Tai YC, Mastrenghelo CH, Muller RS (1988) Thermal conductivity of heavily doped low-pressure chemical vapor deposited polycrystalline silicon films. *J Appl Phys* 63(5):1442–1447
- Volkein F, Baltes H (1992) A microstructure for measurement of thermal conductivity of polysilicon thin films. *J Microelectromech Syst* 1(4):193–196
- Wang ZD, Zhao XX, Jiang SQ, Lu JJ (2005) Determining thermal expansion coefficient of stressed thin films at low temperature. *Polym Test* 24: 839–843
- Wang Z, Fiorini P, van Hoof C (2009) CMOS-compatible surface-micromachined test structure for determination of thermal conductivity of thin film materials based on Seebeck effect. In: *Proceedings of MEMS 2009*, pp 623–626
- Xu G, Huang Q-A (2006) An online test microstructure for thermal conductivity of surface-micromachined polysilicon thin films. *IEEE Sensors J* 6(2):428–433
- Xu G, Huang Q-A, Jiang Y (2002) A new test structure for measuring thermal conductivity of polysilicon thin films. In: *Proceedings of SPIE (SPIE, shanghai 2002) vol 4928*, pp 267–271
- Zhang X, Zhang T, Zohar Y (1998) Measurements of residual stresses in thin films using micro-rotating structures. *Thin Solid Films* 335(1–2):97–105
- Ziang X, Grigoropoulos CP (1995) Thermal conductivity and diffusivity of free standing silicon nitride thin films. *Rev Sci Instrum* 66(2):1115–1120
- Zou Q, Li Z, Liu L (1995) New methods for measuring mechanical properties of thin films in micromachining: beam pull-in voltage (VPI) method and long beam deflection (LBD) method. *Sensors Actuators A Phys* 48(2):137–143



# Fracture Properties of MEMS/NEMS Thin Films

Jinling Yang and Quan Yuan

## Contents

Overview .....	304
Analytical Model .....	305
Load-Deflection .....	305
Stress Profiles .....	308
Fracture Analysis .....	309
Experimental Apparatus .....	312
Key Research Findings .....	313
LPCVD Silicon Nitride Films .....	313
Silicon Carbide Film .....	314
SiO <sub>2</sub> /Si <sub>3</sub> N <sub>4</sub> Composite Diaphragm .....	316
References .....	320

## Abstract

Thin film mechanical property evaluation has become increasingly important for micro devices. To assure the reliability of the devices and to predict the lifetime, fracture properties of thin films need to be investigated. Microtensile testing is a powerful technique to characterize the mechanical and fracture properties of microscale thin films. But the specimen preparation always causes a difficulty.

---

J. Yang (✉)

Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

State Key Laboratory of Transducer Technology, Shanghai, China

School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing, China

e-mail: [jlyang@semi.ac.cn](mailto:jlyang@semi.ac.cn)

Q. Yuan

Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

State Key Laboratory of Transducer Technology, Shanghai, China

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,  
[https://doi.org/10.1007/978-981-10-5945-2\\_8](https://doi.org/10.1007/978-981-10-5945-2_8)

303

The bulge test, as a relatively simple, fast, and precise method, is extended to the determination of the fracture properties of thin films, such as bending stiffness and prestress of the membrane material, the Young's modulus, and fracture strength of single layer film and bilayer films. An accurate model describing load-deflection response is applied on thin films made of silicon nitride, silicon carbide, and composite diaphragms of silicon nitride grown on top of thermal silicon oxide films. Fracture reference stresses were computed according to the Weibull model for brittle fracture by integrating the membrane stress over the edge, surface, and volume of the samples, corresponding respectively to the assumption of dominant edge, surface, and volume flaws. This method is very efficient and able to quantify fracture parameters of single and multilayer films.

---

**Keywords**

Fracture · Bulge test · Load-deflection · Weibull analysis

---

**Overview**

Thin films play an indispensable role as structural materials in complementary metal-oxide-semiconductor (CMOS) and microelectromechanical systems (MEMS). In order to guarantee the long-term operation of CMOS and MEMS devices according to their specifications, it is highly desirable to know mechanical and fracture properties of the thin films to avoid catastrophic failure due to fracture in the operation of a microdevice. Fracture properties of micromachined components have been determined using, e.g., the microtensile test (Tsuchiya et al. 1998). This is a powerful, but delicate experimental technique. On the other hand, from the experimental point of view, bulge testing is a rather simple alternative to extract mechanical properties of thin films. It has evolved into a successful method to measure the prestress of thin films and their elastic properties such as the plane-strain modulus. By increasing the differential pressure up to the bursting point of square diaphragms, the bulge test has been used to investigate the effect on the average fracture strain of polysilicon films exposed to hydrofluoric solutions and the average fracture strength dependence on the corner curvature radius of anisotropically etched single-crystal silicon diaphragms, (Walker et al. 1991; Pourahmadi et al. 1991).

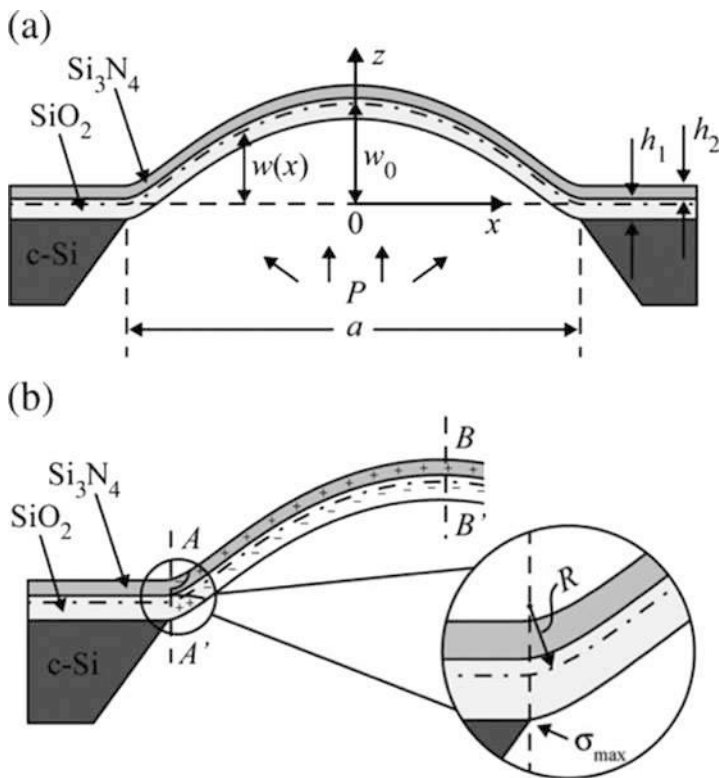
The refined analytical plane-strain load-deflection model was extended to the fracture of the long rectangular membrane and was applied to evaluate the fracture properties of low-pressure chemical-vapor-deposition (LPCVD) silicon nitride films (Yang and Paul 2002). However, the model was proposed under the assumption of a rigidly clamped boundary and restricted to a single-layer thin film with tensile or weakly compressive prestress. The rigidly clamped boundary assumption results in a relatively large error in the local-stress calculation for the membrane structures under investigation. Therefore, with an elastically clamped boundary assumption, the analytical model for the load-deflection response was further refined (Yang et al. 2008) and extended to the composite membranes. The comprehensive analytical model is applicable to the thin films with compressive or tensile stress which enable accurate

computation of stress distribution throughout the composite diaphragm as function of the load, in particular at the critical pressure leading to the fracture of the diaphragms. This model successfully and precisely calculated the fracture stress of thermal silicon oxide thin film, which is difficult to make into a membrane experimentally.

## Analytical Model

### Load-Deflection

The analytical model described here is built up using long rectangular  $\text{Si}_3\text{N}_4/\text{SiO}_2$  composite diaphragms. Figure 1 schematically shows the deflection of such a structure under differential pressure. In view of the horizontal uniformity, the



**Fig. 1** (a) Schematic cross section of silicon oxide/silicon nitride composite diaphragm deflected under differential pressure  $P$ , with center deflection  $w_0$ , width  $a$ , and thicknesses  $h_1$  and  $h_2$  for the  $\text{SiO}_2$  and  $\text{Si}_3\text{N}_4$  layers, respectively. The  $x$ - and  $y$ -axes have their origins in the center of the reference plane: The  $y$ -axis is perpendicular to the drawing plane, and the  $z$ -axis starts from the bottom of the  $\text{SiO}_2$  layer. (b) Enlarged view of the diaphragm edge, with radius of curvature  $R$  and maximum bottom stress  $\sigma_{\max}$ . The symbols “+” and “-” denote tensile and compressive stress, respectively

structure is characterized by properties depending only on the perpendicular coordinate  $z$ : Young's modulus  $E(z)$ , Poisson's ratio  $\nu(z)$ , and the prestress  $\sigma_0(z)$ . The plane  $z = 0$  is defined to coincide with the bottom of the flat  $\text{SiO}_2$  layer. The analytical model described in the following is based on several assumptions. First, the composite diaphragm deforms as a thin plate. This implies that no shear strain involving the  $z$ -direction acts on the material stack, Second, the diaphragms are long, i.e., they have a width  $a$  and length  $b$  with an aspect ratio  $b/a$  larger than 4: 1, Finally, the attachment of the structures to the micromachined substrate is assumed to be well described by an elastic clamping (Yang et al. 2008).

The diaphragm has an overall thickness  $h = h_1 + h_2$ , where  $h_1$  and  $h_2$  denote the thicknesses of the  $\text{SiO}_2$  and  $\text{Si}_3\text{N}_4$  layers, respectively. These two layers have respective elastic moduli  $E_1$  and  $E_2$ , prestress values  $\sigma_{0,1}$  and  $\sigma_{0,2}$ , Poisson's ratios  $\nu_1$  and  $\nu_2$ , and plane-strain moduli  $E_{ps,1} = E_1/(1 - \nu_1^2)$  and  $E_{ps,2} = E_2/(1 - \nu_2^2)$ . Under the influence of an uniform differential pressure  $P$ , the composite diaphragm is deflected into a profile  $w(x)$ . For thin-film diaphragms, bending effects may be neglected throughout most of the diaphragm width.

The mechanical analysis starts with the definition of an appropriate reference plane  $z = z_0$  corresponding to the neutral plane of single-material thin plates. In the case of single-layer diaphragms, this reduces to the usual mid-plane definition  $z_0 = h/2$ . The resistance of the thin structure against plane-strain elongation and plane-strain bending is described by the integral parameters  $D_0$  and  $D_2$ , both defined in Table 1 in general terms and for single, bilayer, and generic multilayer diaphragms. For bilayers,  $D_2$  is similar to the expression proposed for multilayer beams (Soderkvist 1993). The difference is that the elastic modulus is replaced by the plane-strain modulus appropriate for long diaphragms as considered here. Similar to  $D_0$ , the effect of the  $z$ -dependent prestress is condensed into the parameter  $S_0$  also defined in Table 1 as the integral of  $\sigma_0(z)$ . The initial bending moment per unit length  $M_0$  takes into account moments arising in the diaphragm edges due to the different prestress levels of the films in the stack. The supporting torsional stiffness per unit length  $K$  models the finite compliance of the clamping conditions.

Using these parameters, the out-of-plane deformation  $w = w(x)$  of a composite diaphragm under a pressure  $P$  is described by (Paul and Baltes 1999),

$$D_2 \frac{d^4 w}{dx^4} - S \frac{d^2 w}{dx^2} = P \quad (1)$$

where the effective line force  $S$ , as given in (Paul and Baltes 1999), is

$$S = S_0 + \frac{D_0}{2a} \int_{-a/2}^{a/2} \left( \frac{dw}{dx} \right)^2 dx \quad (2)$$

Under deflection, the initial line force  $S_0$  grows to the effective line force  $S$ . This results from the additional strain in the diaphragm due to the elongation of the structure by its out-of-plane deflection. Given an effective line force  $S$ , and under

**Table 1** Integral parameters determining the static deflection of composite long plates under plane-strain deformation

Parameter	General definition	Homogeneous thin plate	Bilayered thin plate	Multilayered thin plate
Reference plane [m]	$z_0$	$\frac{h}{2}$	$\frac{E_{ps,1}h_1^2 + E_{ps,2}h_2^2 + 2E_{ps,2}h_1h_2}{2(E_{ps,1}h_1 + E_{ps,2}h_2)}$	$\sum_{n=1}^N \frac{[E_{ps,n}h_n(\sum_{m=1}^n h_m - h_m/2)]}{\sum_{n=1}^N E_{ps,n}h_n}$
Stretching stiffness [N/m]	$D_0$	$E_{ps}h$	$E_{ps,1}h_1 + E_{ps,2}h_2$	$\sum_{n=1}^N E_{ps,n}h_n$
Bending rigidity [Nm]	$D_2$	$\frac{E_{ps}h^3}{12}$	$E_{ps,1}h_1^3 + E_{ps,2}h_2^3 + \frac{E_{ps,1}^2h_1^4 + E_{ps,2}^2h_2^4}{12(E_{ps,1}h_1 + E_{ps,2}h_2)} + \frac{E_{ps,1}E_{ps,2}h_1h_2(2h_1^2 + 3h_1h_2 + 2h_2^2)}{6(E_{ps,1}h_1 + E_{ps,2}h_2)}$	$\sum_{n=1}^N \left\{ E_{ps,n}h_n \prod_{l=1}^n [\sum_{m=1}^n h_m - h_m/2] \right\} - \left( 1 + i(-1)^l / \sqrt{3} \right) h_n/2 - z_0 \}$
Initial line force [N/m]	$S_0$	$\sigma_0 h$	$\sigma_{0,1}h_1 + \sigma_{0,2}h_2$	$\sum_{n=1}^N \sigma_{0,n}h_n$
Initial bending moment per unit length [N]	$M_0$	0	$\frac{(E_{ps,1}\sigma_{0,2} - E_{ps,2}\sigma_{0,1})h_1h_2(h_1 + h_2)}{2(E_{ps,1}h_1 + E_{ps,2}h_2)}$	$\sum_{n=1}^N [\sigma_{0,n}h_n(\sum_{m=1}^n h_m - h_m/2 - z_0)]$
Torsional stiffness of substrate supports per unit length [N]	$K$	$\frac{E_{ps}h^2}{k_r}$	$\frac{E_{ps}(h_1 + h_2)^2}{k_r}$	$\frac{E_{ps}(\sum_{n=1}^N h_n)^2}{k_r}$

the elastic clamping boundary conditions [7], the solution of Eqs. (1) and (2) is Gaspar et al. (2007)

$$w(x) = \frac{Pa^2}{2S} \left[ \frac{1}{4} - \left( \frac{x}{a} \right)^2 \right] - \left[ \frac{P}{S} \left( D_2 + \frac{Ka}{2} \right) + M_0 \right] \times \frac{\cosh\left(\sqrt{S/D_2}a/2\right) - \cosh\left(\sqrt{S/D_2}x\right)}{\text{Scosh}\left(\sqrt{S/D_2}a/2\right) + K\sqrt{S/D_2}\sinh\left(\sqrt{S/D_2}a/2\right)} \quad (3)$$

where the torsional stiffness of the supporting edges is described as  $K = E_{ps}^* h^2/k_r$  with  $k_r = 0.78$  for diaphragms clamped to a micromachined Si substrate (Gerlach et al. 1996) and  $E_{ps}^*$  is the plane-strain modulus of the silicon substrate (168 GPa). By using this solution, the integration in Eq. (2) is carried out explicitly and Eq. (2) is rewritten as Eq. (4) (Gaspar et al. 2007)

$$S = S_0 + \frac{D_2}{a^2} \frac{\sum_{i=1}^3 \left[ \beta_{i,1} + \beta_{i,2} \cosh\left(\sqrt{S/D_2}a\right) + \beta_{i,3} \sinh\left(\sqrt{S/D_2}a\right) \right] \left[ Pa^4 / \left( D_2 \sqrt{D_2/D_0} \right) \right]^{i-1}}{\left[ (Sa^2/D_2) \cosh\left(\sqrt{S/D_2}a/2\right) + (Ka^2/D_2) \sqrt{S/D_2} \sinh\left(\sqrt{S/D_2}a/2\right) \right]^2} \quad (4)$$

Equation (4) is easily solved for  $S$  by numerical iteration with the starting value  $S^{(0)} = S_0 + D_0 \pi^2 w_0^2 / 4a^2$  (Landau and Lifshitz 1986), where  $w_0 = w(0)$  denotes the center deflection, and provided that this starting value is larger than the instability buckling effective line force, as analytically modeled elsewhere (Gaspar et al. 2007). Equations (3) and (4) constitute the load-deflection law of a composite prestressed long clamped diaphragm. The load-deflection response is nonlinear.

In practical terms, load-deflection curves for the center of diaphragms are calculated by obtaining series of  $\{w_0, P\}$  pairs. First,  $z_0, D_0, D_2, S_0, K,$  and  $M_0$  are computed from the properties of the thin-film materials and substrate, and thicknesses. Then, for each  $w_0$ , the corresponding  $P$  is computed by first finding the effective line stress  $S$  using Eq. (4), and then calculating  $P$  using Eq. (3) with  $x = 0$ . If the material properties  $E_{ps,1}, \sigma_{0,1}, E_{ps,2},$  and  $\sigma_{0,2}$  are not known, they can be used as fitting parameters for modeling experimental curves, thus providing experimental values for these properties. The procedure for the extraction of  $E_{ps,1}, \sigma_{0,1}, E_{ps,2},$  and  $\sigma_{0,2}$  avoids the multiplicity of solutions usually found in multiparameter fittings by (1) first extracting  $E_{ps,2}$  and  $\sigma_{0,2}$  of the nitride layer from the measurement of homogeneous  $\text{Si}_3\text{N}_4$  diaphragms and then (2) determining  $E_{ps,1}$  and  $\sigma_{0,1}$  of the oxide films in the composites, while having  $E_{ps,2}$  and  $\sigma_{0,2}$  fixed.

## Stress Profiles

In order to extract fracture properties of the composite structures, in addition to their load-deflection response, quantitative knowledge of the mechanical state at every



point of the diaphragm cross section is required. Consider a diaphragm point with the coordinates  $x$  and  $z$  in the structure at rest position. In the deflected diaphragm, the material at the considered point experiences three stresses adding up to the total stress  $\sigma_{xx}(x, z)$ . The first contribution is the prestress  $\sigma_0(z)$ . The second contribution is caused by the additional strain due to the diaphragm elongation. The third contribution results from the bending of the structure and the fact that the point is located away from the reference plane. The additional bending strain is given by  $(z - z_0)/R(x)$ , where  $(z - z_0)$  is the distance of the considered point from the reference plane and  $R(x)$  denotes the radius of curvature of the structure at  $x$  (Landau and Lifshitz 1986). This strain component corresponds to the third stress component  $E_{ps}(z)(z - z_0)/R(x)$ . Since  $R(x)$  is expressed in terms of the second derivative of the deflection profile as  $R(x) = (d^2w/dx^2)^{-1}$ , the total stress  $\sigma_{xx,n}(x, z)$  in both component layers of the  $\text{SiO}_2/\text{Si}_3\text{N}_4$  samples considered here is conveniently written as

$$\sigma_{xx,n}(x, z) = \sigma_{0,n} + \frac{E_{ps,n}}{2a} \int_{-a/2}^{a/2} \left( \frac{dw}{dx} \right)^2 dx - E_{ps,n}(z - z_0) \frac{d^2w}{dx^2} \quad (5)$$

where the layer number  $n$  takes the values 1 and 2 for  $\text{SiO}_2$  and  $\text{Si}_3\text{N}_4$ , respectively.

As schematically shown in Fig. 1b for the enlarged diaphragm edge, the bending at the structure edges at fracture results in high local stresses. The signs “+” and “-” denote tensile and compressive stress, respectively. At the moment of diaphragm fracture, near the bottom of the edge area in the  $\text{SiO}_2$  layer, the residual compressive stress is converted into a tensile regime. Conversely, in the edge area close to the top of the  $\text{Si}_3\text{N}_4$  layer, the strong tensile stress in the  $\text{Si}_3\text{N}_4$  is changed into compressive values. Compared to these dramatic variations, the stress change resulting from the stretching and the weak bending in the central part of the diaphragm is much smaller. The maximum stress  $\sigma_{\max}$  at fracture occurs at the bottom of the  $\text{SiO}_2$  layer in the composite diaphragm edge. Using Eq. (5),  $\sigma_{\max}$  is found to be

$$\sigma_{\max} = \sigma_{0,1} + \frac{E_{ps,1}}{2a} \int_{-a/2}^{a/2} \left( \frac{dw_{fr}}{dx} \right)^2 dx + E_{ps,1} z_0 \frac{d^2w_{fr}}{dx^2} \Big|_{x=\pm a/2} \quad (6)$$

where  $w_{fr}$  denotes the deflection profile at fracture.

Clearly, the bilayer model developed in this section is easily applied to single material diaphragms, e.g., diaphragms made of LPCVD  $\text{Si}_3\text{N}_4$  only. This is achieved by using the mechanical expressions in Table 1 in the equations for homogeneous thin plates or by for setting  $h_1 = 0$  in the expressions for the composite bilayer thin plate.

## Fracture Analysis

The analysis of fracture in materials may be performed using different approaches. In traditional fracture mechanics, energies are measured and material parameters such

as energy release rate and fracture toughness are extracted. Xiang et al. (2007) developed a method where flaws of known size, extending through the film thickness, are introduced by focused ion beam along the centerline of rectangular diaphragms, which in turn are pressurized until fracture. These authors take advantage of the quasi-uniform stress state in the central part of the diaphragms, where the precracks are located, which enables the accurate extraction of energy release rates and, consequently, of the fracture toughness. Indeed, it combines the energy release rate with technological parameters leading to a particular distribution of flaws in a material.

The materials characterized here are both brittle, and their fracture is most likely induced by fabrication-related flaws distributed in regions of the thin films (lines, surfaces or volumes) that have been exposed to higher values of tensile stress. For analyzing brittle material strength data, the Weibull distribution has been widely adopted (Weibull 1951). When fracture-causing flaws are distributed uniformly throughout a diaphragm specimen loaded to an arbitrary stress distribution  $\sigma_{\text{total}}(x, y, z)$ , the fracture probability of the sample is

$$F = 1 - \exp \left[ - \iiint_{V^+} \left( \frac{\sigma_{\text{total}}(x, y, z) - \sigma_u}{\sigma_{0,w}} \right)^m dx dy dz \right] \quad (7)$$

where  $m$  is the Weibull modulus, which describes the scattering of the observed strength values;  $\sigma_{0,w}$  is the so-called Weibull scale parameter, which has the awkward dimensions of [stress] [volume]<sup>1/m</sup>, and is proportional to the mean fracture strength  $\mu$ ;  $\sigma_u$  is the threshold stress below which the fracture probability is zero; and  $V^+$  is the volume of the sample where  $\sigma_{\text{total}}(x, y, z) > \sigma_u$ . Most often,  $\sigma_u = 0$  is assumed. When the strength is limited by edge or surface flaws rather than volume flaws, the volume integration in Eq. (7) is replaced by a line or surface integration, respectively (Yang and Paul 2002).

For uniformly loaded samples with volume  $V_E$  under tensile stress  $\sigma_{\text{ref}}$ , Eq. (7) simplifies to (Yang et al. 2008)

$$F = 1 - \exp \left[ (\sigma_{\text{ref}} / \sigma_{0,w})^m V_E \right] \quad (8)$$

From this distribution, it can be shown that the mean fracture strength  $\mu$  is given by (Yang et al. 2008)

$$\mu = \sigma_{0,w} V_E^{-1/m} \Gamma \left( \frac{m+1}{m} \right) \quad (9)$$

By combining Eqs. (8) and (9), one obtains the theoretical Weibull distribution for uniformly loaded specimens

$$F(\sigma_{\text{ref}}, \mu, m) = 1 - \exp \left\{ - \left[ \frac{\sigma_{\text{ref}}}{\mu} \Gamma \left( \frac{m+1}{m} \right) \right]^m \right\} \quad (10)$$

For samples with different geometries and nonuniform stress distribution, the analysis is slightly more complex. The experimental data consists of  $J$  samples with different dimensions, each stressed to an individual fracture stress  $\sigma_{\text{total},j}(x, y, z)$ , where  $j = 1, \dots, J$ . The goal is then to compare these data, to compute their fracture probability using Eqs. (8) or (10), and to extract the Weibull modulus  $m$  and mean fracture strength  $\mu$  from these data. One method to achieve this is to translate the samples into an equivalent set of  $J$  uniformly stressed samples with identical reference volume  $V_E$ . Then, for the  $j$ th sample, e.g., sample with number  $j$ , an equivalent uniform reference stress  $\sigma_{\text{ref},j}(m)$  can be defined from  $\sigma_{\text{ref}}^m V_E = \int_{V^+} \sigma^m dV$  by comparing Eqs. (7) and (8), yielding

$$\sigma_{\text{ref},j}(m) = \left\{ \frac{1}{V_E} \iiint_{V^+} [\sigma_{\text{total}}(x, y, z)]^m dx dy dz \right\}^{1/m} \quad (11)$$

Equation (10) is the standardized form of the Weibull distribution which, along Eq. (11), is particularly useful in discussing variability since it is valid for any specimen geometry and load arrangement and able to predict size effects.

For a given choice of Weibull modulus  $m$ , the samples can then be reordered in increasing size of  $\sigma_{\text{ref},j}(m)$ , with reordered index  $j' = 1, \dots, J$ . Each sample is then assigned the fracture probability  $F_j = j/(J+1)$  resulting in a list of data pairs  $\{\sigma_{\text{ref},j}(m), F_j\}$ . Each sample is then assigned the fracture probability  $F_j = j/(J+1)$  resulting in a list of data pairs  $\{\sigma_{\text{ref},j}(m), F_j\}$ . Clearly,  $\sigma_{\text{ref},j}(m)$  depends on the choice of the Weibull modulus  $m$ . Similar to the case of uniformly stressed samples,  $m$  and  $\mu$  assume optimal values when the Weibull function  $F(\sigma_{\text{ref}}, \mu, m)$ , as defined in (10), fits the experimental Weibull data  $\{\sigma_{\text{ref},j}(m), F_j\}$  best. The quality of the fit is measured by the root mean-square deviation

$$\delta(\mu, m) = \sqrt{\frac{1}{J} \sum_{j'=1}^J [F(\sigma_{\text{ref},j'}(m), \mu, m) - F_j]^2} \quad (12)$$

To find optimal values for  $\mu$  and  $m$ , both parameters are varied until the minimum of  $\delta(\mu, m)$  is found. For each choice of  $\mu$  and  $m$  during such an iterative search, a new set of  $\sigma_{\text{ref},j}(m)$  values has to be computed, a new reordering has to be assessed and performed if necessary, and  $\delta(\mu, m)$  needs to be recalculated. To take advantage of the plane-strain situation portion of each diaphragm analyzed, Eq. (11) is simplified into the 2-D integral form

$$\sigma_{\text{ref},j}(m) = \left\{ \frac{b-ca}{V_E} \int_0^{h_1} \int_{-a/2}^{a/2} \{\max[0, \sigma_{xx,1,j}(x, z)]\}^m dx dz \right\}^{1/m} \quad (13)$$

where  $\sigma_{xx,1,j}(x, z)$  denotes the plane-strain stress distribution, as defined in Eq. (5), in the oxide layer of sample number  $j$ . Similar expressions obtained by identical reasoning are obtained for the single-layer  $\text{Si}_3\text{N}_4$  diaphragms. The function  $\max[0, \sigma_{xx,1,j}(x, z)]$  takes the maximum of 0 and  $\sigma_{xx,1,j}(x, z)$ ; thus, the cross-sectional integral includes only the region where the oxide layer is under tensile stress. The influence of the silicon nitride in the composite diaphragms is neglected because its own fracture strength is far above the stress levels reached in the composite structures at fracture, and consequently, the resistance against fracture is limited by the oxide layer. Also, in Eq. (13), the integration in  $y$ -direction has been replaced by the multiplication with  $b - ca$ , i.e., with the diaphragm length  $b$  corrected for corner effects and for the stress distribution along the small sides using a correction factor  $c$  to be determined, which, along with  $\mu$  and  $m$ , may be used as a minimization parameter for the standard deviation given by Eq. (12). For  $V_E$ , the volume  $2.5 \times 10^5 \mu\text{m}^3$  is chosen. This is of the same order of magnitude as the volumes of the diaphragms characterized.

The hypothesis of diaphragm fracture dominated by surface flaws instead of volume flaws is analyzed by replacing the volume integration in Eq. (13) with the surface integral over the bottom surface  $z = 0$  of the composite diaphragms.

$$\sigma_{\text{ref},j}(m) = \left\{ \frac{b - ca}{A_E} \int_{-a/2}^{a/2} \{ \max[0, \sigma_{xx,1,j}(x, 0)] \}^m dx \right\}^{1/m} \quad (14)$$

with the effective area  $A_E = 5 \times 10^6 \mu\text{m}^2$ , similar to the area of the samples. Finally, the hypothesis of fracture originating from the lower edge line of the silicon oxide is analogously implemented through

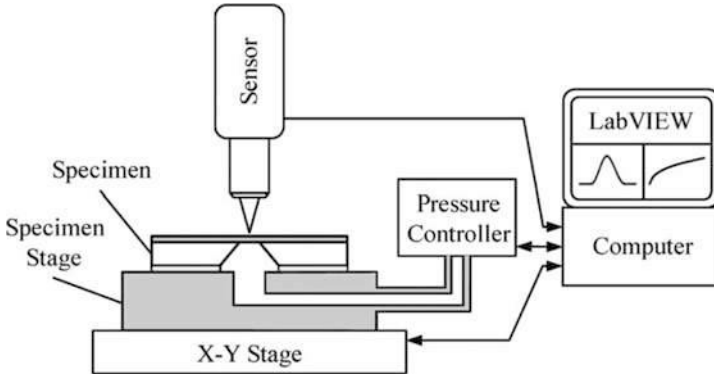
$$\sigma_{\text{ref},j}(m) = \left( \frac{b - ca}{L_E} \right)^{1/m} \sigma_{\text{max},j} \quad (15)$$

where  $\sigma_{\text{max},j}$  denotes the maximum oxide stress, as defined in Eq. (6), in the  $j$ th sample. The reference length  $L_E$  is chosen as  $10^4 \mu\text{m}$ , again similar to the real sample length.

---

## Experimental Apparatus

The experimental apparatus is schematically shown in Fig. 2. It consists of a pressure controller and a scanning autofocusing optical profilometer. Both apparatus are controlled by a LabVIEW program. The membrane deflection profile  $w(x)$ , particularly the center deflection  $w_0$ , as a function of the applied differential pressure  $P$ , was monitored by the optical profilometer.



**Fig. 2** Schematic diagram of the load-deflection-measurement apparatus

## Key Research Findings

### LPCVD Silicon Nitride Films

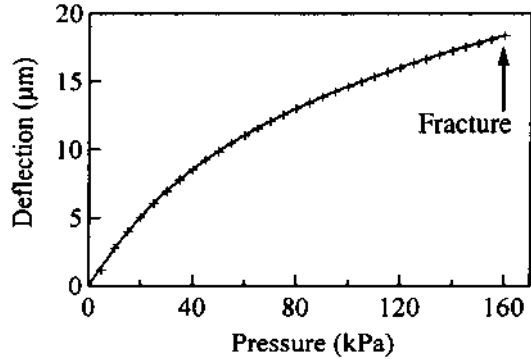
LPCVD silicon nitride thin films with thicknesses between 72.6 and 83.4 nm were fabricated and characterized with respect to their load-deflection response and fracture parameters. The silicon nitride was deposited on a 300 nm thick silicon oxide on silicon wafers. Five types of membranes with designed widths  $a$  from 200 to 600  $\mu\text{m}$ , lengths  $L$  from 2400 to 11,000  $\mu\text{m}$ , and aspect ratios between 6:1 and 25:1 were fabricated by rear KOH bulk micromachining of the substrate. Due to underetching, membrane widths and lengths increased between 20 and 25  $\mu\text{m}$  over the designed values. A total number of 52 membranes were progressively loaded up to the fracture pressure  $P_{fr}$  leading to a deflection  $w_{0,fr}$ . A typical load-deflection response for silicon nitride membrane with dimensions  $a = 424 \mu\text{m}$ ,  $b = 6421 \mu\text{m}$ , and  $h = 78 \text{ nm}$  is shown in Fig. 3.

From the load-deflection results,  $S_0$  and  $E_{ps}$  of the 52 samples were extracted using the model. Extracted prestress levels and plane-strain moduli are  $1.04 \pm 0.16$  and  $287.7 \pm 42.3 \text{ GPa}$ , respectively.

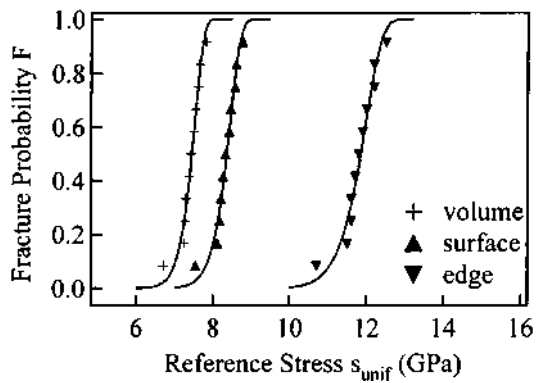
Typical values of the maximum stress  $S_{max}$  in the loaded membranes are in the range from 10.8 to 11.7 GPa. Since CVD silicon nitride is brittle, its strength is governed by the largest flaw in the membranes. The fracture is most likely initiated at defects near the bottom of the edge or at the location where the membrane meets its silicon oxide support. It should be appreciated that the scatter of  $S_{max}$  values is amazingly small in comparison with the values reported, e.g., for polysilicon (Tsuchiya et al. 1998).

In LPCVD silicon nitride thin film, mainly two kinds of defects may contribute to the membrane fracture, namely, volume defects resulting from the amorphous structure, and surface defects mainly from the fabrication process. Weibull fitting plots for membranes ( $a = 420\text{--}425 \mu\text{m}$  and  $L = 2420\text{--}2435 \mu\text{m}$ ) and for the entire set

**Fig. 3** Typical load-deflection response of silicon nitride membrane and fit with the long rectangular plate model (*solid line*). Membrane dimensions are  $a = 424 \mu\text{m}$ ,  $b = 6421 \mu\text{m}$  and  $h = 78 \text{ nm}$



**Fig. 4** Fracture probability for type 3 membranes for the three cases of edge, surface, and volume integration. *Solid lines* are fits with Weibull distribution function

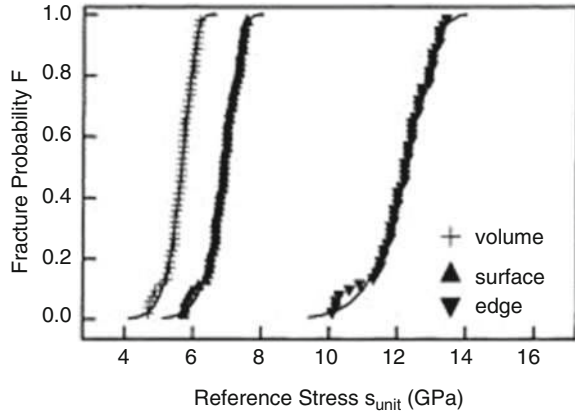


of 52 membranes are shown in Figs. 4 and 5, respectively. The quality of the edge, surface, and volume fits could, in principle, enable one to decide where the fracture originates from, i.e. edge, surface, or volume. The three fitting procedures give only a slightly better fit with edge integration, so that surface and volume effects cannot be unambiguously ruled out. Therefore, a conclusion about the fracture origin cannot be reached at this point. Nevertheless, the conclusion of the present study that LPCVD nitride is able to withstand local stresses larger than 10 GPa is untouched by this uncertainty.

## Silicon Carbide Film

The 3C-SiC thin films were deposited on 2-in (100)-oriented silicon wafers in an LPCVD system using  $\text{SiH}_4$  and  $\text{C}_2\text{H}_4$  as Si and C precursors, respectively (Zhou et al. 2008). The deposition temperature and pressure were 1200 °C and 40 torr, respectively. The films is 0.40  $\mu\text{m}$  thick (wafer A). During the SiC film deposition process, the  $\text{SiH}_4$  and  $\text{C}_2\text{H}_4$  flow rates for wafer A were 1.2 and 3.6 sccm, respectively.

**Fig. 5** Fracture probability for all membranes with all different sizes for the three cases of edge, surface, and volume integration. Solid lines are fits with Weibull distribution function



**Table 2** Geometrical dimensions and fracture properties of four types of membranes in wafer A

Membrane type	1	2	3	4	All
$a(\mu\text{m})$	215–228	317–328	417–426	518–527	215–527
$b(\mu\text{m})$	1624–2627	2425–3932	3217–3623	4022–4029	1624–4029
$N$	9	5	4	4	22
$E_{\text{ps}}(\text{GPa})$	$433 \pm 72$	$402 \pm 47$	$415 \pm 70$	$390 \pm 45$	$415 \pm 61$
$s_0(\text{MPa})$	$319 \pm 46$	$347 \pm 56$	$326 \pm 49$	$293 \pm 30$	$322 \pm 47$
$s_{\text{max}}(\text{GPa})$	$6.02 \pm 0.9$	$6.88 \pm 0.88$	$6.92 \pm 0.64$	$6.65 \pm 0.79$	$6.49 \pm 0.88$

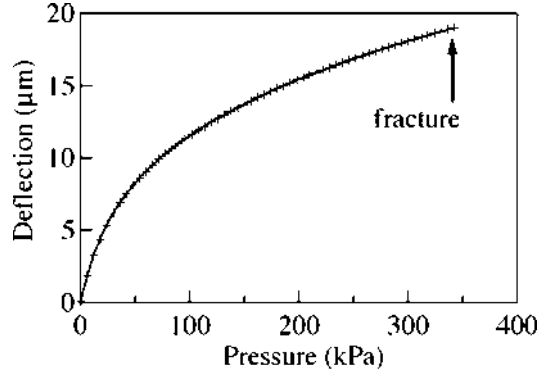
The 3C-SiC membranes were fabricated by anisotropic etching of the silicon substrate from the backside in KOH solution at 80 °C with etch stopping on the SiC layer. Undercutting increased the membrane widths of wafer A between 10 and 35  $\mu\text{m}$  over the designed values. Types of membranes in two wafers were fabricated, as listed in Table 2.

Fig. 6 shows a typical load-deflection response for a 3C-SiC membrane with dimensions  $a = 524 \mu\text{m}$ ,  $b = 4026 \mu\text{m}$ , and  $h = 0.40 \mu\text{m}$ . The fracture occurred without any yielding, the 3C-SiC is brittle.

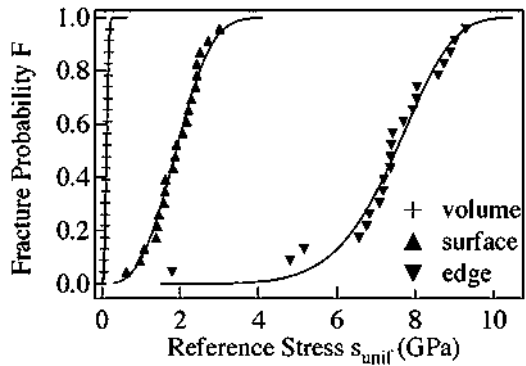
The material properties  $S_0$  and  $E_{\text{ps}}$  were extracted from the load-deflection results of the 40 SiC membranes by using the analytic model. The pre-stresses  $S_0$  and plane-strain moduli  $E_{\text{ps}}$  are  $322 \pm 47 \text{ MPa}$  and  $415 \pm 61 \text{ GPa}$ , respectively, for wafer A. The deflection profile  $w(x)$ , the stress distribution  $S_{\text{total}}(x, z)$ , and the maximum stress  $S_{\text{max}}$  at the membrane bottom edge at the moment of fracture were calculated. By assuming a Poisson's ratio of 0.168 for the 3C-SiC films, the Young's moduli for wafer A is  $403 \pm 59$ . The value of Young's modulus for wafer A is comparable to the reported results for the epitaxial 3C-SiC films characterized by the microtensile and bulge testings (Fu et al. 2005).

Table 2 show that the fracture strength values for the thinner 3C-SiC films on wafer A ( $6.49 \pm 0.88 \text{ GPa}$ ) are high compared with the results reported for the

**Fig. 6** Typical load-deflection response of SiC membranes (“+”) and theoretical fit with the long rectangular plate model (*solid line*). Membrane dimensions are  $a = 524 \mu\text{m}$ ,  $b = 4026 \mu\text{m}$ , and  $h = 0.40 \mu\text{m}$



**Fig. 7** Fracture probability for all membranes in wafer A for three cases of edge, surface, and volume integration. Markers are the experimental Weibull data, and solid lines are the fits with the Weibull distribution function



microbridge testing (Espinosa et al. 2006), the microtensile testing (Jackson et al. 2005), and the bulge testing of square membranes in combination with FEA (Nemeth et al. 2001), but lower than the results of 0.2- $\mu\text{m}$ -thick films by the beam-bending testing (Kahn et al. 2000).

The Weibull fitting plots for the entire set of membranes in wafer A is shown in Fig. 7. For wafer A, the surface fitting leads to a significantly smaller  $\delta$  than the edge and volume fittings. Therefore, the effect of the volume and edge flaws cannot be definitely excluded.

Table 3 summarizes the extracted  $m$ ,  $\mu$ , and  $\delta$  results for each data set. As previously stated,  $m$  is a property of flaw distribution,  $\delta$  is an rms deviation, and  $\mu$  depends on the specimen size. The fitting results of the edge, surface, and volume assumption could, in principle, predict the fracture origin. For wafer A, the surface fitting leads to a significantly smaller  $\delta$  than the volume fittings.

## SiO<sub>2</sub>/Si<sub>3</sub>N<sub>4</sub> Composite Diaphragm

The strongly compressive prestress levels in thermal SiO<sub>2</sub> films (−200 to −300 MPa) and the brittleness of this material make it difficult to obtain single-layer SiO<sub>2</sub>



**Table 3** Weibull modulus  $m$ , mean reference stress  $\mu$ , and rms deviation  $\delta$  extracted from the edge, surface and volume fits with weibull distribution function

$h(\mu\text{m})$	Membrane type	Failure mode	$m$	$\mu(\text{GPa})$	$\delta$
0.40	All	Edge	7.2	7.4	0.048
		Surface	3.0	1.9	0.030
		Volume	2.7	0.1	0.041
	Type 1	Edge	8.8	6.5	0.058
		Surface	8.1	2.5	0.060
		Volume	6.8	0.5	0.043

diaphragms by conventional MEMS fabrication processes. In order to produce diaphragm structures, an LPCVD silicon nitride thin film with strongly tensile prestress is used to balance the compressive stress in the  $\text{SiO}_2$  layer.  $\text{SiO}_2$  thin films with target thicknesses  $h_1$  between 500 and 1000 nm are produced and characterized with respect to the load-deflection response and fracture parameters of  $\text{SiO}_2/\text{Si}_3\text{N}_4$  composite diaphragms.

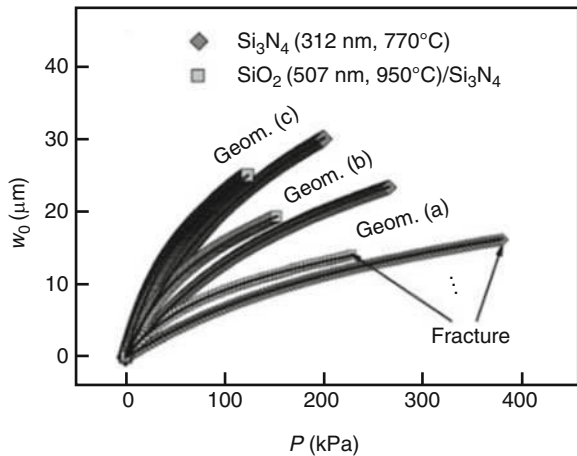
In order to fabricate the bilayer diaphragms, silicon wafers are first thermally oxidized to obtain  $\text{SiO}_2$  layers. The LPCVD silicon nitride films, with a thickness of 312 nm, are then deposited on the top of the silicon oxide layer. Single-layer diaphragms composed solely of  $\text{Si}_3\text{N}_4$  are fabricated and characterized, as well for reference purposes. Diaphragms with three different sizes, referenced hereafter as diaphragms with geometries (a), (b), and (c), have been fabricated by rear potassium hydroxide (KOH) followed by tetramethylammonium hydroxide (TMAH) bulk micromachining of the Si substrates. Geometry (a) includes the smaller diaphragms with  $a = 472.2 \pm 5.0 \mu\text{m}$  and  $b = 4072.3 \pm 13.0 \mu\text{m}$  (aspect ratio  $b/a = 8.62 \pm 0.09$ ), (b) refers to medium-sized diaphragms with  $a = 674.0 \pm 4.5 \mu\text{m}$  and  $b = 6072.6 \pm 17.7 \mu\text{m}$  ( $b/a = 9.01 \pm 0.06$ ), and to the larger structures of geometry of type (c) correspond  $a = 874.5 \pm 4.3 \mu\text{m}$  and  $b = 8072.2 \pm 10.11 \mu\text{m}$  ( $b/a = 9.23 \pm 0.04$ ). Table 4 summarizes the total number of diaphragms that have been characterized, and respective fabrication yield, as a function of the oxide thickness and growth temperature. The relatively large thickness of the  $\text{Si}_3\text{N}_4$  films, 312 nm, in single-layer diaphragms ( $h_1 = 0$ ), along with relatively high levels of tensile prestress, leads to a yield of 47.5% in the  $\text{Si}_3\text{N}_4$  structures, well below the almost 100% observed and reported for 100-nm-thick films (Gaspar et al. 2007).

A total number of 38 silicon nitride diaphragms and 236 composite structures have progressively been loaded up to the fracture pressure  $P_{fr}$  corresponding to the maximum center deflection  $w_{0,fr}$ . As an example, Typical load-deflection responses of single and bilayer diaphragms, with different sizes, are shown in Fig. 8. The fracture occurs without any yielding, in agreement with the well-known brittleness of LPCVD silicon nitride and thermal  $\text{SiO}_2$  thin films.

Table 5 lists some of the elastic data reported in the literature for LPCVD  $\text{Si}_3\text{N}_4$  and wet thermal  $\text{SiO}_2$  films grown with similar conditions as the layers reported in this paper Tabata et al. (1989), Sobek et al. (1993), Petersen (1978), Petersen and

**Table 4** Number of diaphragms characterized  $J$  and fabrication yield as a function of the oxide thickness  $h_1$  in diaphragms composed of  $\text{SiO}_2/\text{Si}_3\text{N}_4$  stacks

$T_{OX} = 950\text{ }^\circ\text{C}$			$T_{OX} = 1050\text{ }^\circ\text{C}$		
$h_1(\text{nm})$	Geometry	$J$	$h_1(\text{nm})$	Geometry	$J$
0	(a)	11	0	(a)	11
	(b)	15		(b)	15
	(c)	12		(c)	12
	Total	38		Total	38
	Yield	48%		Yield	48%
507	(a)	20	525	(a)	18
	(b)	20		(b)	22
	(c)	15		(c)	20
	Total	55		Total	60
	Yield	69%		Yield	75%
710	(a)	8	777	(a)	12
	(b)	11		(b)	12
	(c)	11		(c)	11
	Total	30		Total	35
	Yield	38%		Yield	50%
985	(a)	18	1016	(a)	9
	(b)	14		(b)	2
	(c)	12		(c)	1

**Fig. 8** Center deflection  $w_0$  as a function of the applied pressure for single  $\text{Si}_3\text{N}_4$  and  $\text{SiO}_2/\text{Si}_3\text{N}_4$  composite diaphragms with dimensions corresponding to geometries (a), (b), and (c), and respective fits. The diaphragms are bulged until fracture

Guarnieri (1979), Weihs et al. (1988), and Jaccodine and Schlegel (1966).  $E_{ps,1} = 49.1 \pm 7.4$  GPa is obtained for  $\text{SiO}_2$ , and  $E_{ps2} = 304.8 \pm 12.2$  GPa is found for  $\text{Si}_3\text{N}_4$ , comparable to data previously determined using analytical models for single-layer diaphragms. The values extracted by the different authors for the elastic constant and prestress levels of the nitride films agree with each other. The same applies to the experimentally obtained residual stress of the oxide layers. After

**Table 5** Elastic constants and prestress levels of LPCVD Si<sub>3</sub>N<sub>4</sub> and wet thermal SiO<sub>2</sub> films reported in the literature, grown under similar conditions as reported in this paper. The Young's modulus  $E$  are computed from  $E_{ps}$  and  $E_b$ , assuming  $\nu = 0.22$  and  $\nu = 0.2$  for the nitride and oxide films, respectively

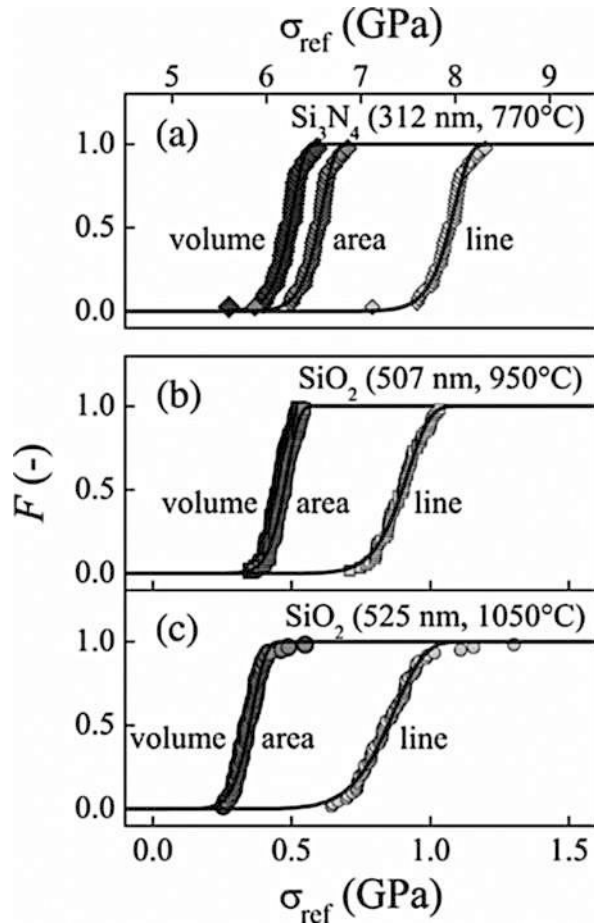
Thin film	Elastic const. [GPa]	Young's mod. $E$ [GPa]	Prestress $\sigma_0$ [MPa]	Reference	
LPCVD Si <sub>3</sub> N <sub>4</sub>	$E_{ps}$	305 ± 12	293 ± 12	1132 ± 34	This paper
	$E_b$	280 ± 31	224 ± 25	1000 ± 30	Sobek et al. (1993)
	$E$	290	290	1000	Tabata et al. (1989)
Thermal SiO <sub>2</sub> (wet)	$E_{ps}$	49 ± 7	47 ± 7	-259 ± 23	This paper
	$E_b$	60 ± 17	48 ± 13	-250 ± 25	Sobek et al. (1993)
	$E$	57 ± 11	57 ± 11	n.a.	Petersen (1978) and Petersen and Guarnieri (1979)
	$E$	64 ± 6	64 ± 6	n.a.	Weihls et al. (1988)
	$E$	66	66	-280–240	Jaccodine and Schlegel (1966)

calculating the Young's modulus  $E$  from the elastic constants reported by the different sources, using the appropriate relations and reasonable values of Poisson's ratio, Table 5 shows that the Young's modulus of the thermally grown films reported in this work are within the range of literature data, although at their lower end.

In order to identify the origin of the fracture, the three fitting procedures, corresponding to line, area, and volume flaws, are applied in a pooled analysis to the entire sets of diaphragms with different dimensions per material and thickness. Nonpooled values corresponding to the individual diaphragm subsets with different dimensions have been evaluated as well. The pooled Weibull fit plots, from which  $\mu$  and  $m$  are extracted, with  $c = 0$ , are shown in Fig. 9. The mean fracture strengths  $\mu$  obtained from the line assumption have identical values and oxide thickness behavior as  $\sigma_{max}$ . The Weibull moduli  $m$  extracted from pooled and nonpooled fits agree with each other. Also, no significant differences were found between the values of  $m$  extracted from line, surface or volume fits. As an example, Weibull moduli of 12.7, 12.0, and 11.8 are obtained for the 507-nm-thick 950 °C oxide films from the line, area, and volume assumptions, respectively. The silicon nitride thin films have an amazingly high Weibull modulus of  $m = 50.0 \pm 3.6$ , in agreement with previous reports (Paul and Gaspar 2007), making this material highly predictable from the fracture point of view.

The mechanical properties and the fracture behavior of MEMS/NEMS thin films are characterized with high accuracy by bulge test combined with a refined load-deflection model of the long rectangular membranes, which takes into account the bending stiffness and prestress of the membrane material. Edge, surface and volume fits using the Weibull distribution function is performed to determine the fracture origins in brittle materials. This approach enables better understanding of mechanical and reliability-related properties of thin films at the design stage and provides a guideline for device lifetime evaluation.

**Fig. 9** Pooled fracture probability  $F$  of (a)  $\text{Si}_3\text{N}_4$  films obtained from the measurement of single layers, and of  $\text{SiO}_2$  grown at (b)  $950^\circ\text{C}$  and (c)  $1050^\circ\text{C}$ , extracted from the bulge test of composites, for the three cases of line, area, and volume distributed flaws. The curves are fits of the Weibull distribution function to the experimental data. The reference stress  $\sigma_{\text{ref}}$  is calculated with the method discussed in the text



## References

- Espinosa HD, Peng B, Moldovan N, Friedmann TA, Xiao X, Mancini DC, Auciello O, Carlisle J, Zorman CA, Mehregany M (2006) Elasticity, strength, and toughness of single crystal silicon carbide, ultrananocrystalline diamond, and hydrogen-free tetrahedral amorphous carbon. *Appl Phys Lett* 89(7):073–111
- Fu XA, Dunning JL, Zorman CA, Mehregany M (2005) Measurement of residual stress and elastic modulus of polycrystalline 3C-SiC films deposited by low-pressure chemical vapor deposition. *Thin Solid Films* 492(1/2):195–202
- Gaspar J, Ruther P, Paul O (2007) Mechanical characterization of thin-film composites using the load-deflection response of multilayer membranes—elastic and fracture properties. In: *Proceedings of materials research social symposium*, vol 977, pp 91–96
- Gerlach G, Schroth A, Pertsch P (1996) Influence of clamping conditions on microstructure compliance. *Sens Mater* 8(2):79–98
- Jaccodine RJ, Schlegel WA (1966) Measurement of strains at Si–SiO<sub>2</sub> interface. *J Appl Phys* 37(6):2429–2434

- Jackson KM, Dunning J, Zorman C, Sharpe WN Jr (2005) Mechanical properties of epitaxial 3C silicon carbide thin films. *J Microelectromech Syst* 14(4):664–672
- Kahn H, Tayebi N, Ballarini R, Mullen RL, Heuer AH (2000) Waferlevel strength and fracture toughness testing of surface-micromachined MEMS devices. In: *Proceedings of materials research social symposium*, vol 605, pp 25–30
- Landau L, Lifshitz EM (1986) *Theory of elasticity*, ch. 2. Pergamon, Oxford
- Nemeth N, Jadaan O, Palko J, Mitchell J, Zorman CA 2001 Structural modeling and probabilistic characterization of MEMS pressure sensor membranes. In: *Proceedings MEMS: mechanical measurement, social experimental mechanical symposium*, Portland, p 46
- Paul O, Baltes H (1999) Mechanical behavior and sound generation efficiency of prestressed, elastically clamped and thermomechanically driven thin film sandwiches. *J Micromech Microeng* 9(1):19–29
- Paul O, Gaspar J (2007) Thin-film characterization using the bulge test. In: Tabata O, Tsuchiya T (eds) *Reliability of MEMS*. Weinheim: Wiley-VCH, , ch. 3, pp. 67–122
- Petersen KE (1978) Dynamic micromechanics on silicon: techniques and devices. *IEEE Trans Electron Devices* ED-25(10):1241–1250
- Petersen KE, Guarnieri CR (1979) Young's modulus measurements of thin films using micro-mechanics. *J Appl Phys* 50(11):6761–6766
- Pourahmadi F, Gee D, Petersen K (1991) The effect of corner radius of curvature on the mechanical strength of micromachined single-crystal silicon structure. In: *Technical digest of sixth international conference on solid-state sensors and actuators (Transducers'91)*, New York, pp 197–200
- Sobek D, Young AM, Gray ML, Senturia SD (1993) A microfabricated flow chamber for optical measurements in fluids. In: *Proceedings MEMS workshop, fort lauderdale, Feb 1993*, pp 219–224
- Soderkvist J (1993) Similarities between piezoelectric, thermal and other internal means of exciting vibrations. *J Micromech Microeng* 3(1):24–31
- Tabata O, Kawahata K, Sugiyama S, Igarashi I (1989) Mechanical property measurements of thin films using load-deflection of composite rectangular membranes. *Sensors Actuators* 20(1/2):135–141
- Tsuchiya T, Tabata O, Sakata J, Taga Y (1998) Specimen size effect on tensile strength of surface micromachined polycrystalline silicon thin films. *J Microelectromech Syst* 7:106–113
- Walker JA, Gabriel KJ, Mehregany M (1991) Mechanical integrity of polysilicon films exposed to hydrofluoric acid solutions. *J Electron Mater* 20:665–670
- Weibull W (1951) A statistical distribution function of wide applicability. *J Appl Mech* 18(3):293–297
- Weihl TP, Hong S, Bravman JC, Nix WD (1988) Mechanical deflection of cantilever microbeams: a new technique for testing the mechanical properties of thin films. *J Mater Res* 3(5):931–942
- Xiang Y, McKinnell J, Ang W-M, Vlassak JJ (2007) Measuring the fracture toughness of ultra-thin films with application to AlTa coatings. *Int J Fract* 144(3):173–179
- Yang JL, Paul O (2002) Fracture properties of LPCVD silicon nitride thin films from the load-deflection of long membranes. *Sens Actuator A, Phys* 97/98:520–526
- Yang JL, Gaspar J, Paul O (2008) Fracture properties of LPCVD silicon nitride and thermal silicon oxide thin films from the load-deflection of long Si<sub>3</sub>N<sub>4</sub> and SiO<sub>2</sub>/Si<sub>3</sub>N<sub>4</sub> membranes. *J Microelectromech Syst* 17(5):1120–1134
- Zhou W, Yang J, Sun G, Liu X, Yang F, Li J (2008) Fracture properties fo silicon carbide thin films by bulge test of long rectangular menbrane. *J Microelectromech Syst* 17:453–461

---

## **Part III**

# **Physical Microsensors**



# High Temperature Silicon Pressure Sensors

Zhuangde Jiang, Yulong Zhao, Libo Zhao, and Tingzhong Xu

## Contents

Introduction .....	326
The Piezoresistive Effect of High-Temperature Pressure Sensor .....	328
Mechanics Model of the Pressure Sensor Chip .....	332
Diaphragm Deflection Model .....	332
Theory Verification .....	344
Structure Designing, Lithography Mask Designing and Fabrication of the Sensor Chip .....	348
Structure Designing .....	349
Lithography Mask Designing .....	354
Micromachining Fabrication Process .....	356
Packaging Structure for the Pressure Sensor Chip .....	361
High-Temperature Packaging .....	361
Ultrahigh Pressure Packaging .....	363
Harsh Environment Packaging .....	365
High-Frequency Response Packaging .....	366
Packaging for Ultrahigh Sensitive Pressure Sensor Chip .....	367
Performance Test Analysis of High-Temperature and High-Frequency Pressure Sensor .....	369
Characteristic Analysis of High-Temperature Piezoresistive Chip .....	369

---

Z. Jiang (✉) · L. Zhao

State Key Laboratory for Manufacturing Systems Engineering, Collaborative Innovation Center of Suzhou Nano Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China

International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, Xi'an Jiaotong University, Xi'an, Shaanxi, China

e-mail: [zdjiang@mail.xjtu.edu.cn](mailto:zdjiang@mail.xjtu.edu.cn); [libozhao@mail.xjtu.edu.cn](mailto:libozhao@mail.xjtu.edu.cn)

Y. Zhao · T. Xu

State Key Laboratory for Manufacturing Systems Engineering, International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technologies, Collaborative Innovation Center of Suzhou Nano Science and Technology, School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China

e-mail: [zhaoyulong@mail.xjtu.edu.cn](mailto:zhaoyulong@mail.xjtu.edu.cn); [tingzhongxu@163.com](mailto:tingzhongxu@163.com)

Static Performance of the Sensor .....	372
Dynamic Performance Indicators of the Sensors .....	379
Conclusions .....	382
References .....	384

### Abstract

This chapter gives a complete development process of a high-temperature silicon pressure sensor. Firstly, the piezoresistive effect was described for sensor chip used in high temperature application. Based on the SiO<sub>2</sub> isolation layer, the leakage current generating through p-n junction can be eliminated, which ensured the proper function of piezoresistors in high-temperature working condition. Secondly, the mechanics models for circular, rectangular, and island-structured diaphragms had been presented. Also a theoretical guide for designing and optimizing a sensor chip was presented in section “Mechanics Model of the Pressure Sensor Chip.” Thirdly, based on the mechanics models, section “Structure Designing, Lithography Mask Designing and Fabrication of the Sensor Chip” presented a designing principle of sensor chip to find a balance point between the sensitivity and dynamic response frequency, in order to take full use of elastic strain energy induced by structure deformation. Then, lithography masks for corresponding pattern structures were presented. Followed by the lithography mask designing, the fabrication process of sensor chip was presented. Fourthly, the packaging technology for sensor chip was presented in section “Packaging Structure for the Pressure Sensor Chip.” The packaging structure not only enabled the sensor chip to work properly in a harsh environment but also ensured the packaged sensor had a good performance in dynamic and sensitivity performance. Finally, in section “Conclusions,” the experimental calibration for the sensitivity, dynamic performance, and cross-sensitivity of developed high-temperature silicon pressure sensor were conducted to obtain a comprehensive performance evaluation.

### Keywords

MEMS · Pressure sensor · Non-linearity · High sensitivity · Stress concentration region · Diaphragm deflection · Analytical modeling · High temperature · High frequency response · SOI strain gauge · Piezoresistive · SIMOX

## Introduction

With the development of the fabrication capacity, the structure of the sensor chip become more exquisite and its performance was also largely improved. Thanks to its improving performance and reducing fabrication cost, the microelectromechanical systems (MEMS)-based pressure sensor chip has received a great deal of attention in recent years. This is due not only to the excitement naturally associated with a nascent technology but also because of the great promise of increased miniaturization and performance of MEMS devices over conventional devices.



Most pressure measurements are undertaken at room temperature, but it is becoming more important to measure pressure in the harsh environments, such as high temperatures above 250 °C up to 600 °C, in various fields including food, automobile engines, petroleum exploration, chemical industries, and so on (Cano et al. 1997; Sato et al. 1996; Fielder 2002; Nwafor 2003).

There are two types of high-temperature pressure sensors (HTPS), one is based on indirect measuring methods and the other is based on direct methods. The HTPSs based on indirect measurement have an insulation layer between the sensor chip and the high-temperature medium in order to reduce the direct heat transmission to the sensor chip. The sensors using the direct measurement method have an intrinsic characteristic that is insensitive to high-temperature changes: they are fabricated using materials insensitive to high-temperature changes, such as sapphire, GaAs, SiC, diamond film, ceramics, and so on. High-temperature pressure sensors are mostly piezoresistive sensors which have a small band gap, but recent sensors based on the above materials beyond Si have a high temperature range (Chen and Mehregany 2008; Dehe et al. 1998; Stuchebnikov 1991; Okojie et al. 1998). However, pressure detection is based on the mechanical deformation of sensor chips, which can change according to temperature variation. Accordingly, even though sensors based on new materials have excellent characteristics for high temperature changes, they have a compensation circuit inside in order to compensate for the sensitivity changes at different temperatures. This means that accurate evaluation of the HTPSs should be undertaken before and after fabrication and during use in the field.

Besides, novel optically interrogated pressure and temperature sensors fabricated directly on optical fibers have been fabricated by anodically bonding an ultrathin crystalline silicon onto a fiber end face (Abeyasinghe et al. 2002). However, the optical pressure sensor usually has limitation on the sensor installation because of its bulky structure.

Silicon-based high-temperature piezoresistive pressure sensor chip based on MEMS and SIMOX (separation by implantation of oxygen) technology was the most widely used sensor chip nowadays, which had been developed and produced in batch. This kind of high-temperature pressure sensor had the ability to work reliably under high temperature of above 200 °C, such as the Qian et al. (2008) presented a SIMOX high-temperature pressure sensor with beam-diaphragm packaging structure which could withstand a transient high-temperature shock of 2000 °C. Niu et al. (2014) presented an SOI-based high-pressure sensor with an oil filling technology had a working range of 150 MPa above 200 °C. Zhao et al. (2010) presented an inverted-cup high-temperature sensors which not only shrink the size of the sensor chip but also improve the dynamic performance of the sensor chip.

In this chapter, a whole development process for a high-temperature pressure sensor had been presented. Based on the working mechanism of silicon piezoresistive effects and its SIMOX technology for high-temperature application, the mechanics model, structure designing, MEMS-based fabrication process, packaging structure, and experimental calibration process of the high-temperature pressure sensor had been discussed.

## The Piezoresistive Effect of High-Temperature Pressure Sensor

The piezoresistive pressure sensor is based on the piezoresistive principle of silicon material, by which the physical signal is transformed into the electrical signal.

In 1954, C. S. Smith reported the piezoresistive effect of silicon and germanium (Smith 1954). This discovery enabled the production of semiconductor-based sensors. In conductor and semiconductor materials, changes in interatomic spacing resulting from strain affect the band gaps, making it easier (or harder depending on the material and strain) for electrons to be raised into the conduction band. This results in a change in resistivity of the material. Within a certain range of strain, this relationship is linear, and the resistivity of semiconductor material can be changed due to applied stress and piezoresistive coefficient as shown in Eq. (1):

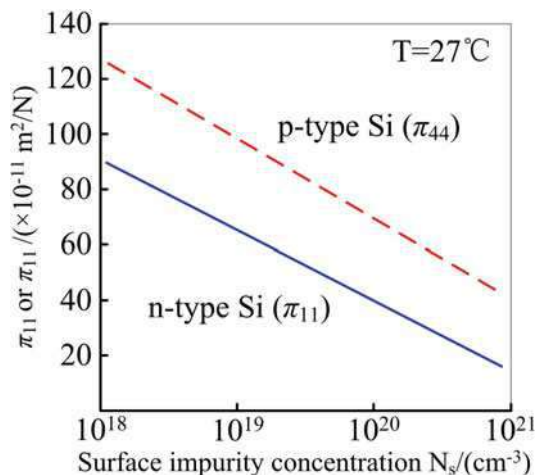
$$\frac{\Delta\rho}{\rho} = \pi_l\sigma_l + \pi_t\sigma_t \quad (1)$$

where  $\rho$  is the resistance value and resistivity of semiconductor material;  $\pi_l$  and  $\pi_t$  denote the longitudinal and transversal piezoresistive coefficients, respectively; and  $\sigma_l$  and  $\sigma_t$  are the longitudinal and transversal stresses of semiconductor material.

The piezoresistive effect of semiconductor materials can be several orders of magnitudes larger than the geometrical effect and is present in materials like germanium, polycrystalline silicon, amorphous silicon, silicon carbide, and single-crystal silicon. The piezoresistive effect has been widely used in sensor mechanism. Hence, the sensing chip based on the piezoresistive effect will have a very high measuring sensitivity.

For MEMS piezoresistive pressure sensor, there are usually four piezoresistors fabricated by doping or ion implantation technology to form the Wheatstone bridge in the sensing chip. As shown in Fig. 1 of the relationship of piezoresistive coefficients and surface impurity concentration, the piezoresistors are always p-type silicon because their piezoresistive coefficients are larger than those of n-type

**Fig. 1** Relationship of piezoresistive coefficients and surface impurity concentration



silicon and have good thermal stability. Based on the piezoresistive effect, the relationship of their resistance variation as well as piezoresistive coefficient and stress is shown in Eq. (2) when the geometrical effect is neglected:

$$\frac{\Delta R}{R} \approx \frac{\Delta \rho}{\rho} = \pi_l \sigma_l + \pi_t \sigma_t \tag{2}$$

where  $R$  is the initial resistance value of piezoresistor and  $\sigma_l$  and  $\sigma_t$  are the longitudinal and transversal stresses of where the piezoresistor locates.

In order to improve the measuring sensitivity of MEMS piezoresistive pressure sensor, it is necessary to utilize maximum piezoresistive coefficient and maximum stress of the piezoresistors. The piezoresistive coefficient of p-type piezoresistor is determined by the doping concentration of boron ion, crystal plane and the direction of single-crystal silicon. For example, on the (100) p-type silicon, the directions of largest absolute piezoresistive coefficient are along the crystal directions  $[\bar{1}10]$  and  $[0\bar{1}1]$ , as shown in Fig. 2. For the stress of piezoresistors, the structure of sensing chip should be designed analytically to utilize the largest stress difference between longitudinal stress and transversal stress in the stress concentration region where the piezoresistors are located; then MEMS piezoresistive pressure sensor with the proposed sensing chip can be guaranteed with high measuring sensitivity.

For example, a layout scheme of circular diaphragm was designed in (100) crystal plane as shown in Fig. 3; there were four piezoresistors located along the edge of

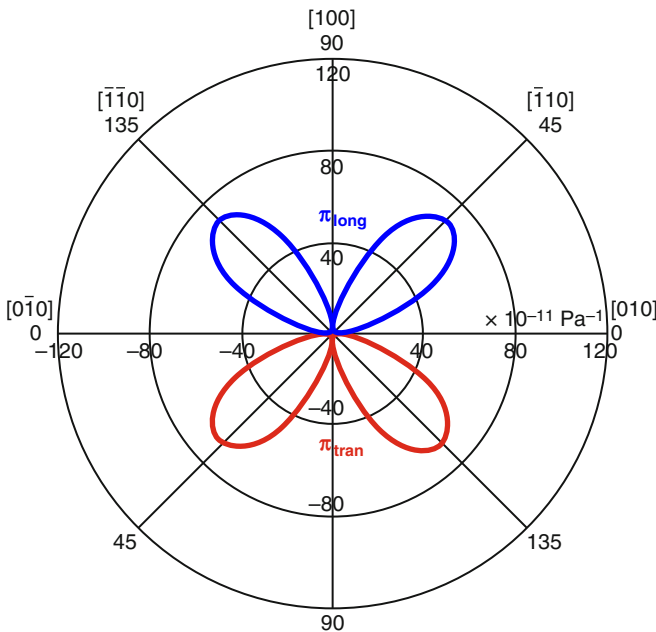
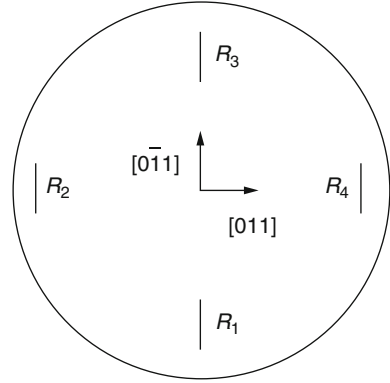
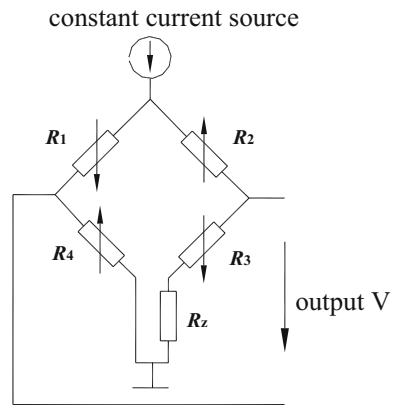


Fig. 2 Piezoresistive coefficient curves in (100) crystal plane

**Fig. 3** Layout scheme of circular diaphragm in (100) crystal plane



**Fig. 4** Wheatstone full bridge consisted of four piezoresistors  $R_1$ – $R_4$  and a zero compensation resistor  $R_z$



diaphragm. Based on Eq. (2), the resistance variations of four piezoresistors are calculated in Eq. (3) when a pressure is applied to the diaphragm:

$$\left\{ \begin{array}{l} \frac{\Delta R_1}{R_1} = \frac{\Delta R_3}{R_3} \approx \frac{1}{2} \pi_{44} (\sigma_{[011]} - \sigma_{[0\bar{1}1]}) \\ \frac{\Delta R_2}{R_2} = \frac{\Delta R_4}{R_4} \approx -\frac{1}{2} \pi_{44} (\sigma_{[011]} - \sigma_{[0\bar{1}1]}) \end{array} \right. \quad (3)$$

where  $\sigma_{[0\bar{1}1]}$  and  $\sigma_{[011]}$  are the stresses of piezoresistors along the crystal directions  $[011]$  and  $[0\bar{1}1]$ , respectively, and  $\pi_{44}$  is the shear piezoresistive coefficient.

The Wheatstone full bridge was consisted of four piezoresistors  $R_1$ – $R_4$  and another resistor  $R_z$  as shown in Fig. 4. In general, the absolute zero output of Wheatstone bridge is not zero and may be larger than 10 mV because the initial values of four piezoresistors cannot be the same under the fabrication errors. The resistor  $R_z$  was used to compensate the zero output of Wheatstone bridge less than 2 mV so that the output signal can be easily amplified.

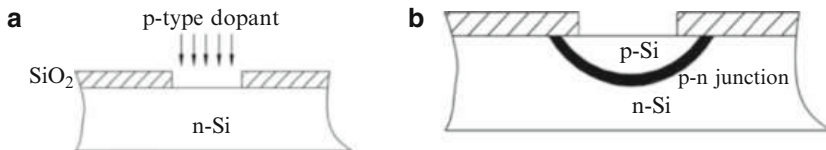
In addition, the Wheatstone full bridge is usually powered with constant current in order to eliminate the effect from temperature coefficients of piezoresistors. Then, the output voltage can be calculated as follows:

$$V_{out} = \left| \frac{R_2 R_4 - R_1 R_3}{R_1 + R_2 + R_3 + R_4} \right| \cdot I_B \tag{4}$$

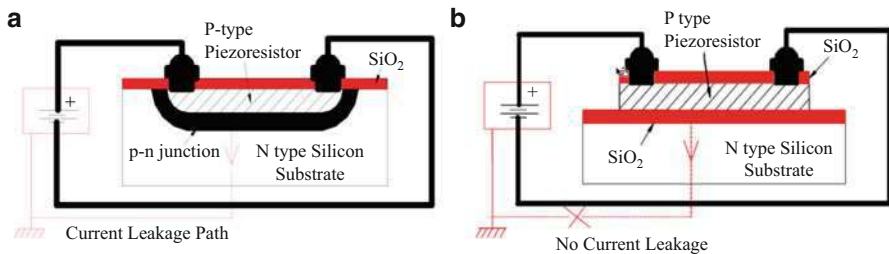
If the initial values of four piezoresistors are equal, that is,  $R_1 = R_2 = R_3 = R_4 = R$ , then the zero output of Wheatstone bridge  $V_{out}$  is zero. When a pressure is applied to the diaphragm, the values of  $R_1, R_3$  are decreased by  $\Delta R$  and the values of  $R_2, R_4$  are increased by  $\Delta R$ . Then, the output voltage of Wheatstone bridge  $V_{out}$  is calculated with the value of  $\Delta R \cdot I_B$  based on Eq. (4). Therefore, the pressure can be measured according to the output voltage of Wheatstone bridge because resistance variation  $\Delta R$  is proportional to the applied pressure (Niu et al. 2014).

For the piezoresistive sensor, the biggest challenge is the temperature influence on sensor performance, because the piezoresistor resistance and piezoresistive coefficient both have thermal coefficients.

For the traditional MEMS piezoresistor pressure sensing chip, the p-type piezoresistors are fabricated by doping and ion implantation technology and isolated by p-n junction from the n-type silicon substrate, as shown in Fig. 5. When the working temperature of sensing chip is beyond 120 °C, the impurity energy level of silicon is close to its intrinsic energy level, and then the leakage current is generated through p-n junction, as shown in Fig. 6a, which will deteriorate the sensing chip to be invalid. For example, leakage current of 1.3 mA will be generated when the piezoresistor is power with 1 mA under 170 °C. In order to solve this problem, a novel method to isolate the p-type piezoresistor by the insulation layer from the Si substrate was presented. As shown in Fig. 6b, a SiO<sub>2</sub> layer was used to isolate the



**Fig. 5** p-type piezoresistor fabricated by doping technology in the traditional piezoresistive sensing chip. (a) p-type dopant to fabricate piezoresistors. (b) p-type piezoresistor isolated by p-n junction



**Fig. 6** Analyses of leakage current when temperature is beyond 120 °C (a) Leakage current generated through p-n junction (b) No p-n junction and no leakage current

piezoresistor from Si substrate, and then this kind of sensing chip can be used under high temperature of 350 °C because it does not generate leakage current. Absolutely, the sensing chip should be designed systematically to guarantee measuring stability and reliability, such as the stress compensation layer design, the piezoresistor structure design, and so on, because different thermal expansion coefficients of various material layers in a wide temperature range will generate large residual stress to deteriorate the sensor performance.

---

## Mechanics Model of the Pressure Sensor Chip

In this part the mechanism model for the piezoresistive pressure sensor will be discussed.

### Diaphragm Deflection Model

Circular and rectangular diaphragm was usually used as the sensing element for the piezoresistive pressure sensor chip. In this part the deflection of the circular and rectangular diaphragm will be introduced first. Based on these two flat diaphragm deflection models, an improved model for structured diaphragm with block was discussed.

#### Circular Diaphragm Deflection Model

The deflection of a circular diaphragm can be equivalent to a mechanical model of circular plate with clamped edges. In this case the slope of the deflection surface in the radial direction must be zero for  $r = 0$  and  $r = a$ . And the deflection of the diaphragm can be presented in Eq. (5) (Timoshenko and Woinowsky-Krieger 1959):

$$w = \frac{q}{64D} (a^2 - r^2)^2 \quad (5)$$

The maximum deflection is at the center of the plate and, from Eq. (5), is equal to

$$w_{\max} = \frac{qa^4}{64D} \quad (6)$$

This deflection is equal to three-eighths of the deflection of a uniformly loaded strip with built-in ends having a flexural rigidity equal to  $D$ , a width of unity, and a length equal to the diameter of the plate.

And the bending moments  $M_r$  and  $M_t$  of the diaphragm can be presented as

$$M_r = \frac{q}{16} [a^2(1 + \nu) - r^2(3 + \nu)] \quad (7)$$

$$M_t = \frac{q}{16} [a^2(1 + \nu) - r^2(1 + 3\nu)] \quad (8)$$

where  $M_r$  and  $M_t$  denote bending moments per unit length. The moment  $M_r$  acts along circumferential sections of the plate.  $M_t$  acts along the diametric section of the plate.

When substituting  $r = a$  in Eq. (7) and (8), the bending moment at the boundary of the plate can be presented as

$$(M_r)_{r=a} = -\frac{qa^2}{8} \tag{9}$$

$$(M_t)_{r=a} = -\frac{\nu qa^2}{8} \tag{10}$$

And the maximum stress is at the boundary of the plate where

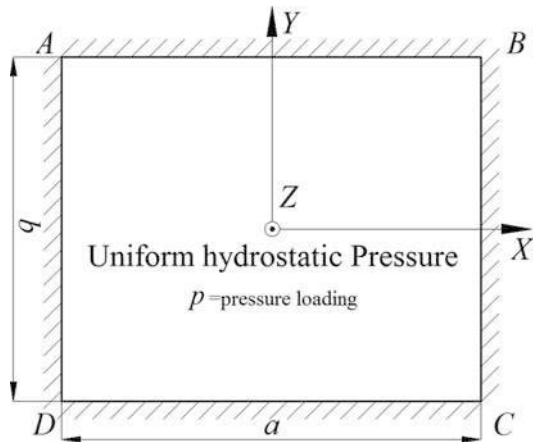
$$(\sigma_r)_{\max} = -\frac{6M_r}{h^2} = \frac{3}{4} \frac{qa^2}{h^2} \tag{11}$$

### Rectangular Diaphragm Deflection Model

Rectangular diaphragm is also widely used in piezoresistive pressure sensor. And its mechanics model can be equivalent to a rectangular plate with clamped edges. Any four-edge fixed diaphragm can be equivalent to a four-edge simply supported diaphragm with additional bending moments applied to its edges. Thus, the deflection of a four-edge fixed diaphragm under any axial load can be obtained based on the corresponding four-edge simply supported diaphragm under the same load with additional bending moments to the edges, in order to offset the rotation along the four edges.

Four-edge fixed rectangular diaphragm under a pressure in Fig. 7 can be decomposed into two models. The first model is a simply supported rectangular diaphragm under a pressure. The second model is a simply supported rectangular diaphragm under the bending moment distributing along four edges to offset the rotations.

**Fig. 7** Four-edge fixed diaphragm loaded with a pressure load



Based on the Navier double trigonometric series (Timoshenko and Woinowsky-Krieger 1959), the deflection of a simply supported diaphragm under a pressure is presented as

$$w_{sp} = \frac{4pa^4}{\pi^5 D} \sum_{m=1,3,5,\dots}^N \frac{(-1)^{(m-1)/2}}{m^5} \cos \frac{m\pi x}{a} \cdot \left( 1 - \frac{\alpha_m \tanh \alpha_m + 2}{2 \cosh \alpha_m} \cosh \frac{m\pi y}{a} + \frac{1}{2 \cosh \alpha_m} \frac{m\pi y}{a} \sinh \frac{m\pi y}{a} \right) \quad (12)$$

$$\alpha_m = \frac{m\pi b}{2a} \quad (13)$$

where  $p$  is the pressure,  $a$  and  $b$  are the length and width of the diaphragm,  $D$  is the flexural stiffness of the diaphragm,  $D = Eh^3/[12(1-\mu^2)]$ ,  $E$  is the Young modulus,  $\mu$  is the Poisson ratio, and  $h$  is the diaphragm thickness.

The rotation along the edge AB caused by the pressure can be represented as

$$\begin{aligned} (dw_{sp})_{y=b/2} &= \left( \frac{\partial w_{sp}}{\partial y} \right)_{y=b/2} \\ &= \frac{2qa^3}{\pi^4 D} \cdot \sum_{m=1,3,5,\dots}^N \frac{(-1)^{(m-1)/2}}{m^4} \cos \frac{m\pi x}{a} \left( \frac{\alpha_m}{\cosh^2 \alpha_m} - \tanh \alpha_m \right) \end{aligned} \quad (14)$$

Similarly, the rotation along the edge BC by the pressure can be represented as

$$(dw_{sp})_{x=a/2} = \frac{2qb^3}{\pi^4 D} \cdot \sum_{m=1,3,5,\dots}^N \frac{(-1)^{(m-1)/2}}{m^4} \cos \frac{m\pi y}{b} \left( \frac{\beta_m}{\cosh^2 \beta_m} - \tanh \beta_m \right) \quad (15)$$

In order to offset the bending rotation along the diaphragm edges of a simply supported diaphragm, to form the four-edge clamped working condition, two pairs of bending moments shall be loaded on the diaphragm. Here, the two bending moments loaded on the diaphragm edges can be presented as

$$(M_y)_{y=\pm b/2} = \sum_{m=1,3,5,\dots}^N (-1)^{(m-1)/2} E_m \cos \frac{m\pi x}{a} \quad (16)$$

$$(M_x)_{x=\pm a/2} = \sum_{m=1,3,5,\dots}^N (-1)^{(m-1)/2} F_m \cos \frac{m\pi y}{b} \quad (17)$$

where  $E_m$  and  $F_m$  are undetermined coefficients

Then, the diaphragm deflections under different bending moments are represented, respectively, as



$$w_{my}(x, y, a, b) = -\frac{a^2}{2\pi^2 D} \sum_{m=1,3,5\dots}^N E_m \frac{(-1)^{(m-1)/2}}{m^2 \cosh \alpha_m} \cos \frac{m\pi x}{a} \left( \frac{m\pi y}{a} \sinh \frac{m\pi y}{a} - \alpha_m \tanh \alpha_m \cosh \frac{m\pi y}{a} \right) \tag{18}$$

$$w_{mx}(x, y, a, b) = -\frac{b^2}{2\pi^2 D} \sum_{m=1,3,5\dots}^N F_m \frac{(-1)^{(m-1)/2}}{m^2 \cosh \beta_m} \cos \frac{m\pi y}{b} \left( \frac{m\pi x}{b} \sinh \frac{m\pi x}{b} - \beta_m \tanh \beta_m \cosh \frac{m\pi x}{b} \right) \tag{19}$$

$$\beta_m = \frac{a^2}{b^2} \alpha_m \tag{20}$$

In order to ensure the form of the solution is consistent with the form of Eqs. (14) and (15), the rotations on the edges  $A''B''$  and  $B''C''$  contributed by the moments are represented as (Timoshenko and Woinowsky-Krieger 1959)

$$\begin{aligned} (dw_M)_{y=b/2} &= \left( \frac{\partial w_{my}}{\partial y} \right)_{y=b/2} + \left( \frac{\partial w_{mx}}{\partial y} \right)_{y=b/2} \\ &= -\frac{a}{2\pi D} \sum_{m=1,3,5\dots}^N E_m \frac{(-1)^{(m-1)/2}}{m} \cos \frac{m\pi x}{a} \left( \tanh \alpha_m + \frac{\alpha_m}{\cosh^2 \alpha_m} \right) \\ &\quad - \frac{4a^2}{\pi^2 D b} \sum_{m=1,3,5\dots}^N \frac{F_m}{m^3} \sum_{i=1,3,5\dots}^N \frac{i(-1)^{(i-1)/2}}{\left( \frac{a^2}{b^2} + \frac{i^2}{m^2} \right)^2} \cos \frac{i\pi x}{a} \end{aligned} \tag{21}$$

$$\begin{aligned} (dw_M)_{x=a/2} &= \left( \frac{\partial w_{mx}}{\partial x} \right)_{x=a/2} + \left( \frac{\partial w_{my}}{\partial x} \right)_{x=a/2} \\ &= -\frac{b}{2\pi D} \sum_{m=1,3,5\dots}^N F_m \frac{(-1)^{(m-1)/2}}{m} \cos \frac{m\pi y}{b} \left( \tanh \beta_m + \frac{\beta_m}{\cosh^2 \beta_m} \right) \\ &\quad - \frac{4b^2}{\pi^2 D a} \sum_{m=1,3,5\dots}^N \frac{E_m}{m^3} \sum_{i=1,3,5\dots}^N \frac{i(-1)^{(i-1)/2}}{\left( \frac{b^2}{a^2} + \frac{i^2}{m^2} \right)^2} \cos \frac{i\pi y}{b} \end{aligned} \tag{22}$$

And based on the boundary condition of the four-edge fixed diaphragm, the relationships among Eqs. (14), (21), (15), and (22) are represented as

$$(dw_{sp})_{y=b/2} = (dw_M)_{y=b/2} \tag{23}$$

$$(dw_{sp})_{x=a/2} = (dw_M)_{x=a/2} \tag{24}$$

Based on Eqs. (23) and (24), a linear matrix equation is also established.

Thus,  $E_m$  and  $F_m$  for compensating the rotation along the edges of simply supported rectangular diaphragm under a pressure can be derived from Eqs. (43) and (24). Base on the deformation superimposing theory, substituting  $E_m$  and  $F_m$  to Eqs. (18) and (19), the deflection of four-edge fixed rectangular diaphragm under a pressure can be presented as

$$w_{fp} = w_{sp} - w_{my}(x, y, a, b) - w_{mx}(x, y, a, b) \quad (25)$$

The diaphragm stress along X and Y direction under a pressure can be calculated by Eqs. (26) and (27):

$$\sigma_x = -\frac{Ez}{1 - \mu^2} \left( \frac{\partial^2 w}{\partial^2 x} + \mu \frac{\partial^2 w}{\partial^2 y} \right) \quad (26)$$

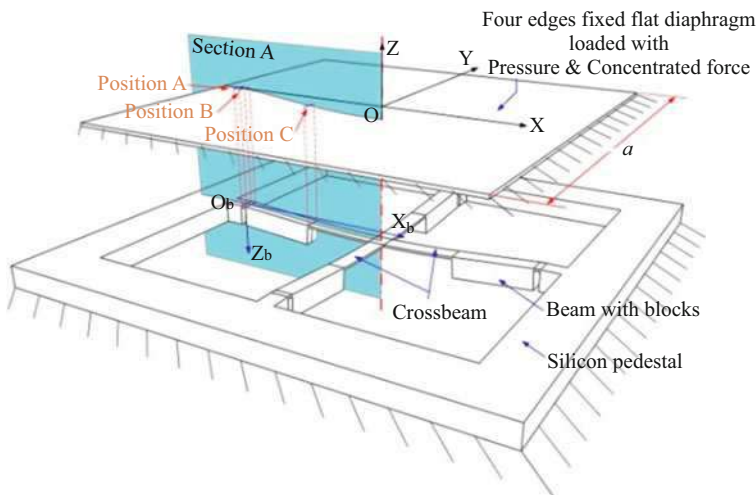
$$\sigma_y = -\frac{Ez}{1 - \mu^2} \left( \frac{\partial^2 w}{\partial^2 y} + \mu \frac{\partial^2 w}{\partial^2 x} \right) \quad (27)$$

where  $w$  is the diaphragm deflection and  $z$  is the vertical distance to the neutral plane.

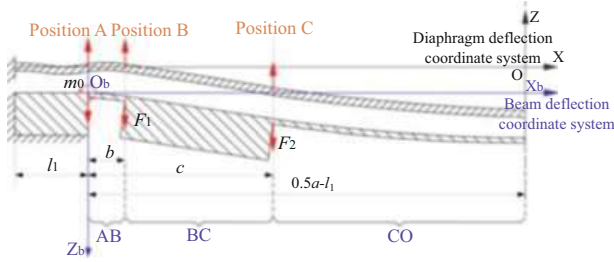
### Structured Diaphragm with Block Deflection Model

A beam-diaphragm coupled model was presented to describe the deformation mechanism of the diaphragm combined with block structures such as peninsula-island structure. And the schematic of the model was shown in Fig. 8.

In the beam-diaphragm coupled model, the effect from the peninsula-island structure to the deformation of the diaphragm was simplified as the interaction force from a beam consisting of two islands to the diaphragm, as shown in Fig. 9. The width of the beam was the same with the width of the peninsula-island structure. And the size of the islands on the beam was also the same as the size of the peninsula-island structure. The interaction force from the beam to the diaphragm can be simplified as three pairs of centrosymmetric concentration forces applying to positions A, B, and C, respectively, which were positioned at the ends of the



**Fig. 8** Schematic of the beam-diaphragm coupled model



**Fig. 9** Schematic of the beam-diaphragm coupled model in section A

peninsula structure and island structure, as shown in Fig. 9. So the beam and the diaphragm can be regarded as bonded together through A, B, and C position. Once a pressure was loaded on the diaphragm, the interaction force would develop at these positions.

According to the displacement compatibility between the diaphragm and beam at the A, B and C position, the deformation of the diaphragm can be determined. Once the deformation of the diaphragm is determined, the stress value at the SCR can be derived from the first derivative of the diaphragm deformation curve along the longitudinal direction of the peninsula-island structure.

The loading condition of the diaphragm can be seemed as the superimposing of three basic loading conditions. The first was a loading condition of simply supported diaphragm loaded with a uniform pressure, which presented the effect of the pressure loaded on the front of the diaphragm. The second was a loading condition of simply supported diaphragm loaded with centrosymmetric concentration force, which presented the effect of the interaction force from the beam model. The third was a loading condition of simple supported diaphragm loaded with bending moment along four edges, which was used to offset the rotation at the diaphragm edges to ensure the diaphragm edges are in a fixed boundary condition.

**Simply Supported Diaphragm Loaded with a Uniform Pressure**

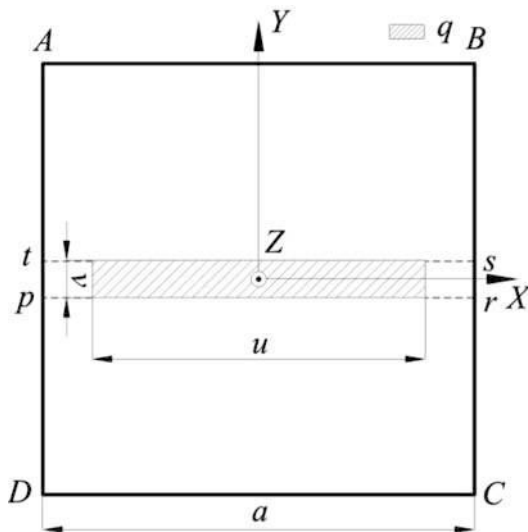
The deformation of a simply supported diaphragm under a pressure was presented in Eq. (12). The rotation along one of the four edges caused by the pressure was presented in Eq. (14).

**Simply Supported Diaphragm Loaded with Centro-Symmetric Concentration Force**

The loading condition of simply supported diaphragm loaded with centrosymmetric concentration force can be derived by superimposing the loading condition of simply supported rectangular diaphragm loaded with partial load, as shown in Fig. 10.

And the partial differential equation governing the deformation and its general solution were shown as

**Fig. 10** Simply supported rectangular diaphragm loaded with partial load



$$\frac{\partial^4 w}{\partial x^4} + 2 \frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} = \frac{q_0(x, y)}{D} \tag{28}$$

$$q(x, y) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{a} \tag{29}$$

$$a_{mn} = \frac{4}{a^2} \int_{-a/2}^{a/2} \int_{-a/2}^{a/2} q_0(x, y) \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{a} dx dy \tag{30}$$

where  $w$  was the diaphragm deformation,  $q_0(x, y)$  was the original load to the diaphragm, and  $q(x, y)$  was the Navier trigonometric expression of  $q_0(x, y)$  which was used to substitute the original load in Eq. (28) to improve the solvability of Eq. (28).

For the loading condition of simply supported rectangular diaphragm loaded with partial load, the  $q(x, y)$  presented in Eq. (29) can be presented as

$$Q(q, x, y, a, u) = \frac{4q}{\pi} \sum_{m=1,3,5,\dots}^N \frac{(-1)^{m-1}}{m} \sin \frac{m\pi u}{2a} \cos \frac{m\pi x}{a} \tag{31}$$

where  $q$  was the loading intensity and  $u$  and  $v$  were the length and width of the area  $tsrp$ .

And the diaphragm deformation in area  $tsrp$  can be presented as

$$w_{s1}(q, x, y, a, u, v) = \sum_{m=1,3,5}^N (-1)^{(m-1)/2} \cos \frac{m\pi x}{a} \left( a_m + A_m \cosh \frac{m\pi y}{a} + B_m \frac{m\pi y}{a} \sinh \frac{m\pi y}{a} \right) \tag{32}$$

$$\begin{cases} A_m = -\frac{a_m}{\cosh \alpha_m} \left[ \cosh(\alpha_m - 2\gamma_m) + \gamma_m \sinh(\alpha_m - 2\gamma_m) + \alpha_m \frac{\sinh 2\gamma_m}{2 \cosh \alpha_m} \right] \\ B_m = \frac{a_m}{2 \cosh \alpha_m} \cosh(\alpha_m - 2\gamma_m) \end{cases} \quad (33)$$

$$\gamma_m = \frac{m\pi v}{4a} \quad (34)$$

Considering the series convergence, the deformation in  $ABst$  area was derived by the method of mirror image presented in Timoshenko and Woinowsky-Krieger (1959). And the diaphragm deformation in area  $ABst$  can be presented as

$$w_{s2}(q, x, y, a, u, v) = w_{\text{munit}}(q, x, y, a, u, v) + \sum_{i=1}^M (-1)^i \quad (35)$$

$$[w_{\text{munit}}(q, x, ia - y, a, u, v) + w_{\text{munit}}(q, x, ia + y, a, u, v)]$$

$$\begin{aligned} w_{\text{munit}}(q, x, y, a, u, v) = \\ \frac{qa^3}{\pi^4 D} \sum_{m=1,3,5}^N (-1)^{(m-1)/2} \cos \frac{m\pi x}{a} \frac{1}{m^4} \sin \frac{m\pi}{2} \sin \frac{m\pi u}{2a} \\ \left[ \left( \frac{2a}{m\pi} + y - \frac{v}{2} \right) e^{-\frac{m\pi(2y-v)}{2a}} - \left( \frac{2a}{m\pi} + y + \frac{v}{2} \right) e^{-\frac{m\pi(2y+v)}{2a}} \right] \end{aligned} \quad (36)$$

The first derivative for  $y$  of Eq. (35) can be represented as

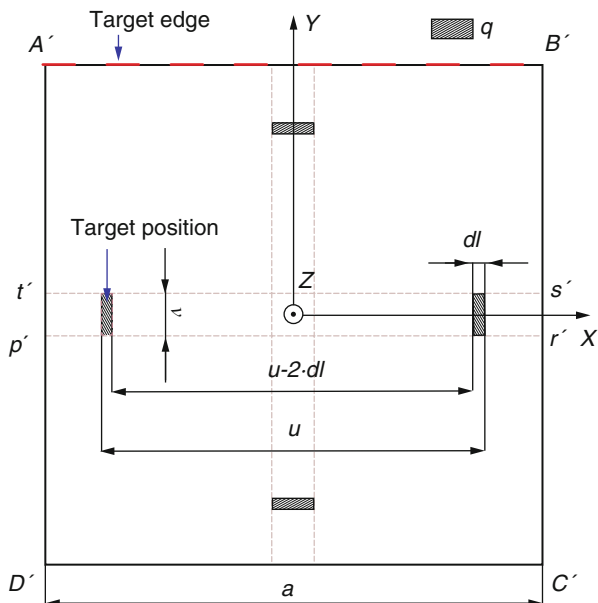
$$\begin{aligned} dw_{s2}(q, x, y, a, u, v) = dw_{\text{munit}}(q, x, y, a, u, v) \\ + \sum_{i=1}^M (-1)^{i-1} [dw_{\text{munit}}(q, x, ia - y, a, u, v) - dw_{\text{munit}}(q, x, ia + y, a, u, v)] \end{aligned} \quad (37)$$

$$\begin{aligned} dw_{\text{munit}}(q, x, y, a, u, v) &= \frac{\partial w_{\text{munit}}(x, y, a, b, u, v)}{\partial y} \\ &= \frac{qa^3}{\pi^4 D} \sum_{m=1,3,5}^N (-1)^{(m-1)/2} \cos \frac{m\pi x}{a} \\ &\quad \cdot \frac{1}{m^4} \sin \frac{m\pi}{2} \sin \frac{m\pi u}{2a} \left\{ \left[ 1 - \frac{m\pi}{a} \left( \frac{2a}{m\pi} + y - \frac{v}{2} \right) \right] e^{-\frac{m\pi(2y-v)}{2a}} \right. \\ &\quad \times \left. - \left[ 1 - \frac{m\pi}{a} \left( \frac{2a}{m\pi} + y + \frac{v}{2} \right) \right] e^{-\frac{m\pi(2y+v)}{2a}} \right\} \end{aligned} \quad (38)$$

where  $M$  was the total amount of unit loading condition. The larger the value of  $M$  was, the more accurate the result can be.

When two simply supported rectangular diaphragms under different partial loads were superimposed together, the loading condition of simply supported rectangular diaphragm with axisymmetric concentration force can be obtained. And the loading

**Fig. 11** Simply supported rectangular diaphragm loaded with centrosymmetric concentration force



condition of simply supported diaphragm loaded with centrosymmetric concentration force can be obtained by superimposing the loading condition of simply supported rectangular diaphragm with axisymmetric concentration force, as shown in Fig. 11.

Considering the central symmetry of the proposed structure, only the displacement and rotation at the target places, as shown in Fig. 11, were investigated.

The diaphragm displacement at the target position shown in Fig. 11 can be presented as

$$\begin{aligned}
 w_f(q, x, y, a, u, v) &= [w_{s1}(q, x, y, a, u, v) - w_{s1}(q, x, y, a, u - 2dl, v)] \\
 &\quad + [w_{s1}(q, x, y, a, v, u) - w_{s1}(q, x, y, a, v, u - 2dl)] \quad (39) \\
 x &\in \left(-\frac{a}{2}, \frac{a}{2}\right) \quad y \in \left(-\frac{v}{2}, \frac{v}{2}\right)
 \end{aligned}$$

And the rotation along the  $A'B'$  edge can be presented as:

$$\begin{aligned}
 dw_f(q, x, a, u, v) &= [dw_{s2}(q, x, 0.5a, a, u, v) - dw_{s2}(q, x, 0.5a, a, u - 2dl, v)] \\
 &\quad + [dw_{s2}(q, x, 0.5a, a, v, u) - dw_{s2}(q, x, 0.5a, a, v, u - 2dl)] \quad x \in \left(-\frac{a}{2}, \frac{a}{2}\right) \quad (40)
 \end{aligned}$$

**Simply Supported Diaphragm Loaded with Bending Moment**

Considering the fixed boundary condition of the diaphragm, the diaphragm rotation along the four edges should be kept zero. In order to offset the rotation along the diaphragm edges generated by pressure and interaction force, bending moments were introduced to the four edges. The bending moments loaded on each edge can be presented in Eqs. (16) and (17).

Then, the diaphragm deformation under the bending moments was represented as

$$w_m(x, y, a) = -\frac{a^2}{2\pi^2 D} \sum_{m=1,3,5\dots}^N E_m \frac{(-1)^{(m-1)/2}}{m^2 \cosh \alpha_m} \cos \frac{m\pi x}{a} \left( \frac{m\pi y}{a} \sinh \frac{m\pi y}{a} - \alpha_m \tanh \alpha_m \cosh \frac{m\pi y}{a} \right) - \frac{a^2}{2\pi^2 D} \sum_{m=1,3,5\dots}^N E_m \frac{(-1)^{(m-1)/2}}{m^2 \cosh \alpha_m} \cos \frac{m\pi y}{a} \left( \frac{m\pi x}{a} \sinh \frac{m\pi x}{a} - \alpha_m \tanh \alpha_m \cosh \frac{m\pi x}{a} \right) \tag{41}$$

And the rotation along one of the four edges can be presented as

$$(dw_M)_{y=b/2} = -\frac{a}{2\pi D} \sum_{m=1,3,5\dots}^N E_m \frac{(-1)^{(m-1)/2}}{m} \cos \frac{m\pi x}{a} \left( \tanh \alpha_m + \frac{\alpha_m}{\cosh^2 \alpha_m} \right) - \frac{4a}{\pi^2 D} \sum_{m=1,3,5\dots}^N \frac{E_m}{m^3} \sum_{i=1,3,5\dots}^N \frac{i(-1)^{(i-1)/2}}{\left(1 + \frac{i^2}{m^2}\right)^2} \cos \frac{i\pi x}{a} \quad x \in \left(-\frac{a}{2}, \frac{a}{2}\right) \tag{42}$$

**Superimposing for the Loading Condition of Four Edges Fixed Diaphragm Loaded with Centro-Symmetric Concentration Force**

As shown in Fig. 12, by superimposing the three loading condition presented in sections “Introduction,” “The Piezoresistive Effect of High Temperature Pressure Sensor,” and “Mechanics Model of the Pressure Sensor Chip,” the condition of the four-edge diaphragm loaded with centro-symmetric concentration force was obtained. And diaphragm deformation can be presented as

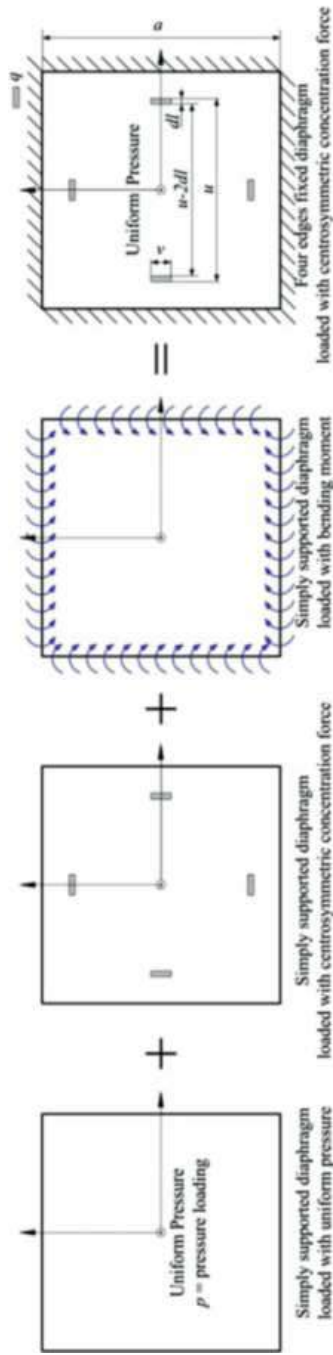
$$w_c(p, q, x, y, a, u, v) = w_p(p, x, y, a) - w_f(q, x, y, a, u, v) - w_m(x, y, a) \quad x \in \left(-\frac{a}{2}, \frac{a}{2}\right) \quad y \in \left(-\frac{v}{2}, \frac{v}{2}\right) \tag{43}$$

which was the basic loading condition for diaphragm loaded with the pressure and the interaction force from the beam (Timoshenko and Woinowsky-Krieger 1959).

Base on the beam-diaphragm coupled model, shown in Figs. 8 and 9, the interaction from the beam to the diaphragm can be simplified as three pairs of centrosymmetric concentrated forces, as shown in Fig. 13. And the deformation of the diaphragm can be presented as

$$w_d(p, q, x, y, a, u, v) = w_p(p, x, y, a) - w_f \left( \frac{F_3}{v \cdot dl}, x, y, a, a - 2l_1 + 2dl, l_w \right) - w_f \left( \frac{F_2}{v \cdot dl}, x, y, a, a - 2l_1 - 2b, l_w \right) - w_f \left( \frac{F_1}{v \cdot dl}, x, y, a, a - 2l_2 + 2dl, l_w \right) - w_m(x, y, a) \quad x_i \in \left(-\frac{a}{2}, \frac{a}{2}\right) \quad y \in \left(-\frac{l_w}{2}, \frac{l_w}{2}\right) \tag{44}$$

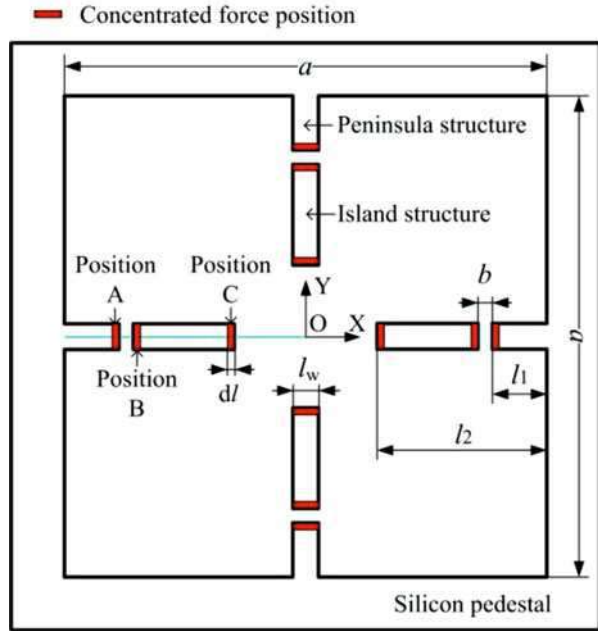
where  $b$  was the distance between the island and peninsula structure and  $l_w$  was the width of the peninsula-island structure.



**Fig. 12** The composition of the loading condition of four-edge fixed rectangular plate under centrosymmetric concentration force



**Fig. 13** Loading positions of the centrosymmetric concentrated forces



**Interaction Between Beam and Diaphragm**

Considering the width of the peninsula-island structure was very small compared to the dimension of diaphragm edge, the deformation of the crossbeam shown in Fig. 8 can be simplified as a two-sided fixed beam structure, as shown in Fig. 9. The deflection of the beam model under the interaction forces  $F_1$  and  $F_2$  from the diaphragm can be derived from Eqs. (45), (46), and (47), which is under the boundary conditions presented in Eq. (48) (Bao 2000);

$$EI_1 w_1'' = -m_0 + (F_1 + F_2)x_b \quad x_b \in [0, b] \tag{45}$$

$$EI_2 w_2'' = -m_0 + F_1 b + F_2 x_b \quad x_b \in [b, c] \tag{46}$$

$$EI_1 w_3'' = -m_0 + (F_1 + F_2)c \quad x_b \in [c, 0.5a - l_1] \tag{47}$$

$$\begin{cases} w_1(0) = 0 & , & w_1'(0) = 0 \\ w_1(b) = w_2(b) & , & w_1'(b) = w_2'(b) \\ w_2(c) = w_3(c) & , & w_2'(c) = w_3'(c) \\ w_3'(0.5a - l_1) = 0 \end{cases} \tag{48}$$

where  $w_1$ ,  $w_2$ , and  $w_3$  were the deformations for the beam segments AB, BC, and CO, respectively,  $m_0$  was the bending moment at the fixed end of the beam model,  $I_1$  was the moment of inertia for segments AB and CO, and  $I_2$  was the moment of inertia for segment BC.

By combining the diaphragm and beam deformation model together, the beam-diaphragm coupled model was established.

Based on the displacement compatibility between the beam and diaphragm deformation model at A, B, and C position, shown in Fig. 9, the interaction forces of  $F_1$ ,  $F_2$ , and  $F_3$  presented in Eq. (44) were determined by the following boundary condition:

$$\begin{cases} (w_d)_{x=-(0.5a-l_1)} = 0 \\ (w_d)_{x=-(0.5a-l_1-b)} = (w_1)_{x_b=b} \\ (w_d)_{x=-(0.5a-l_1-c)} = (w_3)_{x_b=c} \end{cases} \quad (49)$$

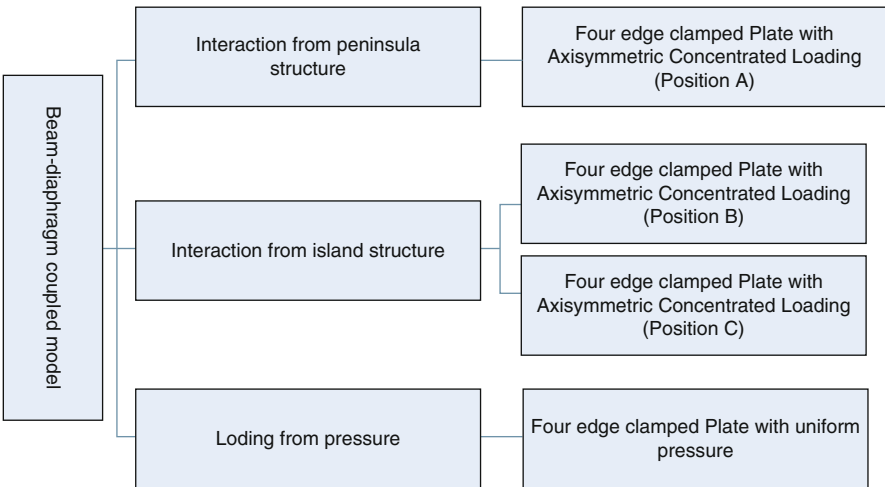
By substituting the values of  $F_1$  and  $F_2$  to the beam deformation model, the stress value at the gap position can be derived by the deformation of the beam structure as

$$\sigma_{x_b} = -0.5Eh \frac{\partial^2 w_1}{\partial x_b^2} \quad x_b \in [0, b] \quad (50)$$

The systematical diagram of loading condition for the beam-diaphragm coupled model of the sensor chip can be concluded in Fig. 14.

### Theory Verification

In order to validate the analytical solutions obtained above, simulations are carried out using commercially available FEM software (ANSYS15.0, ANSYS Inc.).



**Fig. 14** Systematical diagram for the sensor's loading condition

### Discussion of the Basic Loading Condition

The four-edge fixed diaphragm separately loaded with the axisymmetric concentrated force or a pressure is the basic model to analyze the proposed novel diaphragm structure. Because the bending moment and stress of the diaphragm are both derived from the deflection, the diaphragm deflections under different loads are the first to be discussed. The detailed material properties and the geometry parameters are shown in Table 1 (Maluf 2000; Yu and Lan 2001; Hopcroft et al. 2010).

The four-edge fixed diaphragm loaded with axisymmetric concentrated forces was shown in Fig. 15; the deflections of the diaphragm in different cross sections are obtained by Eqs. (12) and (14) as well as FEM simulation as shown in Fig. 15. Figure 16 shows its deflection results in different cross sections from the theoretical calculation and FEM simulation for the four-edge fixed diaphragm loaded with centrosymmetric concentrated forces. And the theoretical results are derived by superimposing the deflections of four-edge fixed diaphragms under different axisymmetric concentrated forces together. The deflection for four-edge fixed diaphragm loaded with the pressure results in different cross sections are shown in Fig. 17.

Figure 18 shows the relative errors of the deflections between the theoretical calculation and FEM simulation across the section line  $y = 0$  in Figs. 15, 16 and 17. Which indicated that the relative errors are minimal. The relative errors close to the edge of the diaphragm are somewhat larger, because of the boundary error caused by

**Table 1** Material properties and the geometry parameters for the basic loading condition

	Four-edge fixed diaphragm loaded with axisymmetric concentrated force	Four-edge fixed diaphragm loaded with centrosymmetric concentrated force	Four-edge fixed diaphragm loaded with uniform pressure
Geometry parameters			
$a$ (m)	3500e-6	3500e-6	3500e-6
$b$ (m)	2000e-6	3500e-6	2000e-6
$u$ (m)	3140e-6	2900e-6	/
$v$ (m)	160e-6	160e-6	/
$dl$ (m)	1e-6	1e-6	/
$h$ (m)	10e-6	10e-6	10e-6
Material property			
$E$ (Pa)	1.6e11	1.6e11	1.6e11
$\gamma$ (Pa)	0.278	0.278	0.278
$\rho$ (kg/m <sup>3</sup> )	2330	2330	2330
Loading condition			
$F$ (N)	0.001	0.001	/
$q$ (Pa)	$F/(v \cdot dl)$	$F/(v \cdot dl)$	500

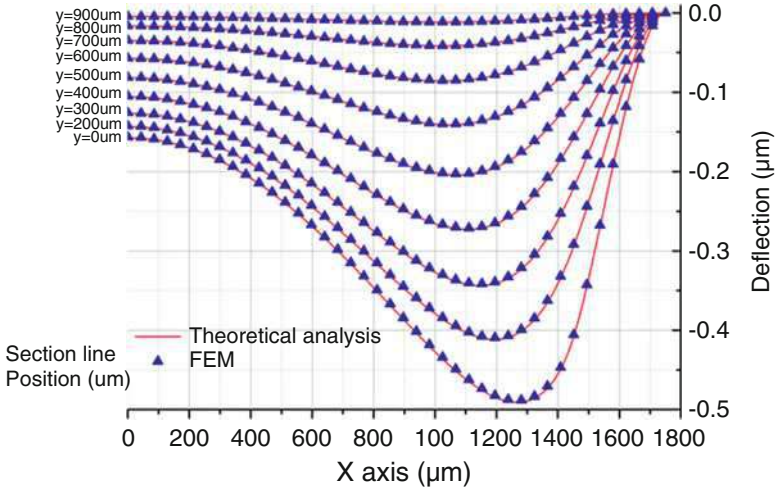


Fig. 15 Deflection results of four-edge fixed diaphragm loaded

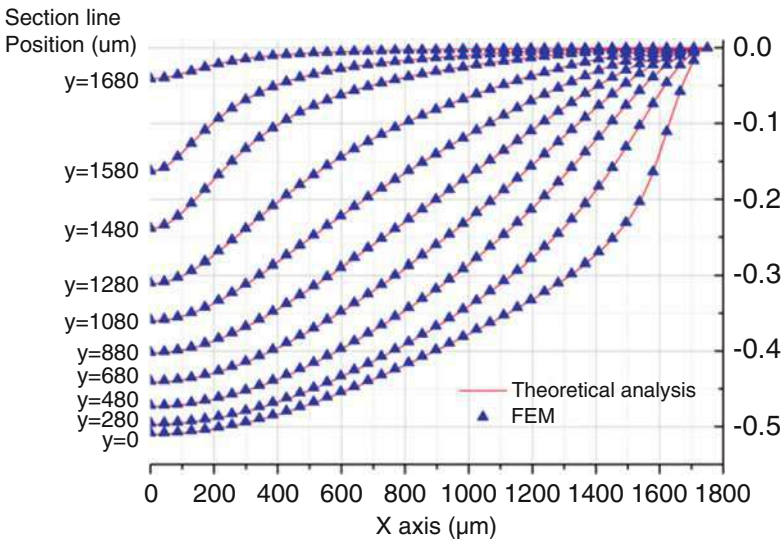


Fig. 16 Deflection results of four-edge fixed diaphragm

trigonometric series. But the errors are still less than 2%. Thus, it demonstrates that the mechanics theory is sufficiently accurate.

**Discussion of the Structured Diaphragm with Blocks**

Here the theory for the structured diaphragm with blocks was verified by FEM and makes a comparison to the models presented in Bao (2000) and Xu et al. (2016a).

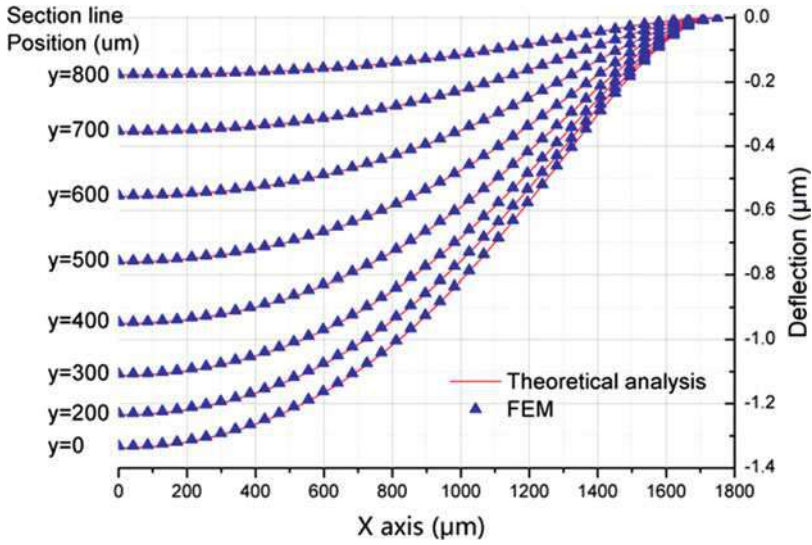


Fig. 17 Deflection results by theory and FEM for the four-edge fixed diaphragm loaded with the pressure load

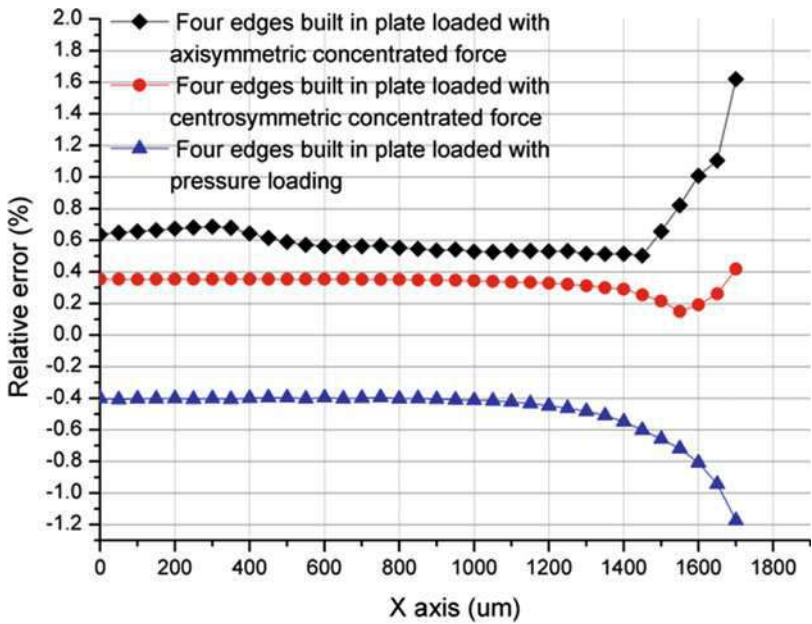
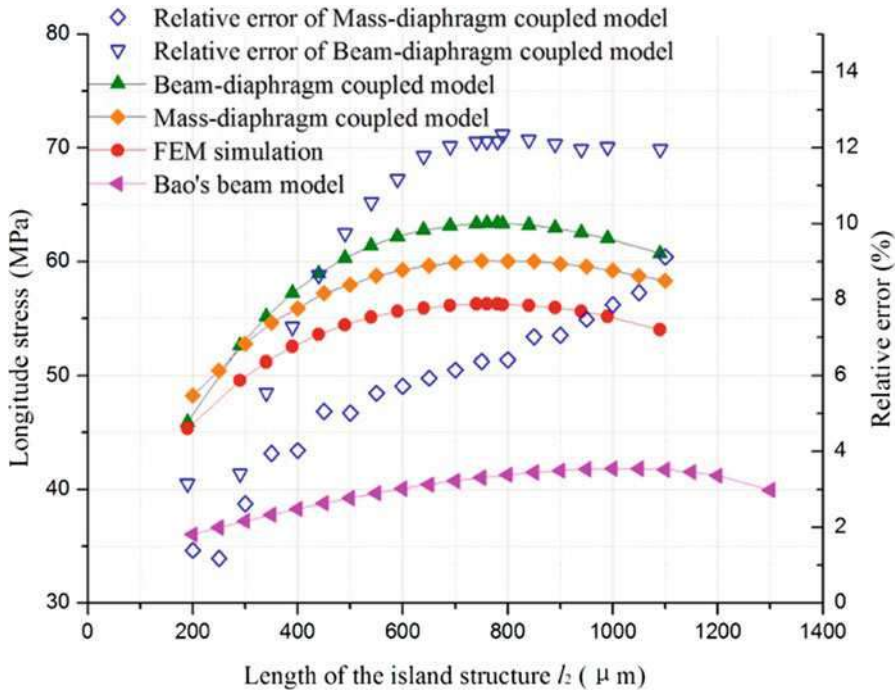


Fig. 18 Relative errors between theory and FEM results



**Fig. 19** The relationships between longitudinal stress at the SCR and  $l_2$  when  $l_1$  was 175  $\mu\text{m}$

Figure 19 indicated that the analytic results of the beam-diaphragm coupled model were in accordance to the finite element method (FEM) results, and the maximum relative error was less than 13% which was a little higher than that of the mass-diaphragm coupled model presented in Xu et al. (2016a). However, it could reduce the calculation amount and improve the quality of the solution derived from the system of linear equation. Also the beam-diaphragm coupled model could predict the optimized island length as the mass-diaphragm coupled model did, and the accuracy of the beam-diaphragm coupled model was greatly improved compared to the most widely used beam model theory presented by Bao (2000).

## Structure Designing, Lithography Mask Designing and Fabrication of the Sensor Chip

In this part, the structure designing and corresponding lithography mask designing method were presented.

In the field of piezoresistive pressure sensor designing, there are two problems that hinder the development of the micromachined piezoresistive pressure sensor

chip. One is the tradeoff between the sensitivity and linearity, and the other was the strain energy dissipation outside the stress concentration region (SCR). In order to overcome these two problems, a well-designed structured diaphragm should be presented.

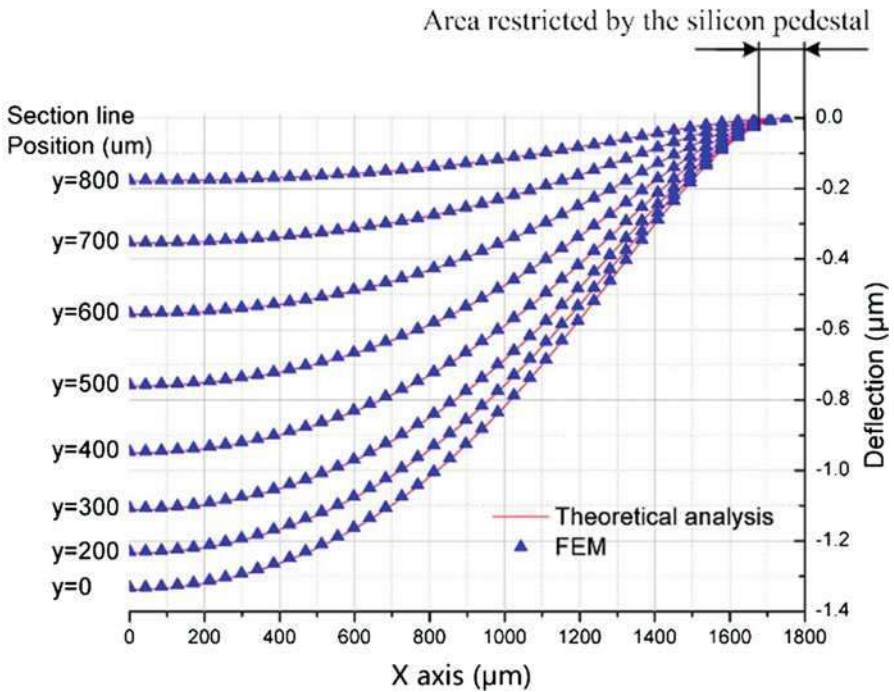
### Structure Designing

The designing principle of the sensor chip can be presented as below:

**Step 1:** Lowering the constrain from the structure

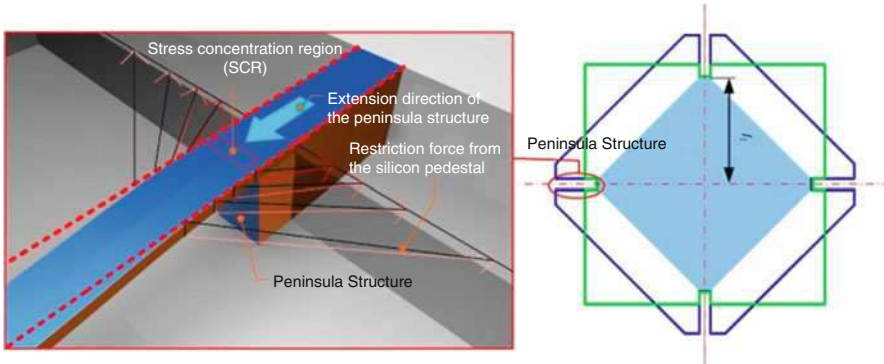
According to the plate and shell theory (Timoshenko and Woinowsky-Krieger 1959), the silicon pedestal has much constrain on the deformation at the edge of the diaphragm. And the curvature at the edge of the diaphragm was nearly zero. So the region close to the diaphragm can be regarded as silicon pedestal restriction region, as shown in Fig. 20.

In order to get rid of the constrain from the silicon pedestal to the SCR, a peninsula structure can be introduced to the sensor chip, as shown in Fig. 21. The peninsula structure makes the strain energy concentrated at the tip of the peninsula



**Fig. 20** Deflection results by theory and FEM for the four-edge fixed diaphragm loaded with the pressure load





**Fig. 21** The working mechanism of using peninsula structure to relieve the constrain on SCR from silicon pedestal

structure. By this means, the SCR was positioned in a place with a distance from the silicon pedestal. Then the constrain restricted on the deformation at the SCR was largely relieved.

**Step 2:** Create stiffness mutation to enhance the stress concentration at SCR

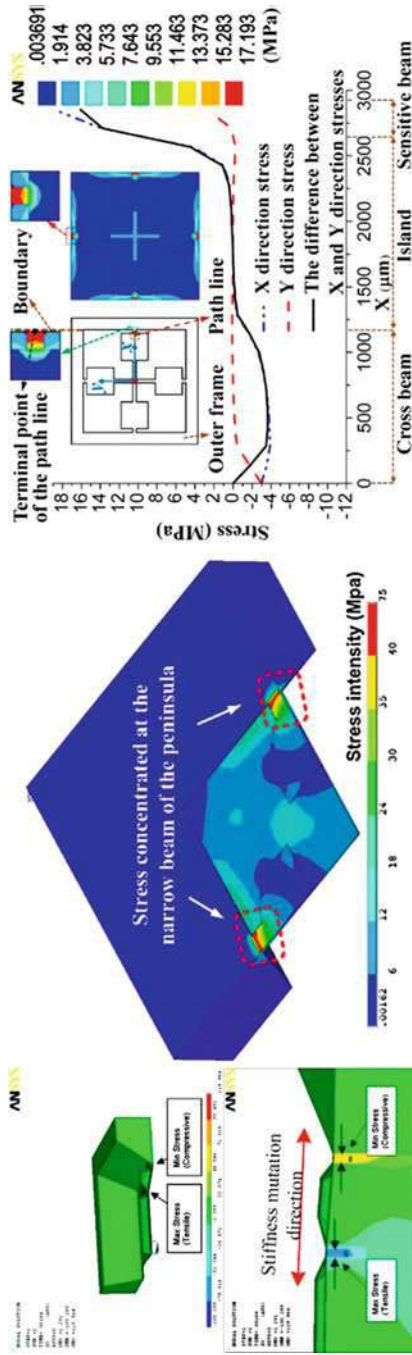
According to the test results presented in paper Kinnell et al. (2010), Xian and Zhang (2014), and Yu et al. (2015a), as shown in Fig. 22, those with the smaller SCR area perform better in pressure sensitivity. This indicated that designing a structure with small SCR area and less strain energy dissipation around SCR can help the piezoresistor to gain more strain energy intensity. Then the piezoresistor can take full use of the strain energy and convert them into electrical signal output.

Because the stress distributed on the diaphragm was in direct proportion to the bending moment distributed in the diaphragm and the bending moment was in direct proportion to the second derivative of the deflection in a certain direction, a structure should be designed to upgrade increase the changing rate of the first derivative in a certain direction. And a stiffness mutation structure such as the peninsula-island structure shown in Fig. 23 can be presented to reach this target (Xu et al. 2016a, b; Zhao et al. 2016). Assuming the slope of the island structure was fixed, the changing rate of the deflection derivative in longitudinal direction gets bigger when the gap width becomes smaller. And the smallest size of the gap width can be determined by the piezoresistor size.

Here, the sensor chip performances were analyzed with static and modal analyses of the diaphragm using ANSYS software. Owing to the structure symmetry, only a quarter of finite element model of the diaphragm along diagonal line was established.

Figure 24 showed the stress difference distribution of the proposed diaphragm under 500 Pa. Thanks to the stiffness mutation created around SCR (Bashir et al. 2000), the mechanical stress was concentrated above the gap between the peninsula and island structure. The changing resistance values of piezoresistors and sensor sensitivity were actually determined by the magnitude of stress difference ( $\sigma = S_y - S_x$ ) between the longitudinal and traversal stresses. It was evident that



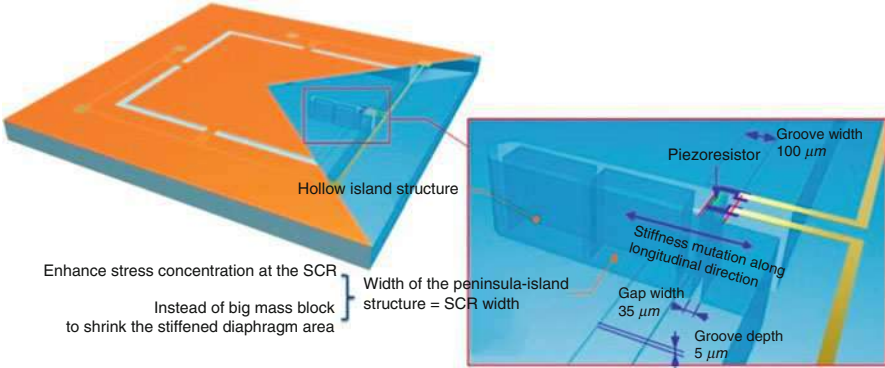


Yu Z, Zhao Y, Li L, et al. Realization of a micro pressure sensor with high sensitivity and overload by introducing beams and Islands[J]. *Microsystem Technologies*, 2015, 21(4):1-9.

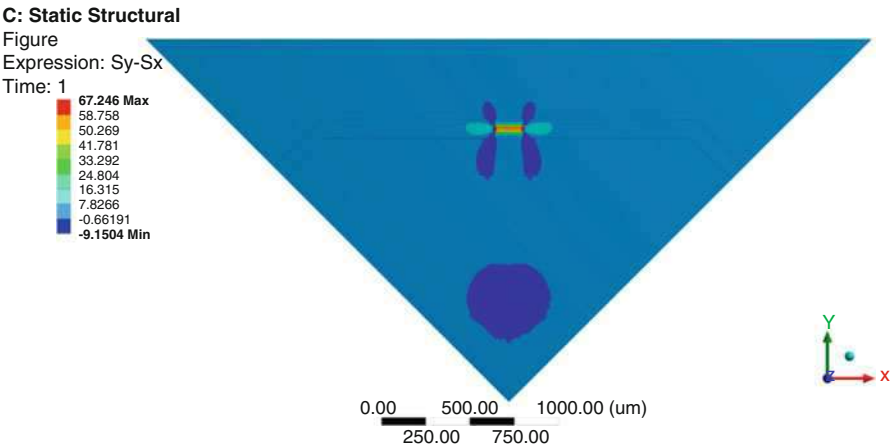
Huang X, Zhang D. A high sensitivity and high linearity pressure sensor based on a peninsula-structured diaphragm for low-pressure ranges[J]. *Sensors & Actuators A Physical*, 2014, 216(3):176-189.

Kimccl P, Kng J, Lester M, et al. A Hollow Shifting Structure for Low Pressure Sensors[J]. *Sensors & Actuators A Physical*, 2009, 111(1):100-105.

**Fig. 22** Stress concentration characteristic of three difference structured pressure sensor chip



**Fig. 23** Stiffness mutation structure

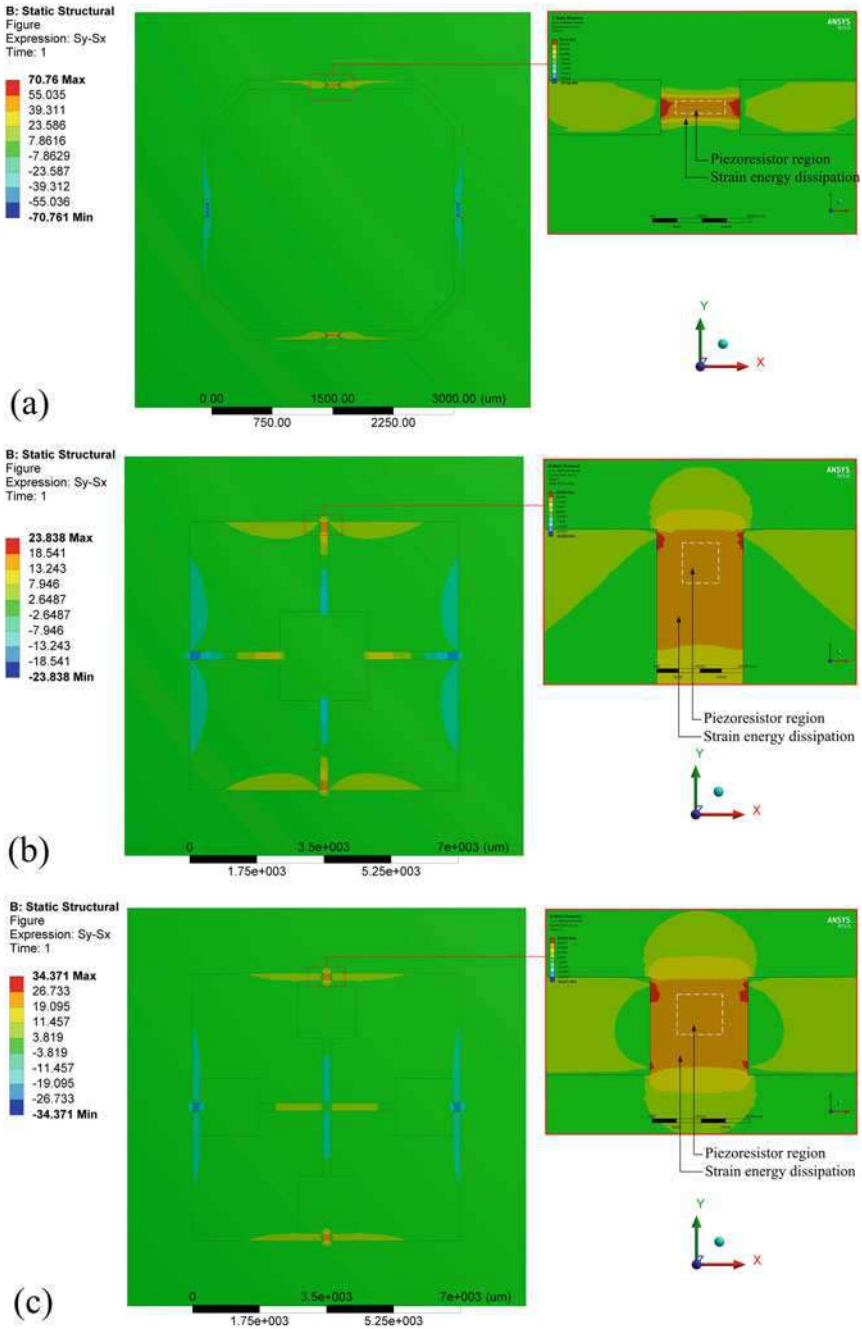


**Fig. 24** Stress difference ( $\sigma = S_y - S_x$ ) distribution for a 1/4 model of the diaphragm

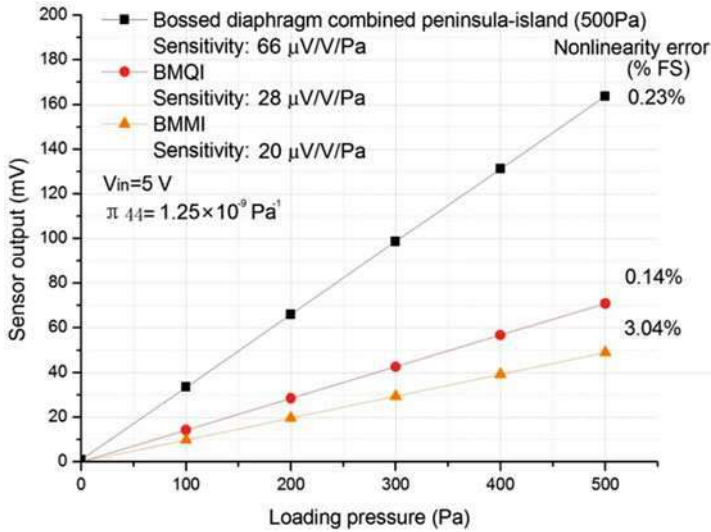
the mechanical stress of diaphragm reached its maximum at the position above the gap. The maximum stress of the diaphragm combined with peninsula-island was increased up to 380% compared with that of the flat diaphragm in the same dimensions.

Figure 25a–c shows that the stress difference distributions of different sensor chips under a pressure of 500 Pa. It was shown that for the sensor chip with the same working range, the smaller the size of SCR was, the less strain energy would be dissipated outside the piezoresistor region, which helped the piezoresistor make the full use of the strain energy to improve the sensitivity.

In order to eliminate the influence from the resistance of piezoresistors, the piezoresistive coefficient  $\pi_{44}$  for all the structures was set as  $1.25 \times 10^{-9} \text{ Pa}^{-1}$ . The sensor chips outputs and nonlinearity errors under 500 Pa were simulated and shown in Fig. 26. Compared with BMMI (Yu et al. 2013) and BMQI



**Fig. 25** The relationship between the SCR and piezoresistor region of comparing diaphragms: (a) Static FEM analysis for the bossed diaphragm combined with peninsula-island structure, (b) static FEM analysis for the BMMI structure, (c) static FEM analysis for the BMQI structure



**Fig. 26** Sensor chip outputs versus loading pressure and non-linearity errors under 500 Pa

(Yu et al. 2015b), the sensitivity of the proposed diaphragm was increased by about 230% and 135.71%, respectively. Although the proposed diaphragm size of the proposed sensor chip was  $3500 \times 3500 \mu\text{m}^2$  which was smaller than the BMMI and BMQI's diaphragm of  $7000 \times 7000 \mu\text{m}^2$  dimension, the proposed sensor chip gave better sensitivity. This is because the design of the peninsula-island structure created a stiffness mutation at the position between the gap, which reduces the strain energy dissipation around the SCR.

And the designing process of the silicon-based piezoresistive pressure sensor could be concluded as below (Fig. 27):

## Lithography Mask Designing

In this section, taking the peninsula-island structured diaphragm as an example, the lithography mask designing of the proposed sensor chip had been discussed.

The lithography mask was fabricated based on a 4-inch n-type 100 oriented SOI wafer. And five lithography masks were used to fabricate the sensor chip, and the whole layout of the lithography masks was shown in Fig. 28.

The front side of the sensor chip was fabricated by the following lithography mask, as shown in Fig. 29.

Figure 29a, b shows two patterns of the piezoresistor. In order to form the Wheatstone full bridge, the resistor of the four piezoresistors should be equal which means the total length of each piezoresistor's pattern should be equal. P-layer mask was used to pattern the piezoresistor region of the sensor chip; P+ layer was used to pattern the heavily doped section to create the ohmic contact between the

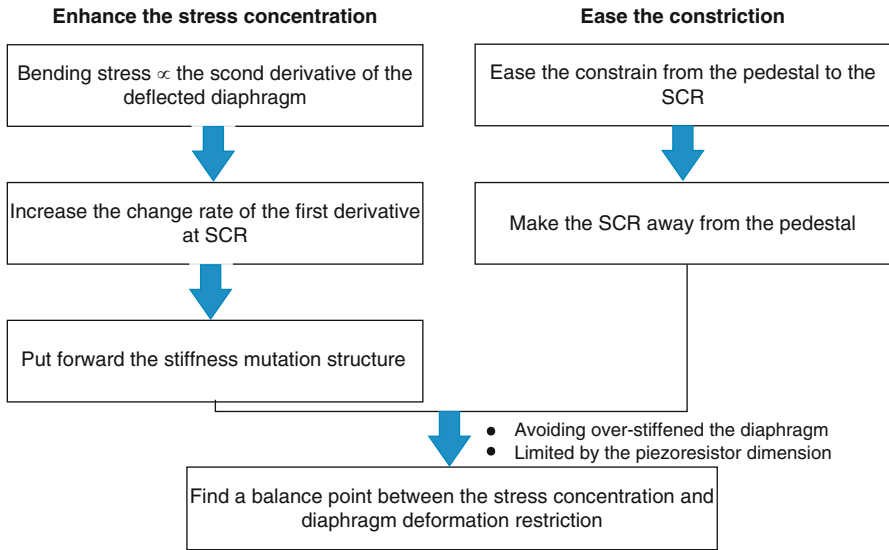


Fig. 27 A brief introduction for the pressure sensor designing process

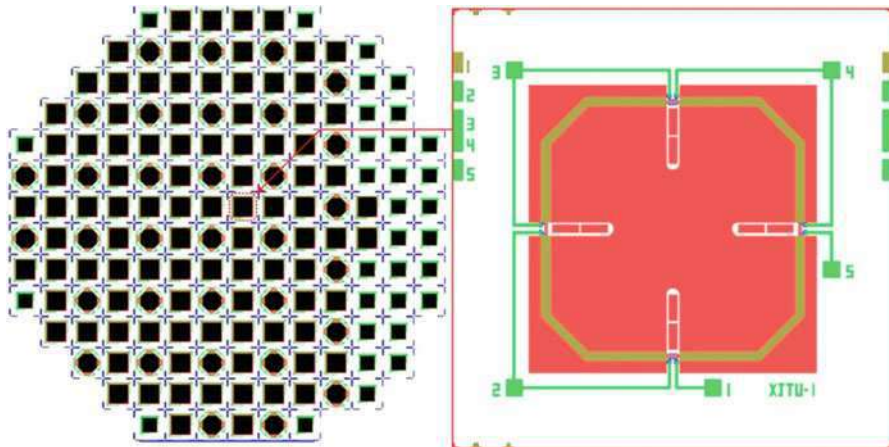
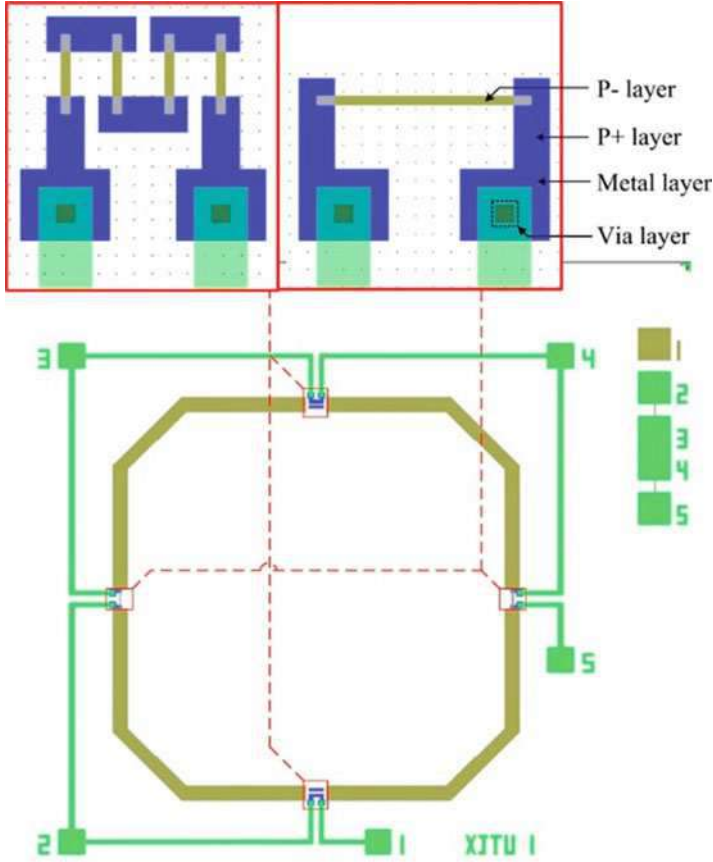


Fig. 28 Whole layout of the lithography mask

metal conductor to the piezoresistor; in order to prevent the short-circuit contact between the silicon wafer and the metal layer, a SiO<sub>2</sub> layer was formed on the surface of the wafer before sputtering the metal layer, which means the piezoresistors formed through P+ and P- layer will be separated from the metal layer by the SiO<sub>2</sub> layer. So pin holes are needed to be etched through the SiO<sub>2</sub> via layer mask and enabled the contact between the metal layer and the piezoresistors. Then, the front etch layer was used to form four grooves on the front surface of the sensor chip by RIE (reactive iron etching).



**Fig. 29** Lithography masks for the front structure of the sensor chip

The back cavity of the sensor chip was fabricated by DREI (deep reactive iron etching). And only one lithography mask was used in this step, as shown in Fig. 30.

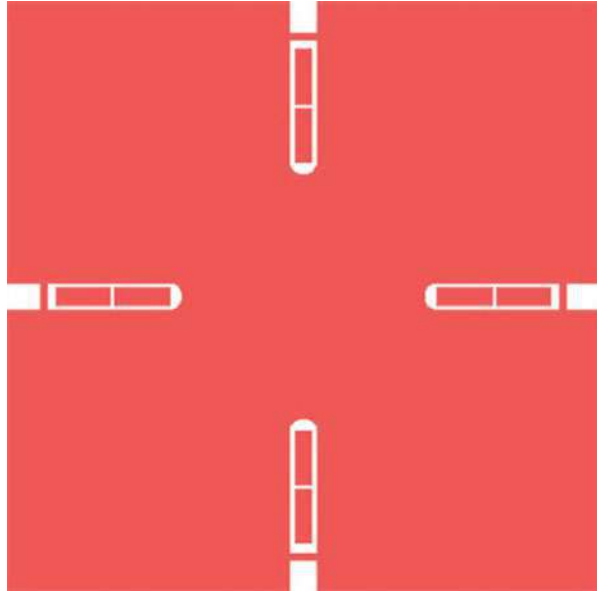
Based on the same fabrication process, sensor chip with different pattern shapes can be fabricated on one layer, as shown in Fig. 31. Another two patterns were also fabricated on the same lithography mask.

### Micromachining Fabrication Process

It is known that the SOI single silicon materials are popular in IC (integrated circuit) manufacture and sensor chip fabrication for high-temperature application. The buried SiO<sub>2</sub> isolation layer in the SOI material guarantees little leakage current when the temperature is up to 350 °C.

The main fabrication processes of the sensor chip were shown in Fig. 32.

**Fig. 30** Lithography masks for the back structure of the sensor chip

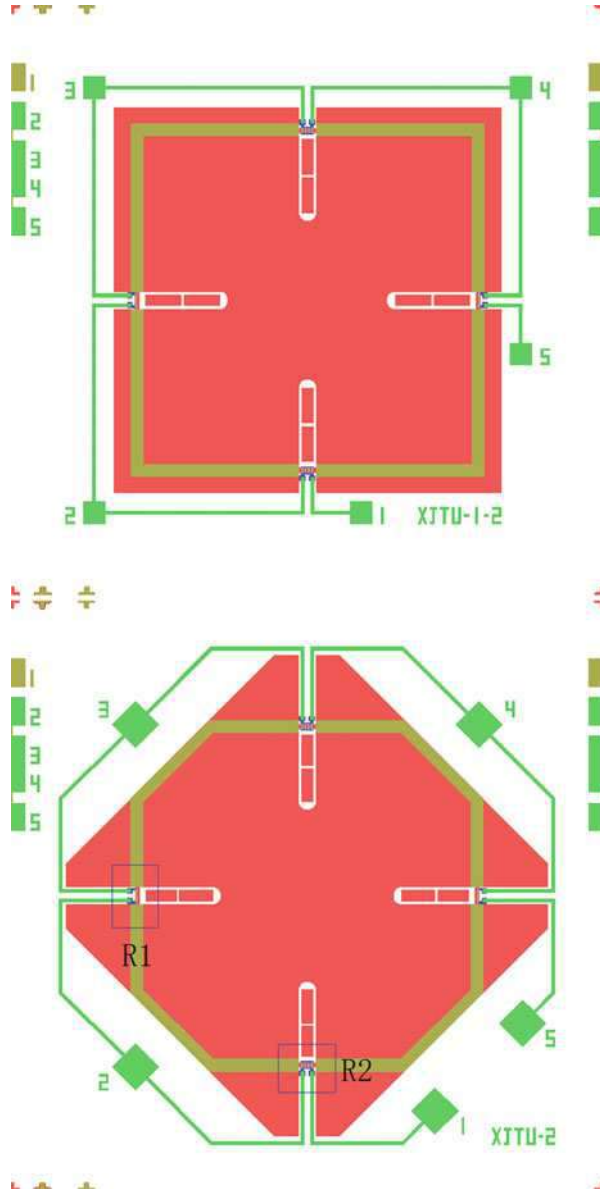


According to Fig. 32a, the starting material is p-type (100) single-crystal silicon which has been fabricated with SIMOX technology. In the SIMOX process, it is important to control the dose of oxygen. If the dose of implanted oxygen is low, the buried silicon oxide  $\text{SiO}_2$  would not be formed, and only the layer of deposited oxygen is generated. However, if the dose of implanted oxygen is high, the oxygen must irretrievably damage the crystallizability of the superficial silicon layer. And the subsequent annealing steps could not restore the crystal quality yet. In the implanting process, the substrate temperature was one of the important parameters to produce SOI material, which may affect the quality of the top silicon layer. The practice shows that the optimal temperature of the target substrate with additional heating resource is kept at  $600\text{ }^\circ\text{C}$ – $650\text{ }^\circ\text{C}$ . In the case of the common SOI structure with 200 nm top Si layer and 400 nm buried  $\text{SiO}_2$  layer, the implanting energy of 200 KeV accompanied with the oxygen dose of at least  $1.4 \times 10^{18}/\text{cm}^2$  is necessary to produce an obviously buried  $\text{SiO}_2$  layer. After the oxygen implantation, the subsequent annealing steps had to transform the defective post-implantation structure into a viable substrate. Annealing at temperatures close to the melting point of silicon ( $1300\text{ }^\circ\text{C}$ ) segregated the top silicon from the oxide precipitates under argon with 0.5%–2% oxygen and then formed an atomic craggedness in the top Si and buried  $\text{SiO}_2$  interface. The oxygen protected the superficial silicon layer from pitting during the high temperatures. In order to increase the thickness of the top silicon layer that satisfies the need of piezoresistivity, the LPCVD (low pressure chemical vapor deposition) technique is used, as shown in Fig. 32b.

In Fig. 32c, the implantation and diffusion of boron was performed from a  $\text{B}_2\text{O}_3$  constant source into the top silicon layer, and then annealing and activation process was carried out in  $\text{N}_2$  for 30 min at  $1100\text{ }^\circ\text{C}$  in order to eliminate some crystal lattice



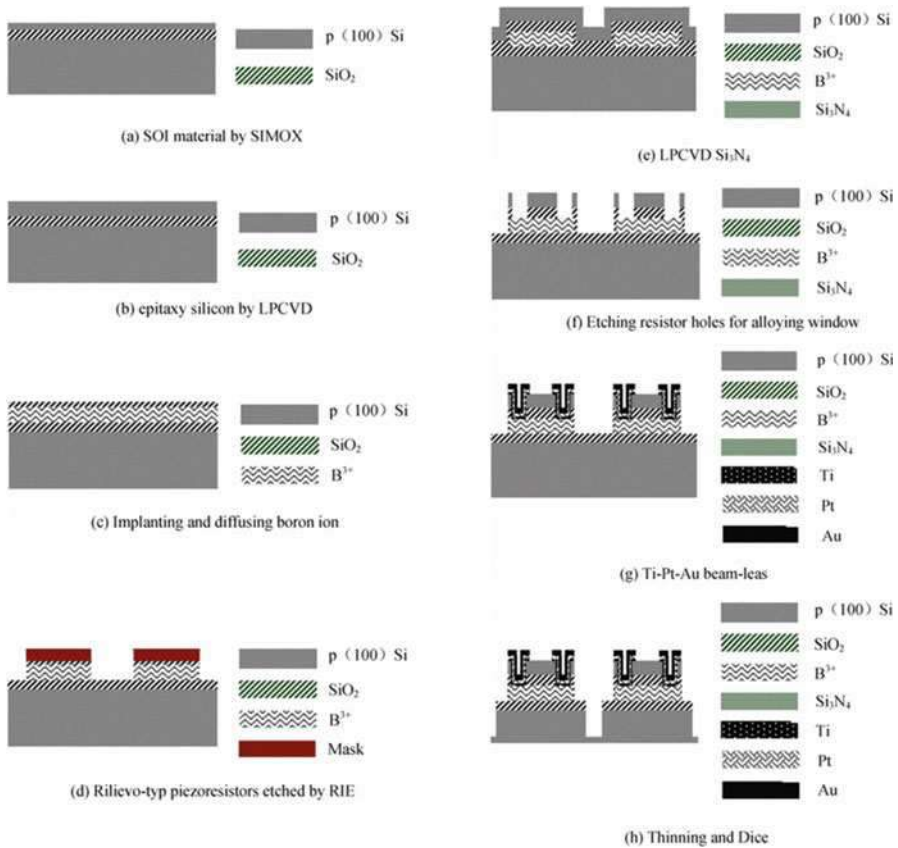
**Fig. 31** Sensor chips with another two patterns



defects and improve the conducting power of the top silicon layer. The top measuring circuit silicon layer was obtained with the boron concentration  $2 \times 10^{20}/\text{cm}^3$ . Another  $\text{SiO}_2$  layer was formed onto the top silicon layer with the oxidation process in the  $\text{O}_2$  for 30 min at  $1100^\circ\text{C}$ .

Next, four piezoresistors with threefold structure were fabricated by RIE (reactive ion etch) technology, as shown in Fig. 32d. The single piezoresistor's width size was





**Fig. 32** Mainly fabrication process (a) SOI material by SIMOX. (b) epitaxy silicon by LPCVD. (c) Implanting and diffusing boron ion. (d) Rilevo-typ piezoresistors etched by RIE. (e) LPCVD Si<sub>3</sub>N<sub>4</sub>. (f) Etching resistor holes for alloying window. (g) Ti-Pt-Au beam-leas. (h) Thinning and Dice

12.5 μm and the total length was 750 μm. The sheet resistance of the four piezoresistors was 20Ω/γ. The piezoresistors were designed and fabricated to rilievo type; in other words, the piezoresistors were embossed from the substrate. And outer region of the piezoresistors, such as redundant implanted silicon, top SiO<sub>2</sub>, were all etched in order to decrease the residual stress induced by the mismatch of thermal expansion coefficient of different membrane materials. So the performances of sensor chip would be improved in the high-temperature circumstance.

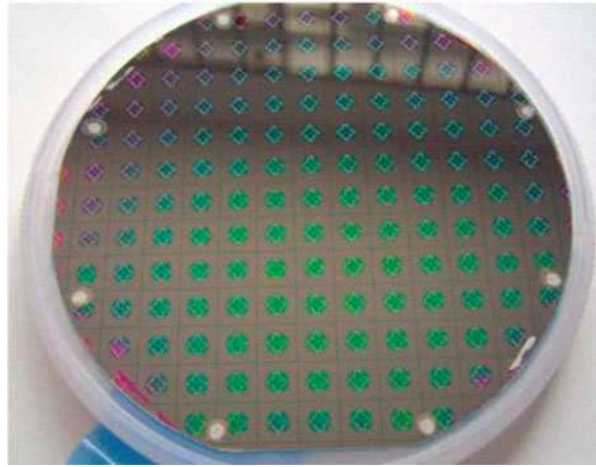
In order to enhance the stability of the top silicon layer and match the internal stresses, a Si<sub>3</sub>N<sub>4</sub> layer was fabricated onto the top SiO<sub>2</sub> layer by LPCVD technology, as shown in Fig. 32e, which also was used to resist the diffusion of contaminant ion and improve the thermal stability of sensor chip.

And then, the resistor holes were etched by the RIE or ICP (inductively coupled plasma) in order to form the inner lead, as shown in Fig. 32f. In order to make the sensor chip work under high temperature of over 200 °C, the Ti-Pt-Au multimetal

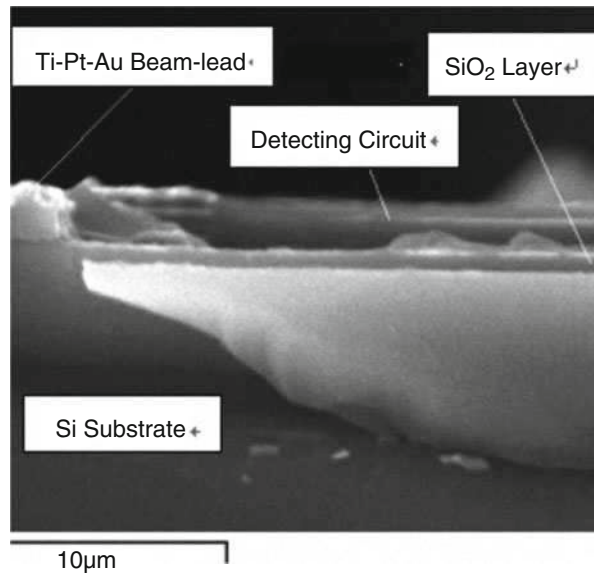
layer was grown by sputtering technology with the thickness of 50 nm, 50 nm, and 500 nm, respectively. And the metal layers were etched to the beam lead to connect the four piezoresistors to form the Wheatstone bridge, as shown in Fig. 32g. Thermal treatment for ohms connection was carried out.

According to the pressure measurement range, the sensor chip was thinning to the theory height by CMP (chemical mechanical polishing). In this paper, the sensor chip with the height of 0.35 mm can be used to measure the pressure with the value of 25 MPa. As shown in Fig. 32h, the wafer was thinned and diced to dies by the scribing technology. Figure 33 showed the fabricated piezoresistive sensor chip, and Fig. 34 shows the SEM photo of the sensor chip in the cross view.

**Fig. 33** Photo of sensor chip



**Fig. 34** SEM photo of the sensor chip in cross section



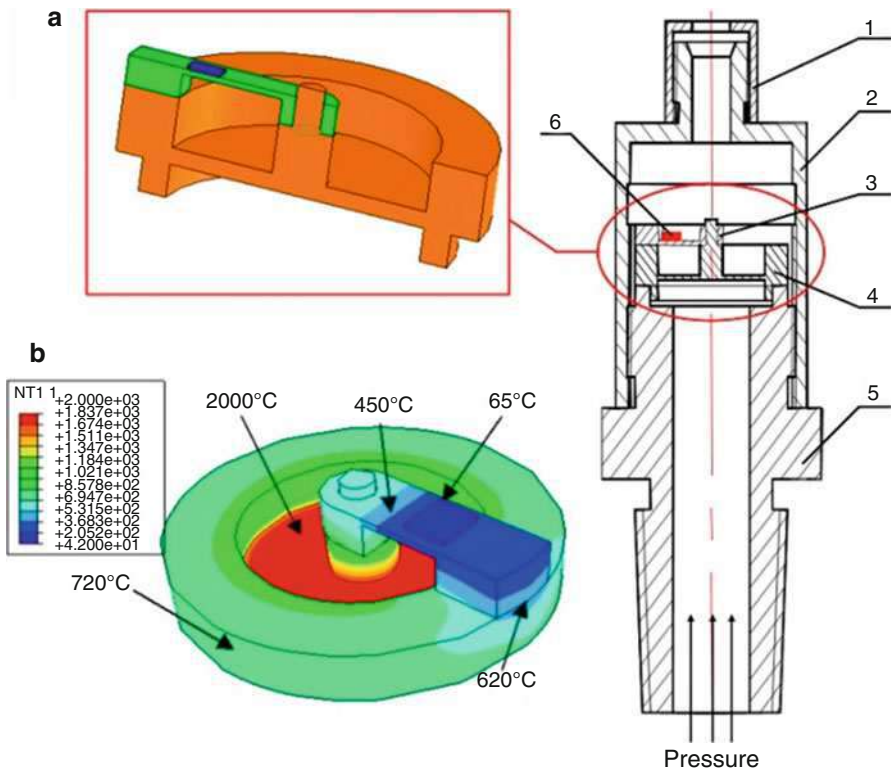
### Packaging Structure for the Pressure Sensor Chip

Packaging technology plays a critical role in MEMS, which makes the sensing chip integrate into an integral system. And the reliability and fabrication cost of a whole MEMS is usually decided by its packaging technology. So the proper designing of the packaging structure and process for a MEMS sensing chip is very important.

Here four packaging structures for high temperature, high pressure, harsh working environment, high working frequency, and high sensitivity application were discussed.

### High-Temperature Packaging

In order to make the sensor chip withstand the 1000 °C thermal shock, a beam-membrane package structure was presented, as shown in Fig. 35a (Jiang 2013). The high-temperature pressure flow first touches the circle membrane, and the deformation of the diaphragm was transmitted to the cantilever beam by transmission rod which makes the stress concentrated at the root of the cantilever beam. This will be



**Fig. 35** Schematic for the beam-membrane packaging structure: (a) the beam-membrane packaging structure, (b) temperature distribution on the packaging structure

detected by the SOI-based high-temperature resisting pressure sensor chip. And the output of the pressure sensor chip was in direct proportion to the pressure value of the high temperature pressure flow. Thanks to the beam-membrane structure, the circle membrane of the structure prevents the direct contact from the ultrahigh temperature thermal shock to the sensor chip. By this means, the high-temperature protection problem of the sensor chip was solved out.

The final size of the beam-membrane structure was determined by the finite element simulation optimization to make the structure meet the requirement of the working range, intrinsic frequency, and resistance ability to high-temperature shock.

And the simulation of the temperature influence to the structure was shown in Fig. 35b. According to the simulation, when the sensor experiences a 2000 °C thermal shock within 500 ms, the temperature values on the sensor chip and membrane were only 65 °C and 350 °C, respectively, which means that the beam-membrane structure played well in the working condition with transient ultrahigh temperature shock.

According to the working mechanism of the sensor, the free end displacement of the cantilever  $w_x$  was equivalent to the center displacement of the membrane  $w_p$ :

$$w_x = \frac{4PL_x^3}{Eb_x h_x^3} = w_p = \frac{3(1 - \mu^2)q_0 R_p^4}{16Eh_p^3} - \frac{3(1 - \mu^2)PR_p^2}{4\pi E h_p^3} \tag{51}$$

where  $L_x$ ,  $b_x$ , and  $h_x$  were the length, width, and thickness of the cantilever and  $R_p$  and  $h_p$  were the radius and thickness of the circular membrane structure, respectively.

According to the structure, the value of  $L_x$  was equivalent to that of the  $R_p$ , and then Eq. (51) can be written as

$$P = \frac{\frac{3(1 - \mu^2)q_0 R_p^2}{16Eh_p^3}}{\frac{3(1 - \mu^2)}{4\pi h_p^3} + \frac{4R_p}{b_x h_x^3}} = \frac{3\pi(1 - \mu^2)q_0 R_p^2 b_x h_x^3}{12(1 - \mu^2)b_x h_x^3 + 64\pi R_p h_p^3} \tag{52}$$

And the maximum displacement at the membrane center was

$$w_p = \frac{3\pi(1 - \mu^2)q_0 R_p^5}{3(1 - \mu^2)Eb_x h_x^3 + 16\pi ER_p h_p^3} \tag{53}$$

And the structure size limitation deduced from the structure can be presented as

$$\begin{cases} |\sigma_p| = \frac{9(1 - \nu^2)q_0 R_p^2 b_x h_x^3 + 96\pi q_0 R_p^3 h_p^3}{24(1 - \nu^2)h_p^2 b_x h_x^3 + 128\pi R_p h_p^5} \leq [\sigma] = 786.25 \text{ MPa} \\ \frac{\omega_p}{h_p} = \frac{3\pi(1 - \nu^2)q_0 R_p^5}{3(1 - \nu^2)Eh_p b_x h_x^3 + 16\pi ER_p h_p^4} \leq 0.2 \\ \frac{h_p}{2R_p} \leq 0.2 \end{cases} \tag{54}$$

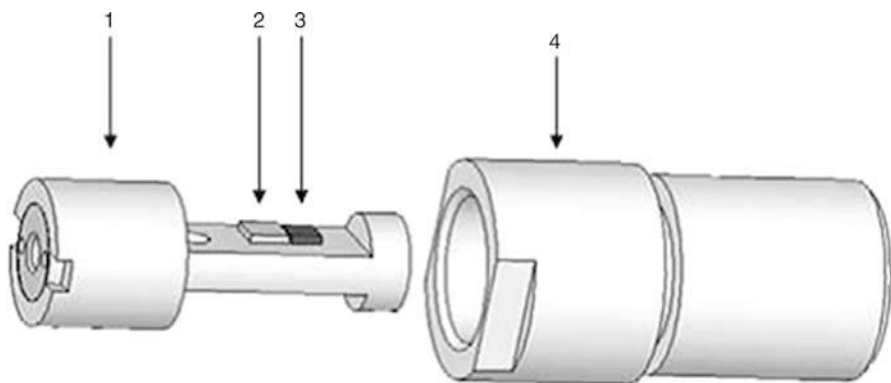
Then the maximum stress value can be presented as

$$|\sigma_x| = -\frac{6PL_x}{b_x h_x^2} = \frac{9\pi(1 - \nu^2)q_0 R_p^3 h_x}{6(1 - \nu^2)b_x h_x^3 + 32\pi R_p h_p^3} \leq 196 \text{ MPa} \quad (55)$$

According to the above analysis, the size of the package structure could be determined.

### Ultrahigh Pressure Packaging

The mechanical structure is designed based on the piezoresistive effect of MEMS sensitive chip, as shown in Fig. 36. It consists of four parts. Part 4 is the outer shell, which protects the chip and cylindrical elastic body inside. Simultaneously, the top of part 4 is a circular plate, which senses pressure at first. The plate is a part of part 4, which is formed by taking materials out of the structure and it is 5 mm thick. Part 1 is the elastic body connected to part 4 with screw thread. Besides part 2 and part 3 are on the surface of it as shown in Fig. 36. Part 2 is a circuit diagram which is used to compensate for zero drift and drift caused by temperature changes, part 3 is the sensor chip connected to the elastic body by welding. Meanwhile, it is connected to part 2 with gold wire to output signal precisely. After the accomplishment of assembly of the sensor, the plate will be in intimate contact with part 1, which is called cylindrical elastic body. When pressure is loaded on the plate, it will get elastic deformation and transmit the pressure to the elastic body. Then the elastic body will engender compressive deformation. As a result, the compressive strain can be measured by the SOI chip with voltage signal output. In the process, the deformation and stress can be calculated with the following formulas.



**Fig. 36** Schematic of mechanical structure. (1) Cylindrical elastic body, whose mechanical model is shown in Fig. 44b. (2) Circuit diagram, to compensate for temperature and zero drift. (3) Sensor chip. (4) Outer shell, in front of which lies circular plate whose mechanical model is as shown in Fig. 44a

The maximum radial stress locates at the edge of the plate, as shown in Eq. (56) (Hsu 2002):

$$(\sigma_{rr})_{\max} = \frac{3W}{4\pi h^2} \quad (56)$$

The maximum circumferential stress locates at the edge of the plate, as shown in Eq. (57) (Timoshenko and Woinowsky-Krieger 1959):

$$(\sigma_{\theta\theta})_{\max} = \frac{3\nu W}{8\pi h^2} \quad (57)$$

The stress in the center of the plate is

$$\sigma_{rr} = \sigma_{\theta\theta} = \frac{3\nu W}{8\pi h^2} \quad (58)$$

While the maximum deflection lies in the center of the plate (Timoshenko 1936):

$$w_{\max} = \frac{3W(m^2 - 1)a^2}{16\pi E m^2 h^3} \quad (59)$$

In all the formulas,  $W = (\pi a^2) \times p$ , and  $m = 1/\nu$ , where  $E$  is Young's modulus of the plate,  $\nu$  is Poisson's ratio,  $h$  is the plate's thickness,  $a$  is the plate's radius, and  $p$  is the equally distributed load. From the upper formulas, it can be discerned that the maximum stress lies at the edge of the plate, which is beyond the range of sensor chip if directly measured by splicing chip on it. Moreover, it is too difficult to splice chip on the edge of the plate. The stress in the center of the plate is also beyond the range of chip under the equally distributed pressure of 2 GPa. Thus, the maximum deflection in the center is utilized because it transforms pressure measurement into strain measurement of the elastic body, which is obviously in the range of chip.

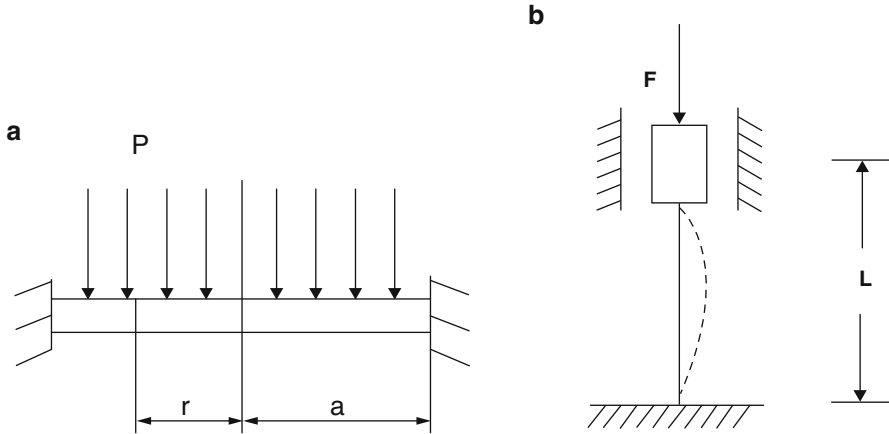
In the cylindrical elastic body, strain is shown in Eq. (60):

$$\xi = \frac{\sigma}{E} = \frac{F}{EA} = \frac{p\pi a^2}{EA} \quad (60)$$

where  $p$  is the equally distributed pressure on the circular plate,  $a$  is the plate's radius, and  $A$  is the elastic body's sectional area. Simultaneously, the elastic body is also a compressive bar as shown in Fig. 37b. It should satisfy the following conditions, shown as Eqs. (61), (62), (63), and (64):

$$F_{\text{cr}} = \frac{\pi^2 EI}{(\mu l)^2} \quad (61)$$

$$n_{\text{st}} = \frac{F_{\text{cr}}}{F} \geq [n_{\text{st}}] \quad (62)$$



**Fig. 37** Working principle of sensor

where  $F_{cr}$  is the critical load of the bar;  $I$  is the minimum moment of inertia of the cross section;  $\mu$  is the length factor, in this situation  $\mu = 0.5$ ;  $l$  is the bar’s length,  $n_{st}$  is the factual safe factor of the bar;  $[n_{st}]$  is the safe factor;  $F$  is the force working on the bar; and  $\sigma_{cr}$  is the bar’s critical stress. To keep the bar from buckling, it should satisfy the following conditions (Cai and Yang 2004; Roark 2003):

$$\sigma_{cr} = \frac{F_{cr}}{A} = \frac{\pi^2 EI}{(\mu l)^2 A} \tag{63}$$

$$\sigma_{cr} \leq \sigma_p \tag{64}$$

where  $\sigma_p$  is the bar’s proportional limit, which is determined by the bar’s material property.

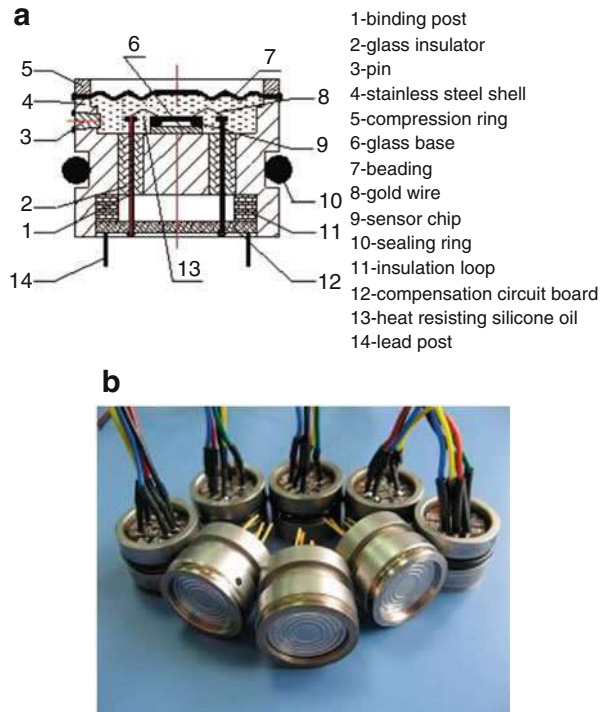
It is determined from the chip’s working principle that  $\xi$  is proportional to the chip’s output voltage. Thus,  $\xi$  can be determined by the chip’s range, and then the series of parameters such as radius  $a$ , sectional area  $A$ , and plate’s thickness  $h$  can be determined by the upper formulas. The process of design is completed when all the parameters are confirmed.

### Harsh Environment Packaging

In order to ensure the working stability and corrosion resistance ability of the sensor chip in an oil and gas field working condition with high temperature, silicone oil filling isolated packaging technology was presented. Deep cavity welding was used to bond the wire of the sensor chip to ensure wire bonding quality and improve the stability of the sensor chip.

Silicon oil filling technology was used to package the sensor chip, and oil-filled cavity was sealed by a corrugated diaphragm. This kind of packaging structure is

**Fig. 38** The schematic of packaging and interconnecting structure of the HPTSS, (a) the packaging and interconnecting structure of the HPTSS, (b) the packaged sensors



based on the incompressibility of the silicone oil and good linearity performance of the corrugated diaphragm to transfer the exterior pressure variation to the sensor chip through the corrugated diaphragm and silicone oil.

The isolated packaging structure avoids the direct contact between the exterior harsh working environment and the sensor chip which enhance the sensor chip's adaptation in harsh working environment (Stankevič and Šimkevičius 2000). And the packing structure of the sensor chip was shown in Fig. 38

## High-Frequency Response Packaging

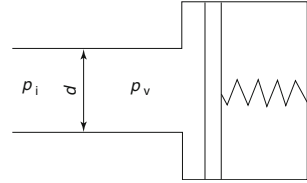
Flush packaging structure was presented to improve the frequency response ability of the sensor chip (Jiang 2013).

The dynamic performance of the sensor chip usually cannot meet the original performance designed through simulation in actual working situation. This is because the "pipeline effect" is induced by the pipeline structure of the packaging shell.

The sensitive part of the sensor cannot always contact with the working medium; there is a pressure transmission pipeline between them. Transmission pipelines and the measured medium constitute a low-pass filter, and the frequency characteristics of low-pass filter influence the accuracy of high-frequency dynamic pressure



**Fig. 39** Typical model for the pressure transmission pipeline



measurement to a great extent. So the package structure with long pipeline will impair the dynamic performance of the sensor a lot.

By analyzing the mechanical properties of this low-pass filter, the high-frequency dynamic pressure measurement transmission pipeline design method is presented. The model for the pressure transmission pipeline was shown in Fig. 39.

This model consisted of a pressure transmission pipeline and a cavity. And the length and diameter of the pipeline were  $l$  and  $d$ , respectively. The volume of the cavity was  $v$ . The measuring pressure at the inlet of the pipeline was  $p_i$ . And the pressure transmitted to the cavity acting on the pressure sensor was  $p_v$ . Assuming that the medium in the pipeline was incompressible and the medium in the cavity is compressible, the inertia and flow velocity of the medium in the cavity can be neglected. So the medium in the pipeline can be equivalent to rigid cylinder with mass  $m$  and the medium in the cavity can be equivalent to a spring without mass. Considering the damping induced by the medium motion, the system is a second-order system.

So the undamped angular frequency of the system was listed in Eq. (56):

$$\omega_n = \frac{d}{2} \sqrt{\frac{\pi E_a}{\rho l V}} = \frac{c}{l} \sqrt{\frac{V_l}{V}} \tag{65}$$

where  $d$  was the diameter of the pressure transmission pipeline,  $l$  was the length of the pipeline,  $V_l$  was the volume of the pipeline,  $\rho$  was the density of the medium,  $E_a$  was the Young’s module of the medium, and  $V$  was the cavity volume toward the sensor chip.

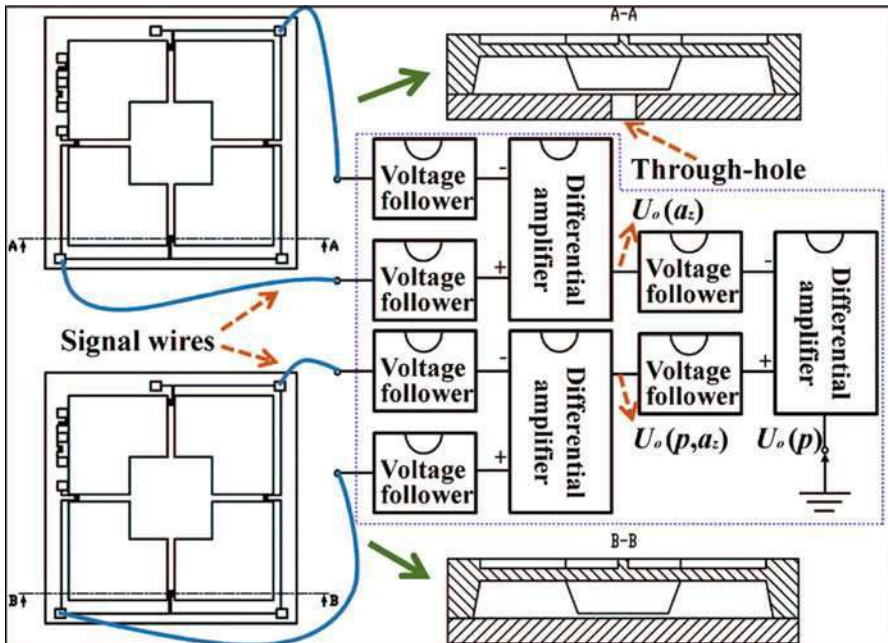
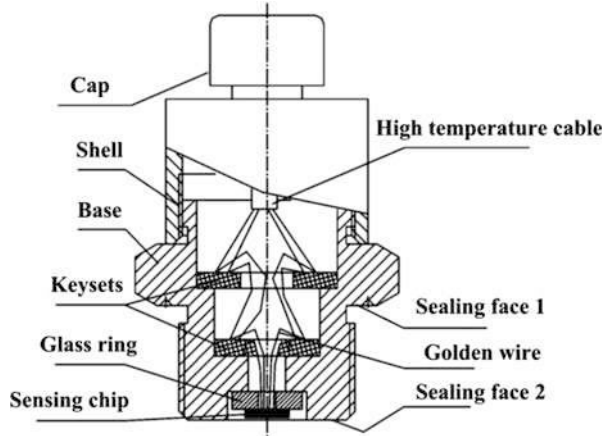
In order to avoid the “pipeline effect” described above, flush packaging method was presented. In this kind of packaging structure, the diaphragm surface of the sensor chip was in a flush position with the inlet surface of the packaging shell, as shown in Fig. 40. This packaging method can avoid the first natural frequency reduction induced by the “pipeline effect.”

### Packaging for Ultrahigh Sensitive Pressure Sensor Chip

Signal decoupling is critical for improving the accuracy of pressure measurements under dynamic circumstance, especially for ultrahigh sensitive pressure sensor. Due to the applied pressure  $p$  and z-axis acceleration,  $a_z$  is linearly related to the output voltage  $U_o$  (Yu et al. 2013).

To decouple the signal, another identical sensor chip is also needed without pressure load on it. Then the sensor chip is only affected by acceleration. Regardless

**Fig. 40** Flush packaging structure



**Fig. 41** The schematic of output signal decoupling

of the fractionally different damping and tiny processing error, the difference caused by the acceleration, namely, the differential readout of the two sensor chips, can accomplish the decoupling. The schematic of differential circuit connection is displayed in Fig. 41, where voltage followers and differential amplifiers are utilized. The dynamic analysis also reveals the feasibility of making the high-sensitive pressure sensor serve as an accelerometer by fabricating a through-hole on the glass.

## Performance Test Analysis of High-Temperature and High-Frequency Pressure Sensor

The temperature characteristic test was carried out on high-temperature pressure sensitive resistance chip, which could analyze the relationship between resistance and temperature. Then, performance test analysis including static and dynamic calibration has been taken in order to satisfy the petrochemical, aerospace, military, and other areas of high-temperature, high-frequency pressure measurement needs. The static calibration was calculated by the relationship of output voltage value under the different pressure input in 250 °C. A relationship curve between the input pressure and output voltage can be established to calculate the static performance index of the sensor, such as nonlinearity, hysteresis, repeatability, accuracy, and thermal zero drift and thermal sensitivity drift. In this paper, the dynamic calibration of the sensor is carried out by using the shock wave tube; the response frequency and the rise time of the sensor are calculated and analyzed according to the measured response curve.

### Characteristic Analysis of High-Temperature Piezoresistive Chip

#### Time Drift and Temperature Drift of Piezoresistors During Test

Four different piezoresistors in a tube core have been tested to verify the temperature characteristic of high-temperature piezoresistive chip, and the value of resistance is shown in Table 2. The distribution curve is shown in Fig. 42.

Temperature coefficient of resistance can be defined as

$$\alpha = \frac{R_T - R_0}{R_0 \Delta T} \times 100\% \quad (66)$$

$R_T$  – Resistance at high temperature

$R_0$  – Resistance at room temperature

$\Delta T$  – Temperature difference

The temperature coefficient of resistance generally cannot be more than  $2 \times 10^{-3}/^\circ\text{C}$ . According to the data in Table 2 and Eq. (66), the temperature coefficients of the four piezoresistors in different temperature ranges can be calculated, as shown in Table 3. According to Table 3, the temperature coefficients of the four piezoresistors are basically the same.

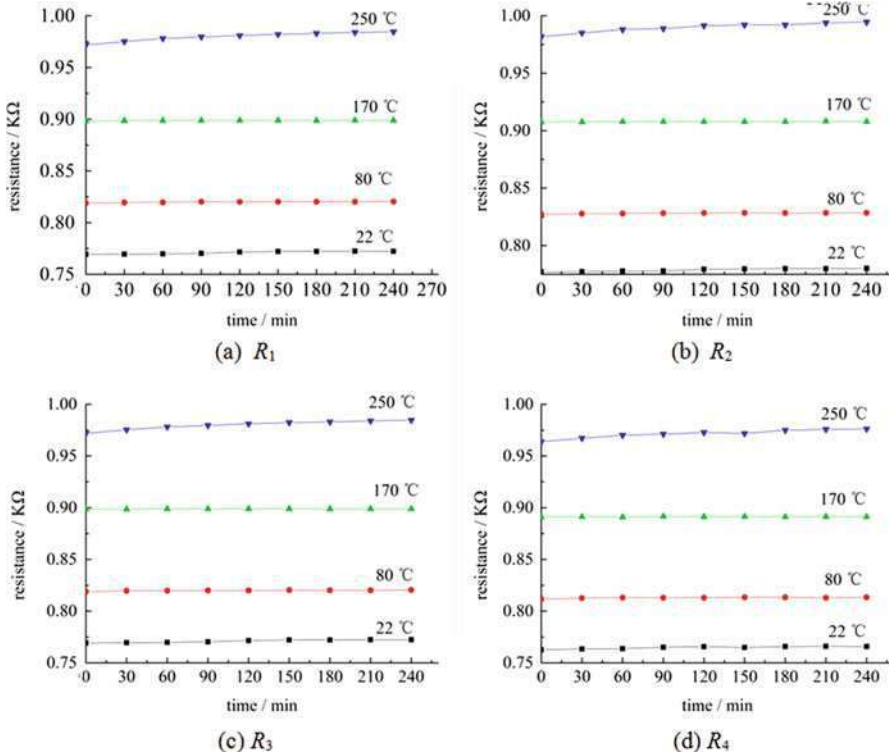
#### The Zero Output of the Piezoresistive Chip

According to the data in Table 2, suppose using constant current source  $I_B = 5 \text{ mA}$  power supply, the zero point output of the uncompensated piezoresistive chip at four temperature points is shown in Table 4.

According to the data in Table 4, the zero point output of the piezoresistive chip without compensation is changed little at 22 °C~80 °C, but when the temperature is above 80 °C, the temperature change is relatively large and compensation must be considered.

**Table 2** Resistance in different temperatures

Resistances	Temperature/ $^{\circ}\text{C}$	Time drift resistance at different temperatures/ $\text{K}\Omega$											Average value/ $\text{K}\Omega$
		0	30 min	60 min	90 min	120 min	150 min	180 min	210 min	240 min			
$R_1$	22	0.7693	0.7695	0.7698	0.7702	0.7716	0.7721	0.7721	0.7724	0.7723	0.7712	0.7712	
	80	0.8188	0.8195	0.8196	0.8201	0.82	0.8202	0.8201	0.8202	0.8203	0.8199	0.8199	
	170	0.8983	0.8985	0.8987	0.8988	0.8989	0.8989	0.8987	0.8989	0.8988	0.8987	0.8987	
	250	0.9718	0.9751	0.978	0.9796	0.9808	0.982	0.983	0.9839	0.9845	0.9804	0.9804	
$R_2$	22	0.7771	0.7773	0.7778	0.778	0.7794	0.7797	0.7799	0.78	0.7802	0.7790	0.7790	
	80	0.8272	0.8278	0.8279	0.8282	0.8283	0.8285	0.8283	0.8284	0.8285	0.8281	0.8281	
	170	0.9075	0.9076	0.9078	0.9079	0.9078	0.9078	0.9078	0.9082	0.9079	0.9078	0.9078	
	250	0.9817	0.985	0.988	0.9888	0.9914	0.992	0.992	0.9938	0.9945	0.9902	0.9902	
$R_3$	22	0.7692	0.7695	0.7698	0.7704	0.7716	0.7721	0.7722	0.7724	0.7724	0.7712	0.7712	
	80	0.819	0.8196	0.8198	0.82	0.82	0.8203	0.8202	0.8202	0.8204	0.8199	0.8199	
	170	0.8986	0.8986	0.8989	0.899	0.899	0.899	0.8988	0.8989	0.8989	0.8989	0.8989	
	250	0.972	0.9754	0.9782	0.9796	0.9811	0.9824	0.983	0.9839	0.9847	0.9806	0.9806	
$R_4$	22	0.7626	0.7635	0.7638	0.7651	0.7657	0.765	0.7658	0.766	0.7658	0.7646	0.7646	
	80	0.8115	0.8126	0.8131	0.8129	0.813	0.8134	0.8134	0.813	0.8134	0.8129	0.8129	
	170	0.891	0.8913	0.891	0.8917	0.8913	0.8915	0.8912	0.8914	0.8912	0.8912	0.8912	
	250	0.9639	0.9672	0.9701	0.9712	0.9728	0.9718	0.9749	0.9758	0.9762	0.9721	0.9721	



**Fig. 42** The distribution curve of the values of resistance at different temperatures. (a)  $R_1$ . (b)  $R_2$ . (c)  $R_3$ . (d)  $R_4$

**Table 3** The temperature coefficients of piezoresistors

	Temperature coefficient of resistance at different temperature ranges/ $10^{-3} \text{ }^\circ\text{C}^{-1}$		
	22 °C~80 °C	22 °C~170 °C	22 °C~250 °C
$R_1$	1.09	1.12	1.18
$R_2$	1.09	1.12	1.19
$R_3$	1.09	1.12	1.19
$R_4$	1.09	1.12	1.19

**Table 4** Zero point output

Temperature/ $^\circ\text{C}$	22	80	170	250
The zero point output of the uncompensated piezoresistive chip/mV	1.416	1.412	1.654	1.521

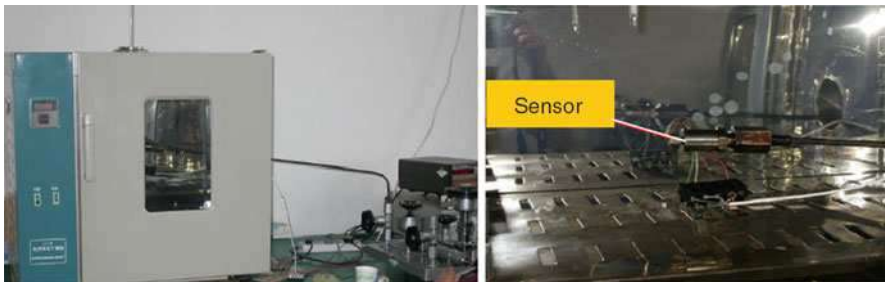
## Static Performance of the Sensor

The static performance of the sensor is tested by equipment: piston pressure gauge, high-temperature drying oven, constant current source, digital multimeter Fluke 45, stainless steel pressure tube, etc. The piston pressure gauge uses the YS-600 type device made by Xi'an Instrument Factory. The working principle is that the pressure generated in the hydraulic container is equal to the pressure generated on piston by the weight of the piston itself and the specific mass. The device has an accuracy level of 0.02 and the measurement range of 0~60 MPa and must be operated at temperature of  $20\text{ }^{\circ}\text{C} \pm 5\text{ }^{\circ}\text{C}$  to ensure their test accuracy. In order to test the static performance of the sensor in high temperature, we have adopted the pressure tube to offer the pressure. One end of the pressure tube is installed in the piston pressure gauge calibration port, and the other end is installed into the high-temperature drying box with sensor. The pressure generated by the piston pressure gauge is applied to the pressure sensor placed in a high-temperature environment by means of a pressure pipe as shown in Fig. 43:

The high-temperature and high-frequency pressure sensor with a range of 25 MPa was presented as an example to calibration. It was calibrated with constant current source  $I_B = 5\text{ mA}$  power supply, at high temperature  $250\text{ }^{\circ}\text{C}$  for testing. Calibration data is recorded in Table 5. The nonlinear, repeatability, hysteresis, accuracy, and other indicators can be calculated by the sensor performance parameters by the calculation method according to GB/T15478-1995.

### Nonlinear

The non-linearity of the pressure sensor refers to the proximity between the relation curve of the output signal  $V$  and the input pressure  $P$  and ideal line. Nonlinearity error is the deviation between the curve of the output-input, and the ideal line is divided by the full-scale output voltage  $V_F$ . The nonlinearity error is unsure for a sensor because the calculated nonlinearity error is different if the calculated linearity is different for the same calibration curve. We use the least squares method to fit the measured data, and the fitting straight line equation is

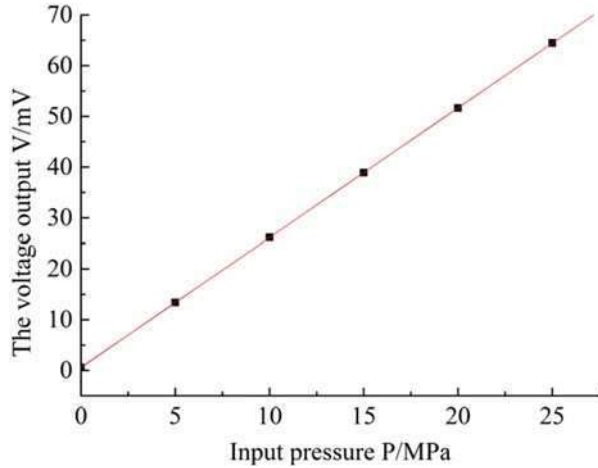


**Fig. 43** High-temperature and high frequency pressure sensor static calibration test stand

**Table 5** Calibration data of sensor

Test temperature/°C	Pressure/MPa	Output signal/mV							
		First time		Second time		Third time			
		Forward stroke	Reverse stroke	Forward stroke	Reverse stroke	Forward stroke	Reverse stroke		
250	0	0.58	0.59	0.59	0.58	0.59	0.58	0.5867	0.58333
	5	13.45	13.43	13.44	13.44	13.43	13.45	13.44	13.44
	10	26.17	26.15	26.16	26.15	26.15	26.15	26.16	26.15
	15	38.92	38.92	38.93	38.93	38.92	38.92	38.9233	38.9233
	20	51.67	51.65	51.67	51.65	51.66	51.67	51.667	51.6567
	25	64.45	64.46	64.45	64.47	64.46	64.46	64.4533	64.4633

**Fig. 44** The curve of input-output by least squares fit



$$V_{LS} = a + bP \tag{67}$$

$$a = \frac{\left(\sum_{i=1}^n V_i\right)\left(\sum_{i=1}^n P_i^2\right) - \left(\sum_{i=1}^n P_i\right)\left(\sum_{i=1}^n V_i P_i\right)}{n \sum_{i=1}^n P_i^2 - \left(\sum_{i=1}^n P_i\right)^2}, \quad b = \frac{n \sum_{i=1}^n V_i P_i - \left(\sum_{i=1}^n V_i\right)\left(\sum_{i=1}^n P_i\right)}{n \sum_{i=1}^n P_i^2 - \left(\sum_{i=1}^n P_i\right)^2}$$

According to the average of the measured data in Table 5 and Eq. (67), we can calculate that  $a = 0.623$ ,  $b = 2.553$ . The curves are shown in Fig. 44:

The nonlinearity of the sensor can be represented as:

$$\delta_L = \frac{|\bar{V}_i - (a + bP_i)|_{\max}}{V_{FS}} \times 100\%FS \tag{68}$$

$\bar{V}_i$  – Average value of the measurements

$V_{FS}$  – The full scale output of the transducer

The nonlinear error is obtained as  $\delta_L = 0.081\% FS$ .

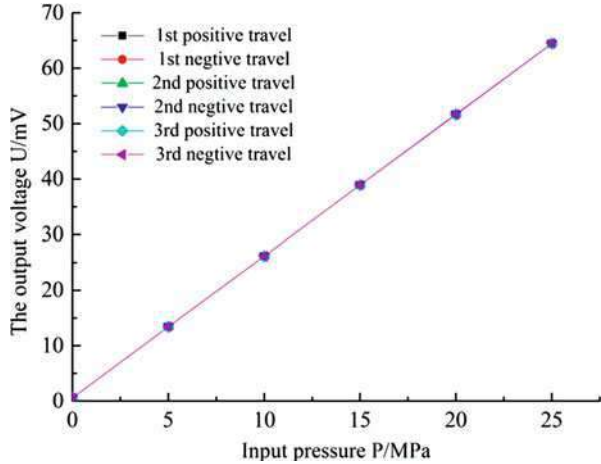
### Repeatability

Repeatability refers to the degree of inconsistency for the output characteristic curve when the sensor test is on full range many times and the input is tested in the same direction. The stand or fall of repeatability is related to many factors; the main reason is whether there are inevitable defects in the mechanical parts of the sensor, such as clearance, the loose of the fasteners, the internal friction of the material, etc. According to the measurements in Table 5, we can get the repeatability curve as shown in Fig. 45.

The indicator of repeatability generally adopts the percentage of the maximum non repetition error of output and full-scale output as



**Fig. 45** Repeatability curve



$$\delta_R = \frac{\Delta V_{Rmax}}{V_{FS}} \times 100\%FS \tag{69}$$

Repeatability error belongs to the random error; it is not very reasonable if calculated by the above method, and the dispersion degree of calibration data is associated with the precision of random error, so the indicator of repeatability can be calculated according to the standard deviation expressed as

$$\delta_R = \frac{\lambda\sigma}{V_{FS}} \times 100\%FS \tag{70}$$

$\lambda$  – Coverage factor, usually take as 2 ~ 3

$\sigma$  – Standard deviation, it can be calculated according to Bessel formula

And  $\sigma_{1i}$  is the standard deviation of forward stroke subsample, expressed as

$$\sigma_{1i} = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (V_{1ij} - \bar{V}_{1i})^2} \tag{71}$$

$\sigma_{2i}$  is the standard deviation of reverse stroke subsample, expressed as

$$\sigma_{2i} = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (V_{2ij} - \bar{V}_{2i})^2} \tag{72}$$

$\sigma$  is the sensor's subsample deviation in the entire measuring range, expressed as

$$\sigma = \sqrt{\frac{1}{2m} \left( \sum_{i=1}^m \sigma_{1i}^2 + \sum_{i=1}^m \sigma_{2i}^2 \right)} \tag{73}$$

$V_{1ij}$  – The measured value in forward stroke

$V_{2ij}$  – The measured value in reverse stroke

$\bar{V}_{1i}$  – The average value of the multiple measurements in forward stroke

$\bar{V}_{2i}$  – The average value of the multiple measurements in reverse stroke

$n$  – The times of the measurement

$m$  – The number of calibration points

According to Eq. (70), the sensor's repeatability error is obtained as  $\delta_R = 0.03\%FS$ .

## Hysteresis

Hysteresis characteristics show the degree of misalignment of the output-input characteristic curve in the forward stroke (increase input) and reverse stroke (reduce input) of the sensor. That is to say, for the same input signal, the sensor's forward stroke output and reverse stroke output are not equal. Hysteresis phenomenon is mainly related to the encapsulation process.

Hysteresis size was generally determined by experimental method. It is expressed as a percentage of the maximum differential output value between forward stroke and reverse stroke and the full-scale output value as

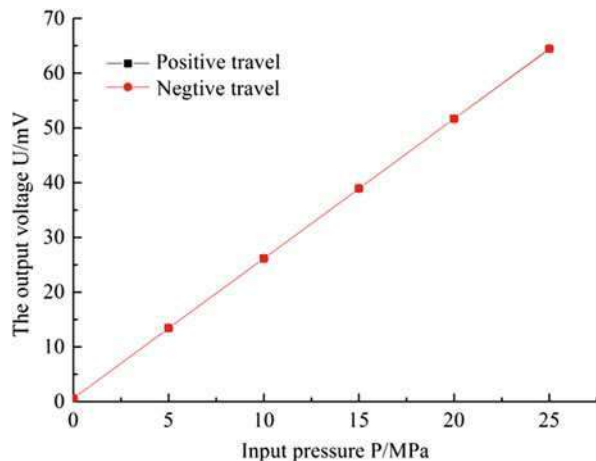
$$\delta_H = \frac{\Delta V_{Hmax}}{V_{FS}} \times 100\%FS \quad (74)$$

According to Eq. (74),  $\delta_H = 0.03\%FS$  can be calculated. Hysteresis curve is shown in Fig. 46.

## Precision

The sensor's precision is defined as the degree of agreement between the measured value and the true value. It depends on the size of the system error band and random error band, and the indicators relevant to temperature generally are ignored. When using least squares method to linear fit, the sensor's system error

**Fig. 46** Hysteresis curve



band  $\gamma_1$  is defined as the average value of forward stroke output and reverse stroke output, respectively, subtracted with the value of line equation which is calculated by least squares fitting, and the biggest absolute value of the difference is the size of system error band  $\gamma_1$ :

$$\gamma_1 = \max(|\bar{V}_{1i} - V_{LS}|, |\bar{V}_{2i} - V_{LS}|) \quad (75)$$

$\bar{V}_{1i}$  – The average value of forward stroke output  
 $\bar{V}_{2i}$  – The average value of reverse stroke output  
 $V_{LS}$  – The line equation get by least squares fitting  
 The sensor's random error band is defined as

$$\gamma_2 = 3\sigma \quad (76)$$

$\sigma$  – Subsample deviation, as shown in the Eq. (73)  
 The sensor's precision is defined as

$$\delta = \pm \frac{\gamma_1 + \gamma_2}{V_{FS}} \times 100\%FS \quad (77)$$

According to Table 5 and Eqs. (67), (73), (76), and (77),  $\delta$  can be calculated as

$$\delta = \pm \frac{0.0566 + 3 \times 0.0054}{63.83} \times 100\%FS = \pm 0.114\%FS \quad (78)$$

### Zero Output, Zero Drift, and Thermal Zero Drift

Zero output is the sensor's output value in the condition of a certain reference temperature and a certain electric excitation (constant current source and constant voltage source power supply) without additional pressure. Sometimes, the absolute value of zero output voltage may be greater than the absolute value of the output voltage range limit. At this time we can take some resistors in series or parallel on a piezoresistor's bridge to make sure that zero output is zero. Zero drift refers to the change of the sensor's zero output in the certain of temperature and time, under the condition of standard power supply. The calculating equation for zero drift  $D$  is expressed as

$$D = \frac{\Delta V_{0\max}}{V_{FS}} \times 100\% \quad (79)$$

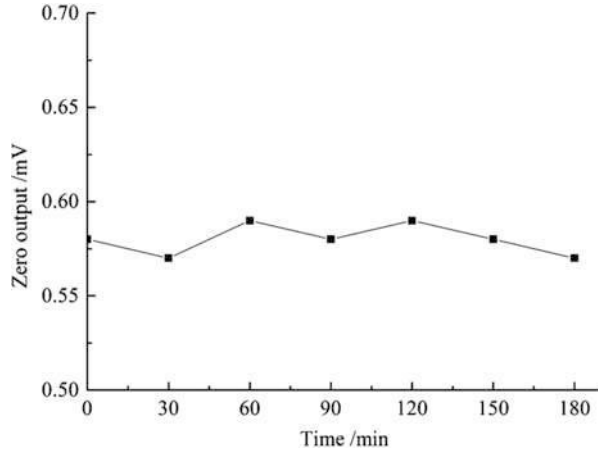
where  $\Delta V_{0\max}$  is the biggest difference between sensor's zero output.

The sensor's zero drift characteristics are measured at high temperature of 250 °C, using constant current source  $I_B = 5$  mA as power supply; the measurement time interval is 30 min, which lasts for 3 h. zero output value derived from the test was shown in Table 6, And its distribution curve is shown in Fig. 47.

According to Eq. (78) and the data in Tables 5 and 6, the sensor's zero drift can be calculated as 0.031%FS/3 h in 3 h.

**Table 6** Zero drift test values

Time/min	0	30	60	90	120	150	180
Zero output/ mV	0.58	0.57	0.59	0.58	0.59	0.58	0.57

**Fig. 47** Schematic of the zero drift curve

The piezoresistive pressure sensor has zero output because of the inconsistent resistance values caused by the designed sizes differences, lithography error, and the dosage concentration differences on each force sensing resistor.; on the other hand, it is also because of the attached compression resistance effect due to the thermal stress caused by the package process. The zero output has close association with temperature, because of inconsistent temperature coefficient in resistance, for which the different dosage concentration of the force sensing sensor is responsible. When the temperature varies within the normal temperature range of the sensor, the zero output also shifts, which is known as the zero drift. Thermal zero drift is an important factor to measure the performance of the pressure sensor, as estimated by the following formula:

$$k_0 = \frac{V_0(T_2) - V_0(T_1)}{V_{FS}(T_1)(T_2 - T_1)} \times 100\%FS/^\circ C \quad (80)$$

where  $V_{FS}(T_1)$  is the output voltage in full range at room temperature,  $V_0(T_2)$  is the zero output voltage at  $T_2$ , and  $V_0(T_1)$  is the zero output voltage at  $T_1$ .

The experiments of the thermal zero drift have been carried out for the 25 MPa high-temperature and high-frequency pressure sensor. The constant flow source  $I_B$  was set to 5 mA; the zero output and the full-range output of the compensated sensor were recorded when  $T_1$  is 22 °C. The voltage is recorded at  $T_1$ ,  $V_0(T_1)$  is recorded at -1.1516 mV, and the full-range output  $V_{FS}(T_1)$  is recorded at 61.682 mV. According to Table 5, the sensor has a zero output  $V_0(T_2)$  at 0.585 mV and the full-range output  $V_{FS}(T_2)$  at 63.8733 mV, both of which are measured at  $T_2$  of 250 °C. Consequently it can be calculated that  $k_0$  is 0.012%FS/°C.

### The Full-Range Output and Thermal Sensitivity Excursion

The full-range output of the pressure sensor is the corresponding difference of the output signal for the specified top and down limit of pressure, at a certain temperature and certain current or voltage, as estimated by the following formula:

$$V_{FS}(T) = V_M(T) - V_0(T) \quad (81)$$

The full-range output has reduced the zero output  $V_0$ , and it is close to the dosage concentration and position of the force sensitive resistance, which has a large output when the resistance is located at the area of the largest stress.

The sensitivity is generally used to describe the sensors' quality, which is defined as

$$S = \frac{1}{I_B} \frac{V_{FS}}{P_{FS}} \quad (82)$$

where  $I_B$  represents the current source and the  $P_{FS}$  is the full-range pressure of the sensor.

Both the pressure sensitivity  $S$  and the full-range output  $V_{FS}$  variate with the temperature. This is because the pressure sensitivity is proportional to the piezoresistive coefficient changing with the temperature. What's more, the difference of the thermal expansion coefficient between the force sensitive resistance and materials such as the surface layer of  $\text{SiO}_2$  and passivation layer of  $\text{Si}_3\text{N}_4$  accounts for the generated thermal stress during the package process and the attached piezoresistive effects. Therefore, the thermal sensitivity drift  $K_S$  is one of the key points applied to judge the stability of the pressure sensors, as calculated by the following formula:

$$k_S = \frac{V_{FS}(T_2) - V_{FS}(T_1)}{V_{FS}(T_1)(T_2 - T_1)} \times 100\%FS/^\circ\text{C} \quad (83)$$

$V_{FS}(T_2)$  – The full-range output voltage at  $T_2$

$V_{FS}(T_1)$  – The full-range output voltage at  $T_1$

When the  $V_{FS}(T_1)$  and  $V_{FS}(T_2)$  recorded in previous research were 61.682 mV and 63.8733 mV, respectively. The  $K_S$  can be calculated by Eq. (82) as 0.016%FS/ $^\circ\text{C}$ .

### Dynamic Performance Indicators of the Sensors

The output-input features as a straight line with one-to-one corresponding relationship when measuring the static signal of the sensors. At the same time, the measured signal will not vary with period, that is to say, the measurements and the records are not limited by time. In practice, there are plenty of dynamic signals needed to be measured, where it is required not only that the signal amplitude should be measured accurately but also the waveform in the process of transformation should be measured and recorded in the meantime. Therefore, it is required that the sensors should

achieve immediate and accurate measurements of the signal amplitude, at the same time, that should catch the waveform corresponding with time unconstrained by distortion.

The sensors' dynamic characteristics refer to how the responses (output) react to the incentives (input). The dynamic characteristic of the sensor, whose output changes over time (the curve of change), will be able to reproduce the time-varying law of the input (change curve), which means that the same time function is used. This is the new requirements of the sensor in dynamic measurement.

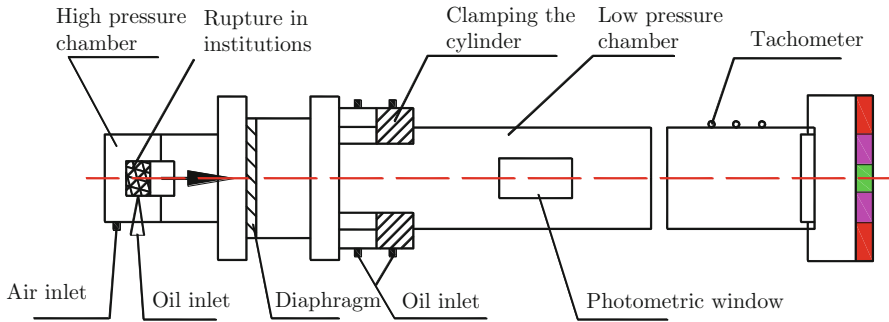
In order to meet the needs of a wider range of applications, a high-temperature and high-frequency pressure sensor with characteristics of high-temperature, high-frequency response is developed in this paper. In order to verify its high-frequency response characteristics, it is necessary to determine its dynamic performance by dynamic calibration test. At present, the main dynamic calibration methods can be concluded as follow:

- (1) Sinusoidal excitation method, which was mainly used for dynamic calibration of sensors below 50 KHz and not suitable for high-frequency dynamic calibration;
- (2) Half sine excitation method. The excitation in this method was usually realized by a dropping hammer breaker. It can be used to calibrate dynamic performance of the sensor which natural frequency is 2~100 KHz. However the main drawback of this method is its poor repeatability and cannot be used for high-frequency dynamic calibration;
- (3) Step stimulation method, which can be realized by shock tube.

### **Stimulation of Step Pressure: Shock Wave Tube**

The primary problem is to get a dynamic excitation signal generator with the bandwidth which can fully cover the bandwidth of verified sensor and the modals of test system, so that the main modals of verified test system can be stimulated (Stankevič and Šimkevičius 2000). Among step pressure generators, shock tubes can generate a steep-front step pressure whose rising edge and platform keeping time are less than 0.1  $\mu$ s and more than 5 ms, respectively. Therefore, this signal is ideal for dynamic calibration of pressure sensor. It is the most commonly used method to calibrate the dynamic characteristics of pressure sensors with shock tubes. They have three characteristics: (1) wide pressure range and easy to change the pressure value, (2) wide frequency range (2 kHz~2.5 MHz), and (3) data processing that is easy to analyze and study. Besides, shock tubes have simple structures and are reliable to use.

Shock wave is a sudden change in pressure somewhere in the gas; the pressure wave will transmit at the speed faster than the velocity of sound. The transmission speed varies with the strength of the pressure: the greater the pressure change is, the higher the wave speed will be. The characteristic of this wave is that the pressure, temperature, and density of the gas change drastically when the wave front reaches a certain location. When the wave front has not yet reached a certain location, the gas is completely free from its disturbance. The pressure, temperature, and density of the



A schematic of shock wave tube

**Fig. 48** The structure of shock wave tube

gas behind the wave front are higher than the characteristics of the gas in the front of the wave front, and the particles of the gas flow in the direction of the wave front, but the velocity is lower than the velocity of the wave front.

Shock wave tube is used to produce a plane shock wave. The structure of a shock wave tube is shown in Fig. 48; it is composed of two straight tubes, the shorter section is the high-pressure chamber and the longer section is the low-pressure chamber. The calibrated sensor will be installed in the end face of the low-pressure chamber. The high-pressure chamber and low-pressure chamber are separated by a diaphragm before the test; the size of the shock pressure is determined by the thickness of diaphragm. High-pressure air transmits through the pressure reducer to the high-pressure chamber and pressure-reducing regulator in order for export pressure to gradually increase. When the pressure of the high-pressure chamber rises to a certain value, the diaphragm ruptures, and at the same time a shock wave is formed in the low-pressure chamber; it spreads at a very high speed (usually 500~600 m/s). A space variant of pressure is formed before and after the shock wave. The sensor generates a damped oscillation according to its natural frequency by the excitation of shock wave; the vibration waveform is recorded, which can be used to analyze the dynamic frequency response characteristics of the sensor.

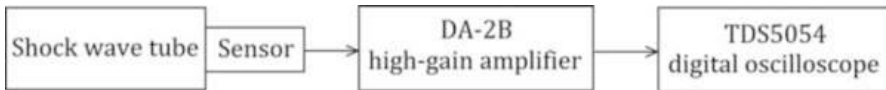
**Dynamic Performance Calibration of the Sensor**

The dynamic performance test of the high-temperature and high-frequency pressure sensor was carried out by means of shock wave tube in Xi’an Institute of Nuclear Technology; the testing site is shown in Fig. 49.

The main test equipment is shock wave tube, Tektronix TDS 5054 digital oscilloscope, IWATSU DA-2B high-gain amplifier (10 times magnification, operating frequency 0~1 MHz), HH1710 dual regulator power supply. During the test, the impact pressure of the shock wave tube is 2 MPa; the sensor is installed on the end face flange of the low-pressure chamber. Vacuum silicon grease is used to fill the gap at the end face of the sensor, so as not to affect the quality of the waveform. The test



**Fig. 49** Testing site of shock wave tube



**Fig. 50** Device connection diagram

equipment is connected, as shown in Fig. 50. After the sensor signal is amplified by the IWATSU DA-2B high-gain amplifier, the response waveform is recorded by the TDS5054 digital oscilloscope.

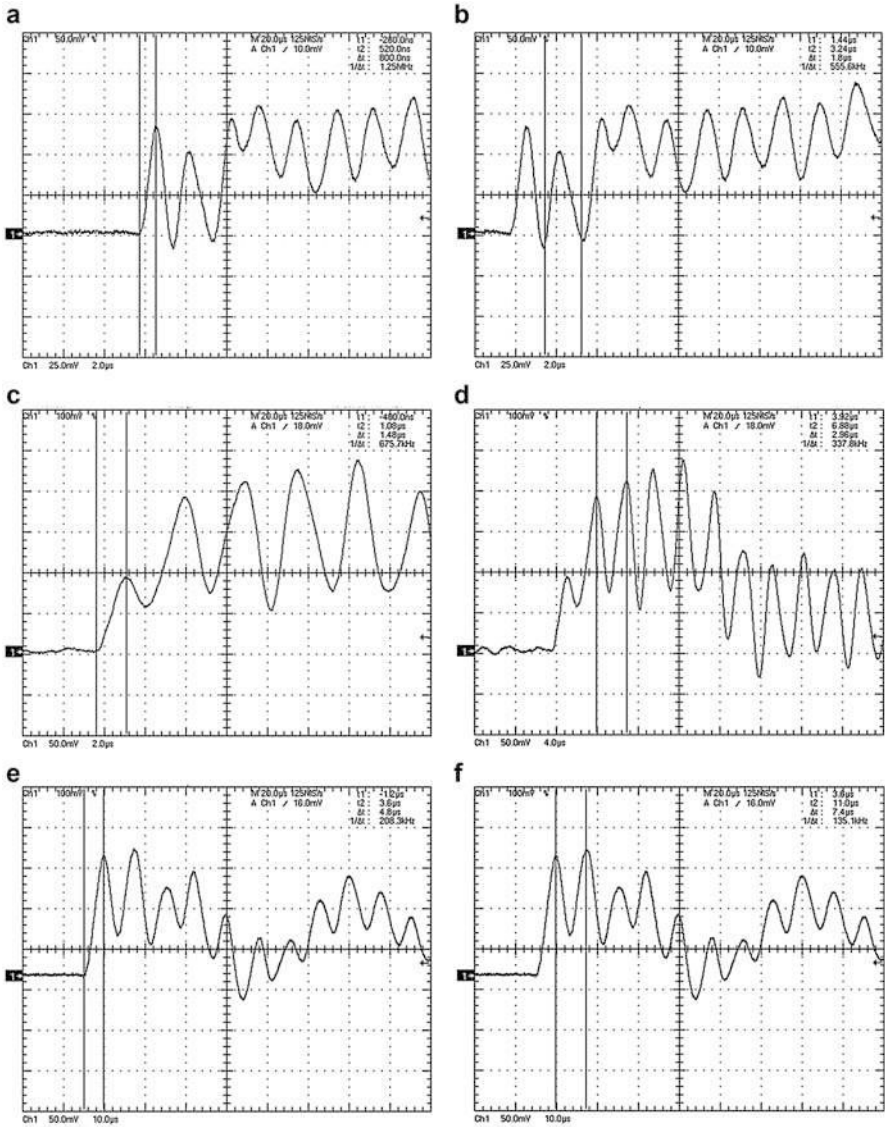
The pressure sensors with ranges of 25, 10, and 5 MPa are tested, respectively; the response waveforms are shown in Fig. 51; the response time and response frequency of the sensors can be roughly calculated from the measurement lines in the figures.

It can be seen from Fig. 51 that for high-temperature and high-frequency pressure sensor with the inverted-cup flush structure, the higher the range is, the thicker the sensitive diaphragm is, so that the higher the response frequency is.

## Conclusions

This chapter presented a complete development process for high-temperature silicon pressure sensor. First, the silicon-based piezoresistive effect together with SIMOX technology was discussed for pressure sensor's high-temperature application. Then the basic mechanics model for the sensor chip had been discussed which was mainly based on Navier trigonometric series and superposition method. The mechanics model gave a designing guide for the sensor chip structure designing, and two designing principles were proposed to take full use of the elastic strain energy of sensor chip and find a performance balance between the sensitivity and dynamic performance. Also, the lithography mask designing for corresponding structure was presented together with the sensor chip designing. Then, the sensor chip fabrication process was presented. In order to ensure the sensor chips' performance in harsh environment, several packaging structures for sensor chip in high-pressure, high-temperature application were presented to improve the sensor chip's performance.





**Fig. 51** Dynamic response curve of the sensor. (a) Measuring range: 25 MPa, response time <math>< 800\text{ ns}</math>. (b) Measuring range: 25 MPa, response frequency: about 555.6 KHz. (c) Measuring range: 10 MPa, response time <math>< 1.48\ \mu\text{s}</math>. (d) Measuring range: 10 MPa, response frequency: about 337.8 KHz. (e) Measuring range: 5 MPa, response time <math>< 4.8\ \mu\text{s}</math>. (f) Measuring range: 5 MPa, response frequency: about 135.1 KHz

Finally, the experimental calibration process and evaluation criterion for the sensor chips' performance were presented and analyzed to show the sensor chip with a good overall performance.

## References

- Abeysinghe DC, Dasgupta S, Jackson HE, Boyd JT (2002) Novel MEMS pressure and temperature sensors fabricated on optical fibers. *J Micromech Microeng* 12:229–235
- Bao MH (2000) *Micro mechanical transducers: pressure sensors, accelerometers and gyroscopes*. Elsevier, Amsterdam, pp 247–250
- Bashir R et al (2000) On the design of piezoresistive silicon cantilevers with stress concentration regions for scanning probe microscopy applications. *J Micromech Microeng* 10(4):483
- Cai HC, Yang H (2004) *Mechanics of materials*. Xi'an Jiaotong University Press, Xi'an
- Cano MP, Hernandez A, Ancos B (1997) High pressure and temperature effects on enzyme inactivation in strawberry and orange products. *J Food Sci* 62(1):85–88
- Chen L, Mehregany M (2008) A silicon carbide capacitive pressure sensor for in-cylinder pressure measurement. *Sensors Actuators A Phys* 145–146(1):2–8
- Dehe A, Fricke K, Mutamba K et al (1998) A piezoresistive GaAs pressure sensor with GaAs/AlGaAs membrane technology. *J Micromech Microeng* 5(2):139
- Fielder RS (2002) High-temperature high-bandwidth fiber optic MEMS pressure-sensor technology for turbine-engine component testing. *Proc SPIE Int Soc Opt Eng* 4578:229–238
- Hopcroft MA, Nix WD, Kenny TW (2010) What is the young's modulus of silicon? *J Microelectromech Syst* 19(2):229–238
- Hsu TR (2002) *MEMS & microsystems: design and manufacture. Integrated manufacturing systems*. McGraw-Hill, Boston, pp 49–50
- Jiang Z (2013) Special micro-electro-mechanical systems pressure sensor. *J Mech Eng* 49(6):187
- Kinnell PK, King J, Lester M et al (2010) A hollow stiffening structure for low-pressure sensors. *Procedia Chem* 160(1–2):35–41
- Maluf N (2000) *An introduction to microelectromechanical systems engineering*. Artech House, Boston/London, p 17
- Niu Z, Zhao Y, Tian B (2014) Design optimization of high pressure and high temperature piezoresistive pressure sensor for high sensitivity. *Rev Sci Instrum* 85(1):015001
- Nwafor OMI (2003) The effect of elevated fuel inlet temperature on performance of diesel engine running on neat vegetable oil at constant speed conditions. *Renew Energy* 28(2):171–181
- Okojie RS, Ned AA, Kurtz AD (1998) Operation of  $\alpha(6H)$ -SiC pressure sensor at 500 °C. *Sensors Actuators A Phys* 66(1–3):200–204
- Qian HE, Zhao YL, Zhao LB et al (2008) Research on packaging technology of high temperature pressure transducer. *Chin J Sensors Actuators* 21(2):310–313
- Roark RJ (2003) *Formulas for stress and strain*. McGraw-Hill Publishing Company LTD, New York
- Sato Y, Yurugi M, Fujiwara K et al (1996) Solubilities of carbon dioxide and nitrogen in polystyrene under high temperature and pressure. *Fluid Phase Equilib* 125(1):129–138
- Smith CS (1954) Piezoresistance effect in germanium and silicon. *Phys Rev* 93:42–49; *Phys Rev* 94(1):42–49
- Stankevič V, Šimkevičius Č (2000) Use of a shock tube in investigations of silicon micromachined piezoresistive pressure sensors. *Sensors Actuators A Phys* 86:58–56
- Stuchechnikov VM (1991) SOS strain gauge sensors for force and pressure transducers. *Sensors Actuators A Phys* 28(3):207–213
- Timoshenko SP (1936) *Theory of elastic stability*. McGraw-Hill Book Co, New York
- Timoshenko SP, Woinowsky-Krieger S (1959) *Theory of plates and shells*. McGraw-Hill, New York, pp 105–225
- Xian H, Zhang DC (2014) A high sensitivity and high linearity pressure sensor based on a peninsula-structured diaphragm for low-pressure ranges. *Sensors Actuators A Phys* 216:176–189
- Xu T, Zhao L, Jiang Z et al (2016a) Modeling and analysis of a novel combined peninsula-island structure diaphragm for ultra-low pressure sensing with high sensitivity. *J Phys D Appl Phys* 49(7):075110
- Xu T, Zhao L, Jiang Z et al (2016b) A high sensitive pressure sensor with the novel bossed diaphragm combined with peninsula-island structure. *Sensors Actuators A Phys* 244:66–76

- Yu JC, Lan CB (2001) System modeling of microaccelerometer using piezoelectric thin films. *Sensors Actuators A Phys* 88(2):178–186
- Yu Z, Zhao Y, Sun L et al (2013) Incorporation of beams into bossed diaphragm for a high sensitivity and overload micro pressure sensor. *Rev Sci Instrum* 84(1):530–224
- Yu ZL, Zhao YL, Li LL et al (2015a) Realization of a micro pressure sensor with high sensitivity and overload by introducing beams and islands. *Microsyst Technol* 21:739–747
- Yu Z, Zhao Y, Li L et al (2015b) Realization of a micro pressure sensor with high sensitivity and overload by introducing beams and islands. *Microsyst Technol* 21(4):1–9
- Zhao L, Zhao Y, Jianbo LI et al (2010) Inverted-cup high-temperature and high-frequency piezoresistive pressure sensor. *J Xi'an Jiaotong Univ* 27(1):96–100
- Zhao L, Xu T, Hebibul R et al (2016) A bossed diaphragm piezoresistive pressure sensor with a peninsula–island structure for the ultra-low-pressure range with high sensitivity. *Meas Sci Technol* 27(12):124012



# A Micromachined Silicon Resonant Pressure Sensor

Junbo Wang, Deyong Chen, Bo Xie, Jian Chen, Lin Zhu, and Yulan Lu

## Contents

Introduction .....	388
Working Principle .....	389
Design and Fabrication .....	392
Electromagnetically Driven and Sensed Resonant Pressure Sensors .....	393
Electrostatically Driven and Capacitive Sensed Resonant Pressure Sensors .....	400
Self-Temperature Compensation .....	409
The Self-Temperature Compensation Algorithm .....	410
Calibration for Self-Compensation of the Resonant Pressure Sensor .....	411
Compensation Results and Analysis .....	411
Conclusions .....	416
References .....	416

## Abstract

Micromachined silicon resonant pressure sensors have been widely used in automotive industry, medical instrument, aerospace, and military fields due to their high accuracy, long-term stability, and quasi-digital output. This chapter begins with the introduction of the working principle of the resonant pressure sensors, illustrating key relationships between (1) intrinsic resonant frequency and structural parameters, (2) pressure under measurement and resonant frequency shift, and (3) device sensitivity and structural parameters. Then, two

---

J. Wang (✉) · D. Chen · B. Xie · J. Chen · L. Zhu · Y. Lu  
State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China  
e-mail: [jbwang@mail.ie.ac.cn](mailto:jbwang@mail.ie.ac.cn); [dychen@mail.ie.ac.cn](mailto:dychen@mail.ie.ac.cn); [xiebo11@mails.ucas.ac.cn](mailto:xiebo11@mails.ucas.ac.cn);  
[chenjian@mail.ie.ac.cn](mailto:chenjian@mail.ie.ac.cn); [lzhu\\_ucas@163.com](mailto:lzhu_ucas@163.com); [luyulan15@mails.ucas.ac.cn](mailto:luyulan15@mails.ucas.ac.cn)

kinds of micromachined silicon resonant pressure sensors based on electromagnetic and electrostatic excitations are presented, respectively, where device design, simulation, fabrication, and packaging are discussed in details. Finally, self-temperature compensation approaches are introduced to improve the performance of the micromachined silicon resonant pressure sensors, which can therefore function in a wide temperature range.

---

**Keywords**

Micromachined silicon · Resonant pressure sensor · Sensor design · MEMS fabrication · Through-glass vias · Through-silicon vias · Anodic bonding · Vacuum packaging · Self-temperature compensation

---

**Introduction**

Micromachined silicon pressure sensors are widely used in the fields of automotive industry, medical instrument, aerospace, and military fields due to their advantages of small sizes, high resolutions, and low costs (Eaton and Smith 1997). Based on the detection mechanisms, they can be classified into capacitive pressure sensors (Kasten et al. 2000; Santo et al. 2010), piezoresistive pressure sensors (Damjanovic 1998; Schulz et al. 2009), piezoelectric pressure sensors (Pramanik et al. 2006; Otmani et al. 2011), and resonant pressure sensors (Kinnell and Craddock 2009). Compared to the other types of micromachined silicon pressure sensors, resonant pressure sensors are featured with the advantage of “quasi-digital” outputs, which allows direct coupling to digital electronics without analog-to-digital converters, leading to high accuracy and reliability. Additionally, resonant pressure sensors have excellent long-term stabilities since the resonance frequency belongs to the intrinsic material properties and geometrical parameters of mechanical resonators, which is not dependent on the unstable or drifting electrical signals (Greenwood and Satchell 1988; Greenwood and Wray 1993).

This chapter introduces the working principle of the micromachined silicon resonant pressure sensors firstly, by highlighting the relationships between (1) intrinsic resonant frequency and device structural parameters, (2) pressure under measurement and device resonant frequency shift, and (3) device sensitivity and structural parameters.

In the following section, two kinds of silicon micromachined resonant pressure sensors, including the electromagnetically driven resonant pressure sensor and the electrostatically driven resonant pressure sensor, are discussed, respectively. The design, simulation, fabrication, and packaging for the micromachined silicon resonant pressure sensor are demonstrated in details. Then, in order to improve the performance of the micromachined silicon resonant pressure sensor, temperature compensation is introduced, enabling the functionality of the pressure sensor in a wide temperature range.

## Working Principle

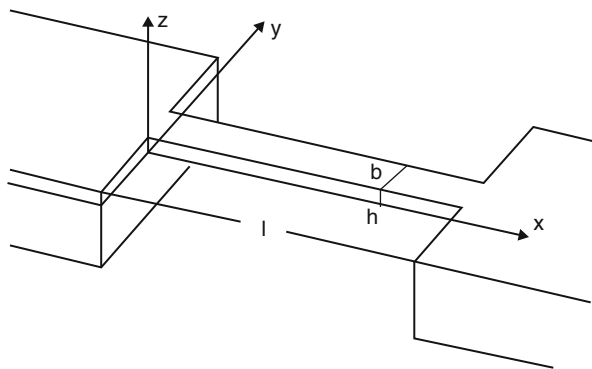
Resonant pressure sensors typically use a resonating mechanical structure such as a vibrating beam as a strain gauge to sense the deflection of the pressure-sensitive diaphragm. The basic principle of a pressure sensor is that the pressure deflects a diaphragm, which in turn applies axial stresses to a beam that is vibrating at its resonant frequency. The change in the resonant frequency is used as an output to indicate the applied pressure based on calibration. Firstly, pressure deflects the diaphragm, which in turn applies an axial stress to the resonant beam strain gauge. Taking the two-part interaction into account, the analytical modeling is divided in two parts: the stress-frequency response of the beam and the pressure-tensile stress response between the diaphragm and the resonant beam. Although a completely analytical model will probably not be numerically accurate, it will be key to understand how the sensor works and provide insights into the most important parameters for the design of the sensor such as pressure sensitivity, resonant frequency, etc.

The resonant beam shown in Fig. 1 is modeled first as a transversely vibrating prismatic beam with the length  $l$  in  $x$  direction, the width  $b$  in  $y$  direction, and the height  $h$  in  $z$  direction, assuming that the transverse and angular deflections of the beam are small and that they only occur in a 2D ( $x$  and  $z$ ) manner (Chen 2002). Thus, the twisting-type torsion of the beam is not considered.

The axial stress is assumed to be constant throughout the length of the beam and not significantly affected by the deflections. Since the deflections are considered small, the rotational inertia effects are also neglected. Finally, the ends are assumed to be clamped, so that both the deflection and its derivative with respect to  $x$  are 0. Regarding geometry, it is assumed that the beam is prismatic and of uniform composition. Following above assumptions, the differential equation of motion of a transversely vibrating beam with axial tension is given by.

$$EI \frac{\partial^4 \omega(x, t)}{\partial x^4} - \sigma A \frac{\partial^2 \omega(x, t)}{\partial x^2} = -\rho A \frac{\partial^2 \omega(x, t)}{\partial t^2} \tag{1}$$

**Fig. 1** A model of a resonant beam



where  $\omega(x, t)$  is the deflection of the beam,  $E$  is the material's Young's modulus,  $I$  is the moment of the inertia,  $\sigma$  is the axial stress,  $\rho$  is the material's density, and  $A$  is the area of the beam cross section. In addition, the moment of the inertia  $I = \frac{bh^3}{12}$ , and the cross-sectional area  $A = bh$ .

When this beam resonator undergoes time-harmonic vibration, we can assume

$$\omega(x, t) = W(x)e^{j\omega t} \quad (2)$$

where  $\omega$  denotes the angular frequency of the vibration. Substituting Eq. 2 into Eq. 1, with the boundary condition of  $\omega|_{x=0, l} = 0$  and  $\frac{\partial \omega}{\partial x}|_{x=0, l} = 0$ , we can obtain the frequency equation with the axial stress  $\sigma$ .

$$2 - 2\cosh\lambda_1 l \cos\lambda_2 l + \frac{\alpha}{\beta} \sinh\lambda_1 l \sin\lambda_2 l = 0 \quad (3)$$

where  $\lambda_1 = \sqrt{\frac{\alpha}{2} + \sqrt{\frac{\alpha^2}{4} + \beta^2}}$ ,  $\lambda_2 = \sqrt{-\frac{\alpha}{2} + \sqrt{\frac{\alpha^2}{4} + \beta^2}}$ ,  $\alpha = \frac{\sigma A}{EI}$ , and  $\beta = \sqrt{\frac{\rho A \omega^2}{EI}}$ .

When the axial stress  $\sigma = 0$ , Eq. 3 can be simplified as.

$$1 - \cosh\beta l \cos\beta l = 0 \quad (4)$$

Solutions for this eigenvalue equation can be found numerically, and the first five solutions of Eq. 4 are  $\beta l = 4.73004, 7.785320, 10.9956, 14.13716, \text{ and } 17.2788$ . And the resonant frequency of the base mode without stress is

$$f_1(0) = \frac{4.73^2}{2\pi l^2} \sqrt{\frac{EI}{A\rho}} = 1.028 \sqrt{\frac{E}{\rho}} \cdot \frac{h}{l^2} \quad (5)$$

When the beam is subjected to an axial stress  $\sigma$  (or strain  $\varepsilon$ ), the first-order resonance frequency will change with the axial strain approximately according to the following equation:

$$f_1(\sigma) = f_1(0) \sqrt{1 + \frac{\sigma}{\sigma_c}} \quad (6)$$

where  $\sigma_c = \frac{\pi^2 E}{3} \left(\frac{h}{l}\right)^2$  is called critical Euler bucking load.

Let  $\Delta f = f - f_0$  be the resonant frequency shift induced by the axial stress, when  $\sigma \ll \sigma_c$ , Eq. 6 can be written into a first-order approximation as

$$\frac{\Delta f}{f_0} = \frac{3}{2\pi^2} \left(\frac{l}{h}\right)^2 \frac{\sigma}{E} \quad (7)$$

Therefore, resonant frequency shift is proportional to the axial stress  $\sigma$  or strain  $\varepsilon(\sigma/E)$ . The total stress is equal to the sum of the individual stress terms,  $\sigma = \sigma_0 + \sigma_T + \sigma_P$ , where  $\sigma_0$  stands for the residual stress resulting from the fabrication

process,  $\sigma_T$  is the thermal stress, and  $\sigma_P$  is the pressure load-induced stress. Both the excitation and packaging will induce the thermal stress when environmental temperature changes. Therefore, temperature compensation methods should be introduced to further improve the accuracy of the sensor, which will be introduced in section “Self-Temperature Compensation.”

The next step is to model the diaphragm-beam interaction to estimate the basic relationship between applied pressure on the diaphragm and axial stress on the resonant beam.

The deflection of a rectangular diaphragm under uniform pressure load may be found using elasticity theory and a plate approximation, assuming a small deflection relative to the plate thickness. Local deflection  $\omega(x,y)$  in the  $z$  direction can be found by solving partial differential equation:

$$\frac{\partial^4 \omega}{\partial^4 x} + 2 \frac{\partial^4 \omega}{\partial^2 x \partial^2 y} + \frac{\partial^4 \omega}{\partial^4 y} = \frac{P}{D} \quad (8)$$

where  $\omega(x,y)$  is the local deflection and  $x$  and  $y$  are the coordinates based on a rectangular area  $-a \leq x \leq a$  and  $-a \leq y \leq a$ . For the geometry and boundary conditions given an applied uniform pressure  $P$  and plate stiffness  $D$ ,  $D$  is expressed by

$$D = \frac{Et^3}{12(1 - \mu^2)} \quad (9)$$

where  $E$  is Young’s modulus of the material,  $t$  is the plate thickness, and  $\mu$  is Poisson’s ratio for the material, which for these equations are assumed to be isotropic.

Galerkin method is utilized to solve the equation of motion of distributed-parameter systems. The deflection shape is approximated as a cosine in both directions based on the clamped edge boundary conditions. Taking the following trial function for the local deflection  $\omega(x,y)$  in the  $z$  direction

$$\omega(x,y) = \frac{1}{4} \omega_0 \left(1 + \cos \frac{\pi x}{a}\right) \left(1 + \cos \frac{\pi y}{a}\right) \quad (10)$$

where  $\omega_0$  is the maximum deflection in the center of the diaphragm where  $x = y = 0$ .

From Eqs. 8, 9, and 10,  $\omega_0$  may be found as

$$\omega_0 = \frac{2a^4 P}{\pi^4 D} = \frac{24(1 - \mu^2)a^4 P}{Et^3} \quad (11)$$

The axial stress  $\sigma_x(z)$  has relationship with the local deflection  $\omega(x,y)$  expressed by

$$\sigma_x(z) = -\frac{Ez}{1 - \mu^2} \left( \frac{\partial^2 \omega}{\partial x^2} + \mu \frac{\partial^2 \omega}{\partial y^2} \right) \quad (12)$$



At the upper surface of the diaphragm ( $z = t/2$ ), the axial stress  $\sigma_x$  along the  $x$  axis ( $y = 0$ ) is given by

$$\sigma_x|_{z=\frac{t}{2}}^{y=0} = \frac{3P}{\pi^2} \left(\frac{a}{t}\right)^2 \left[ (2 + \mu) \cos \frac{\pi x}{a} + \mu \right] \quad (13)$$

When  $x = 0$ , the axial stress  $\sigma_x$  takes a maximum stress  $\sigma_{x\max}$  given by

$$\sigma_{x\max} = \frac{6(1 + \mu)}{\pi^2} \left(\frac{a}{t}\right)^2 P \quad (14)$$

$\sigma_{x\max}$  is positive corresponding to a tensile stress.

When  $x = \pm a$ , the axial stress  $\sigma_x$  takes a minimum stress  $\sigma_{x\min}$  given by

$$\sigma_{x\min} = -\frac{6}{\pi^2} \left(\frac{a}{t}\right)^2 P \quad (15)$$

$\sigma_{x\min}$  is negative corresponding to a compressive stress.

As seen from Eqs. 13, 14, and 15, the axial stress  $\sigma_x$  has a linear relationship with the applied pressure  $P$ ; the ratio of  $\sigma_x/P$  is called the conversion coefficient of applied pressure to axial stress, which varies with the location on the diaphragm and is proportional to the square of the ratio of the width to the thickness.

It may be noted that if the resonant beam is located at the upper center of the diaphragm, from Eq. 14, Eq. 7 can be rewritten as

$$\frac{\Delta f}{f_0} = \frac{9(1 + \mu)}{\pi^2} \left(\frac{l}{h}\right)^2 \left(\frac{a}{t}\right)^2 \frac{P}{E} \quad (16)$$

The pressure sensitivity (denoted by  $S$ ) is defined as the ratio of  $\Delta f/f_0$  to the applied pressure  $P$ ; therefore,  $S$  can be expressed by

$$S = \frac{\Delta f}{f_0} / P = \frac{9(1 + \mu)}{\pi^2 E} \left(\frac{l}{h}\right)^2 \left(\frac{a}{t}\right)^2 \quad (17)$$

From Eq. 17, the pressure sensitivity  $S$  caused by the pressure load on the diaphragm is proportional to the square of both the aspect ratio  $l/h$  and  $a/t$ . Therefore, given the structure parameters of the beam and diaphragm, one can easily predict the sensor performance such as resonant frequencies and frequency response to the applied pressure prior to fabrication through the above analytic solution.

---

## Design and Fabrication

Utilizing the resonant characteristics of the sensor, an oscillator must be constructed. That means the mechanical resonator structure has to be brought into vibration and the vibration has to be detected to obtain the resonant frequency output which will be a reliable indication of the measurement.

Numerous excitation/detection schemes are employed to construct such an oscillator, including electrostatic excitation/capacitive detection or piezoresistive detection, electromagnetic excitation/electromagnetic detection, electrothermal excitation/piezoresistive detection, piezoelectric excitation/piezoelectric detection, optothermal excitation/optical detection, and dielectric excitation/capacitive detection. There are trade-offs associated with each of these schemes. For micromachined silicon resonators, electrostatic, electromagnetic, or thermal excitation mechanisms and capacitive, electromagnetic, piezoresistive detection mechanisms are commonly used. In the following section, we mainly introduce two typical kinds of micromachined silicon resonant pressure sensors, relying on electromagnetic and electrostatic excitations, respectively.

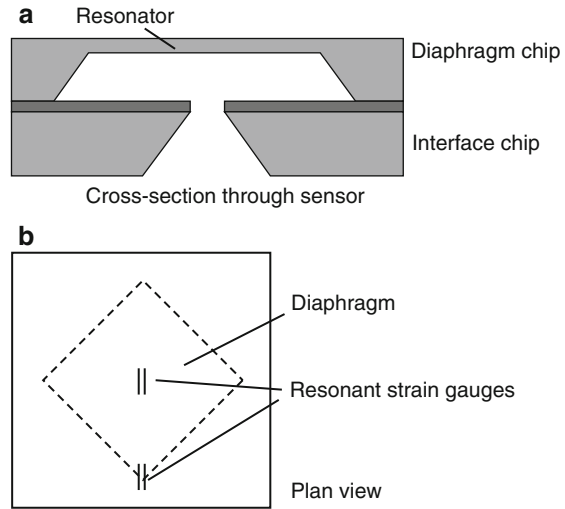
### **Electromagnetically Driven and Sensed Resonant Pressure Sensors**

With electromagnetic excitation/electromagnetic detection, the resonator (usually refer to a beam), which lies perpendicular to a static magnetic field and carries an AC current, experiences a Lorentz force to excite the beam into vibration, and the resulting vibrations are sensed by detecting the magnetically induced voltage developed across the beam as it vibrates in the magnetic field through Faraday's law. If only one beam is used, which is called one-port device, both the drive and sense signal appear superimposed across the beam with the magnitude of the sense voltage much smaller than that of the drive signal, resulting to a difficult detection of the oscillation. To avoid this problem, a two-port device with an "H"-shaped structure is commonly utilized for electromagnetically driven and sensed resonant pressure sensors. By using one side of the "H" structure as the drive arm and the other as the detection arm, the superposition of the drive and sense signals is eliminated.

Based on the two-port electromagnetically driven and sensed mechanism, a successfully commercialized device was developed by the Yokogawa Electric Corporation (Ikeda et al. 1990, 1991; Harada et al. 1996) employing surface micromachining technologies. It consists of two resonators located on a diaphragm, the differential output of which provides the sensor reading. The resonators are driven electromagnetically by placing the device in a magnetic field and running an alternating current through the structure. The pressure sensor arrangement is shown in Fig. 2. The fabrication process associated with this device is particularly impressive. The beams are vacuum encapsulated at wafer level using a series of epitaxial depositions, selective etches, and finally annealing in nitrogen, which drives the trapped gases left by the sealing process through the cavity walls or into the silicon. This leaves a final cavity pressure of below 1 mTorr, and the resonator possesses a Q-factor of more than 50,000.

Using bulk micromachining rather than surface micromachining, D. Chen and J. Wang's team (Wang et al. 2008a, b, 2009; Chen et al. 2009, 2010a, b; Li et al. 2011, 2012, 2013; Luo et al. 2013, 2014a, b, c, 2015) developed two kinds of resonant pressure sensors with lateral vibration modes where BCB adhesive bonding and silicon-glass anodic bonding were used for vacuum packaging, respectively.

**Fig. 2** The cross-sectional view (a) and the top view (b) of the Yokogawa differential resonant pressure sensor

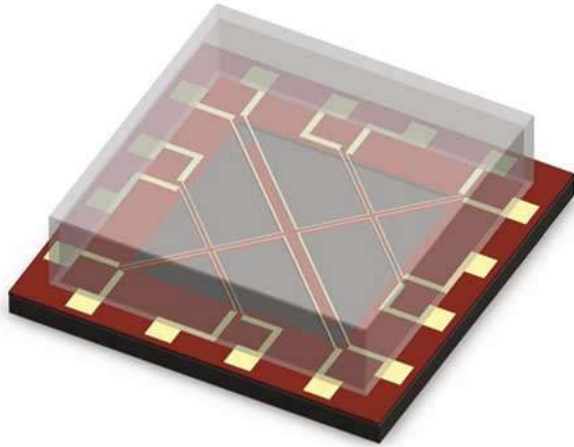


### Resonant Pressure Sensors with BCB Adhesive Bonding

Hermetic packaging is a major bottleneck in fabricating resonant pressure microsensors. In hermetic packaging, electrical interconnection and stress-relieving packaging need to be considered, which significantly add the cost and complexity of packaging. Adhesive wafer bonding possesses a number of advantages compared to established bonding techniques such as silicon-to-glass anodic bonding or silicon-to-silicon bonding, which require complex procedures, high temperature, and critical pretreatments. Since adhesive bonding can be conducted in a much lower temperature, it is compatible with fabrication of microelectronics and can function as a post-process module in IC industry. As one type of organic bonding materials, BCB is widely used in MEMS packaging. The advantages of BCB include low dielectric constant, low dissipation factor at high frequencies, low moisture absorption, high degree of planarization, and high compatibility with various metallization systems and conventional IC processing.

The schematic of the differential resonant pressure sensor is shown in Fig. 3 (Chen et al. 2010; Li et al. 2011, 2013). The resonant elements consist of four clamped-clamped beams suspended on a silicon square diaphragm by three anchors and a silicon rectangular frame, which are organized symmetrically along the diagonal direction. Thick boron-diffused silicon-based H-type beams (30  $\mu\text{m}$  in thickness) are used as resonators so that they can work in a lateral mode to minimize the mechanical coupling between the diaphragm and the resonator as their fundamental resonant modes are perpendicular to each other. A non-photosensitive BCB-based low-temperature adhesive bonding was used to realize vacuum encapsulation. This organic glue-based adhesive bonding has the advantages of low temperature, high flexibility, and low cost. In addition, this approach can facilitate electrical connections by patterning electrodes across the bonding interface.

**Fig. 3** The schematic diagram of a resonant pressure sensor based on electromagnetically driven and sensed H-type lateral beam resonators (Luo et al. 2014a, b)



However, this method suffers from large residual stresses in deep boron-diffused silicon and low long-term vacuum levels limited by the physical properties of BCB.

### Resonant Pressure Sensors with Anodic Bonding

In order to sustain the resonator's oscillation and maintain the desired frequency stability, the energy coupled through the resonator's supports to the surroundings should be minimized by proper resonator design, and the viscous damping from the surrounding atmosphere must also be reduced by placing the device in vacuum. An appropriate resonator design may provide immunity to surrounding structural resonances and vibrations and improves the long-term drift potential of the sensor since the influence of the surrounding structure on the resonator is minimized. The mechanical quality factor of a microresonator deteriorates as the environmental pressure increases due to the air damping effect (Welham et al. 1999; Mandle et al. 1995). To achieve high performance, resonators are commonly isolated from the environment by vacuum packaging (Li et al. 2011; Choa 2005). However, for MEMS-based pressure sensors, vacuum packaging cannot be realized by conventional approaches based on ceramic or metal hermetic sealing, due to their movable and fragile structures.

Thus, wafer-level bonding which can protect MEMS devices from subsequent processes is preferred to realize vacuum packaging (O'Mahony et al. 2009). Wafer-level vacuum packaging has the advantages of small size, low cost, and high compatibility with micro-fabrication processes where a cap wafer is bonded with the device substrate for vacuum packaging. The micro-cap can offer robust protections, protecting the fragile mechanical parts from possible impacts and destructions by subsequent processes (Esashi 2008; Cabuz et al. 1994). A variety of bonding techniques have been proposed for wafer-level vacuum packaging, such as intermediate layer bonding, silicon-silicon fusion bonding, and silicon-glass anodic bonding (Mitchell et al. 2005). Among these approaches, the electrical interconnections between the bonding micro-cap and the device substrate pose key challenges for vacuum packaging.

Through-glass vias (TGV) and silicon-to-glass anodic bonding were utilized to fabricate a resonant pressure sensor consisting of an SOI wafer and a Pyrex 7740 glass wafer (Beeby et al. 2000; Luo et al. 2014a, c). Presently, SOI has been extensively employed in MEMS devices because of its simple yet reliable fabrication steps, high yield, and robust structures. The fabrication procedures for the proposed sensor were based on simplified SOI-MEMS fabrication processes, requesting only three photolithographic steps.

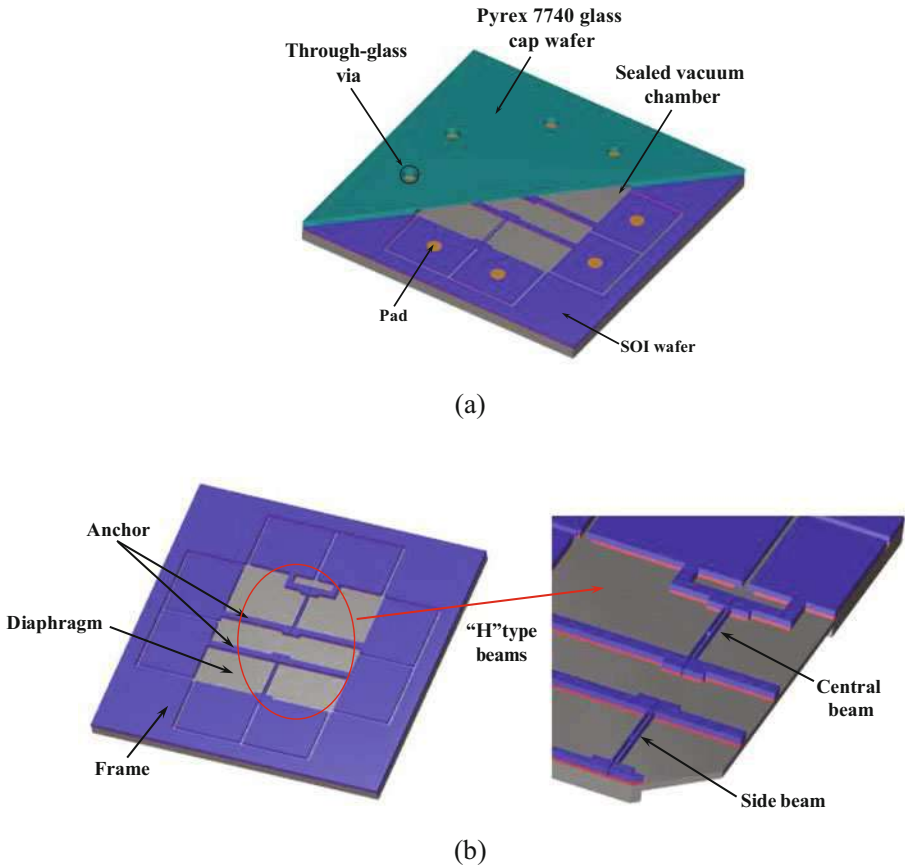
Compared to previous designs, this new device has several advantages (Luo et al. 2014a, c):

1. The top single-crystal silicon layer of the SOI wafer was used to form the resonators due to its super mechanical properties and high intrinsic quality factors. In addition, the proposed pressure sensor adopts a differential structure where two resonators with a comparable temperature dependency were employed to suppress side temperature effects on the resonant frequency drift.
2. A wafer-level vacuum packaging based on silicon-to-glass anodic bonding was utilized to form a sealing vacuum chamber, realizing long-term vacuum tightness. The anodic bonding technique (Mitchell et al. 2005) possesses the advantages of the relatively low required temperature (350–450 °C) and high bonding strength and hermeticity. A Pyrex 7740 glass wafer as the cap wafer was anodic bonded to the patterned SOI wafer for vacuum packaging.
3. The through-glass via (TGV) technology was used. After anodic bonding was completed, an aluminum film was sputtered into the via holes by a shadow mask to form the contacting pads for electrical interconnections. This approach ensured that the resonators can be electrically connected to outside by wire bonding through the front side of the assembled wafers.
4. A non-evaporable thin metal film as the getter material was embedded into the concave of the Pyrex 7740 glass cap to help maintain the high-vacuum condition, producing high-quality factors of resonators. The use of the getter film for wafer-level vacuum packaging was described elsewhere (Moraja and Amiotti 2003), which has two functions. Firstly, the metal film acts as a diffusion barrier to gas atoms. Secondly, getter materials such as titanium have the ability to absorb surrounding gases when activated at certain temperatures. Thus, the getter materials effectively reduce the trapped gases inside the cavity as well as the outgassing generated during the anodic bonding.

### Device Design

As shown in Fig. 4a, through-glass vias were realized using laser drilling into a Pyrex 7740 glass wafer (Luo et al. 2014c). Then, hermetic sealing was achieved by bonding the processed glass cap wafer with the prepatterned SOI wafer. The vias provide electrical connections of the pressure sensor.

The schematic of the differential resonant pressure sensor fabricated on a SOI wafer is shown in Fig. 4b. The resonant elements consist of two “H”-type doubly clamped beams suspended on a pressure-sensitive silicon square diaphragm. The single-crystal silicon “H”-type beams used as resonators work in a lateral mode,



**Fig. 4** The schematic diagram of the resonant pressure sensor (a) with SOI-glass anodic bonding and through-glass via holes (b) with two “H”-type doubly clamped beams as resonators (Luo et al. 2014a)

which are actuated and detected electromagnetically. The two beams named “central beam” (located in the center of the diaphragm) and “side beam” (located near the border of the diaphragm) have almost identical dimensions and thus comparable resonant frequencies at zero pressure loads (Chen et al. 2010; Li et al. 2011). In this device, pressure under measurement causes a deflection of the diaphragm, which is further translated to an axial tensile stress buildup in the central beam, while an axial compressive stress in the side beam, leading to resonant frequency, shifts, respectively, toward opposite directions, enabling a differential output.

**FEM Simulation**

Finite-element modeling was used in the design and optimization of the resonant pressure microsensor to deal with two concerns: mode interference and two-beam sensitivity mismatch. Mode interference or the frequency overlap of two nearby

resonant modes can drive the resonant beams to vibrate in undesirable modes, leading to mode crosstalk and energy loss with compromised quality factors. The sensitivity mismatch between the two resonant beams can lead to compromised performances in the differential resonant frequency output. The structure of the pressure sensor was simulated using the ANSYS finite-element package which was initially defined using the solid module. The key dimensions were parameterized, enabling the rapid simulation of design modifications and the high-efficiency optimization routine. Static simulations were used to calculate the stress distributions along the resonant microbeams as a function of pressure, with the purpose of optimizing the relative positions of the two resonant beams to address the issue of sensitivity mismatch. Then, mode simulations were used to locate the optimum vibration mode, which is capable of transferring stress of resonant beams into the intrinsic resonant frequency shift (Moraja and Amiotti 2003). A series of pressure values were applied on the backside of the diaphragm followed by mode simulations to extract all the intrinsic modes within the frequency range from 0 to 100 kHz to locate the desirable vibration mode. Details of the simulation results are summarized as follows (Luo et al. 2014a).

Figure 5a represents the stress contour along the axial direction of the beams under an applied pressure of 100 kPa. Positive stress (tensile) was induced in the center beam with a characterized value of 19.7 MPa, while negative stress (compressive) was observed in the side beam with a quantified value of  $-19.8$  MPa. Based on the analysis, two resonant beams were under comparable stresses, enabling a differential output.

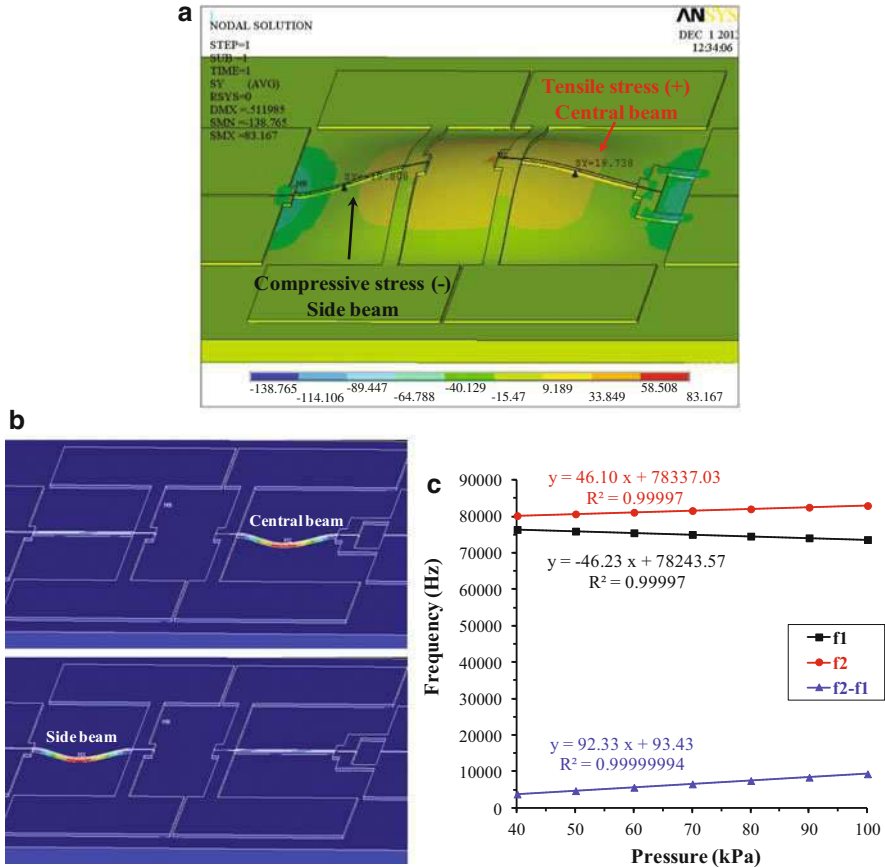
Figure 5b shows the optimum vibration mode for the resonant beams featured with lateral vibration in the wafer plane. Lateral vibrations are preferred in this study since the mechanical coupling between energy loss through the resonator supports and resonant modes of the surrounding structure is minimized since its vibration cannot lead to the co-vibration of the sensitive diaphragm.

Figure 5c shows the FEM simulation results of resonant frequencies of the side beam ( $f_1$ ) and the central beam ( $f_2$ ) as a function of applied pressure where the linear coefficients of the both beams were quantified as 0.9999 and the linear correlation coefficient of the differential output ( $f_2 - f_1$ ) was quantified as 0.9999999. In addition, a twice sensitivity of the differential output compared to the case of single beam was also noticed.

## Device Fabrication

Conventional MEMS bulk-silicon fabrication processes including deep reactive ion etching, photolithography, and sacrificial layer release were used to fabricate the proposed resonant pressure microsensors starting from a standard 4 in. SOI wafer ( $40 + 2 + 300$  ( $\mu\text{m}$ )). The fabrication process is described in Fig. 6 where through-glass holes were realized using laser drilling into a Pyrex 7740 glass wafer ( $500$  ( $\mu\text{m}$ )).

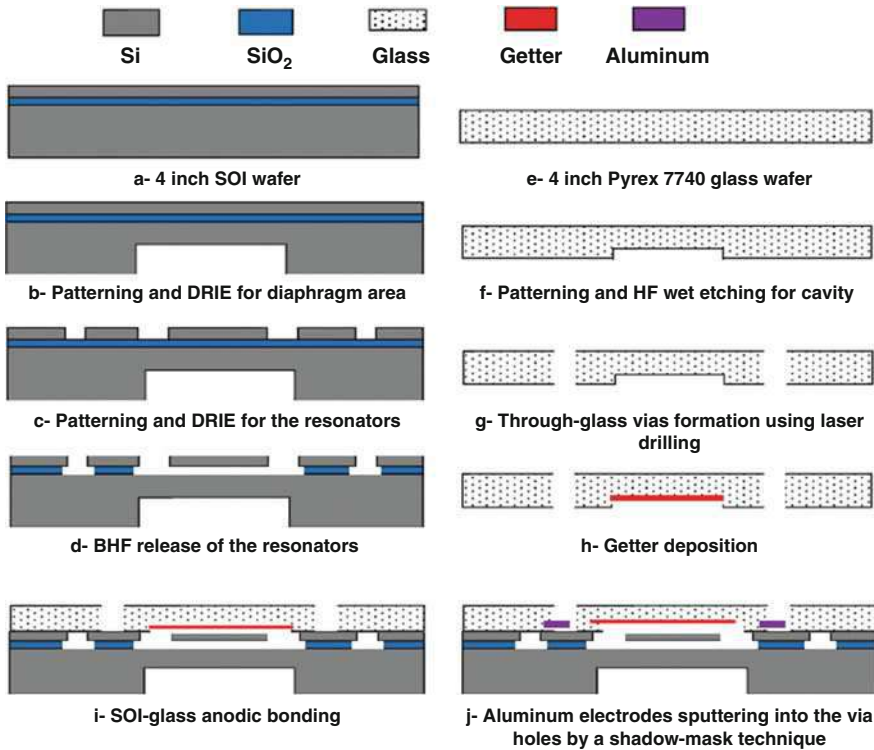
The fabrication of the resonators was based on a 4 in. SOI wafer (Fig. 6a–d) (Luo et al. 2014a). Only two photolithographic steps were needed to fabricate the resonators. Initially, using a patterned positive photoresist as a mask, the handle layer was etched by DRIE to the depth of  $120$  ( $\mu\text{m}$ ), forming the pressure-sensitive diaphragm (Fig. 6b). Secondly, using a patterned aluminum film and positive



**Fig. 5** Simulation results. (a) The stress contour along the axial direction of the resonant beams under an applied pressure of 100 kPa. Positive stress (tensile) was induced in the center beam, while negative stress (compressive) was observed in the side beam. (b) Maximal displacements of the two resonant beams in the first-order lateral vibration mode. (c) Simulated resonant frequencies of the two resonant beams as a function of applied pressure (Luo et al. 2014a)

photoresist as the mask, the exposed device layer of the SOI wafer was etched by DRIE to a depth of 40 μm to define the resonant beams (Fig. 6c). Then, the SOI wafer was immersed in a buffered hydrofluoric acid (BHF) solution, releasing the resonant beams by undercutting the insulation layer in a time-controlled manner (Fig. 6d). For the cap wafer, a Pyrex 7740 glass wafer (500 μm) was used. The fabrication process flow of the cap wafer is shown in Fig. 6(e-h). Based on deposition and patterning of a metallic mask, the cavity was etched using wet HF (Fig. 6f). Then, after formation of through-glass holes using laser drilling, the getter film was deposited inside the cavity (Fig. 6h). Next, silicon-to-glass anodic bonding was utilized to form a sealed vacuum chamber by bonding the patterned SOI wafer with the Pyrex 7740 glass wafer (Fig. 6i). Then, the Al electrodes were sputtered on





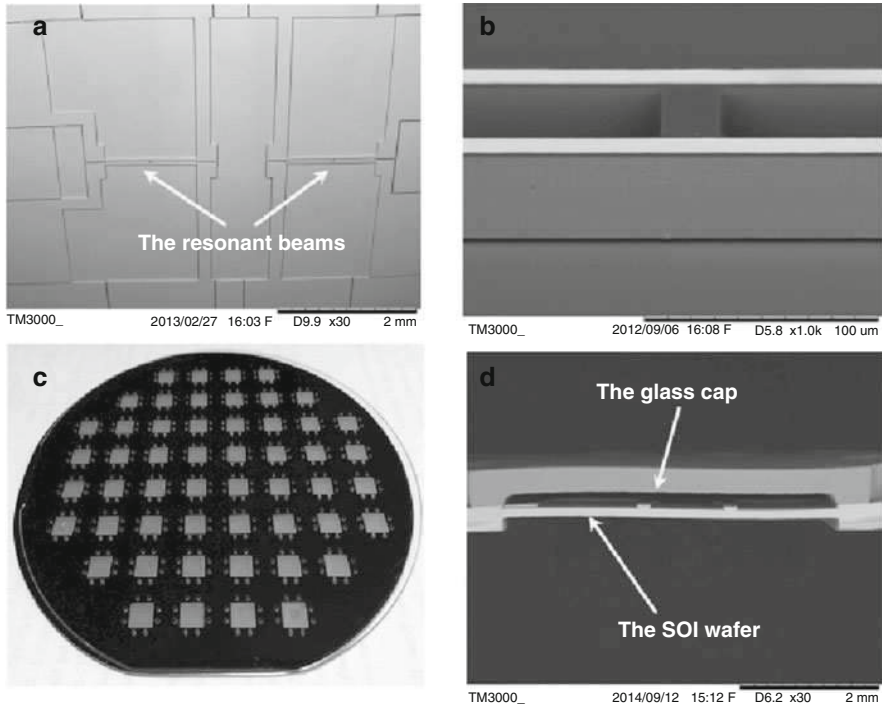
**Fig. 6** The fabrication procedure for the resonant pressure microsensors, relying on SOI-MEMS fabrication processes (Luo et al. 2014a)

the front side of the bonded SOI-glass wafer through the via holes by a shadow mask to electrically connect to the outside world (Fig. 6j). In the end, the fabricated wafer was diced into individual sensor units.

The proposed resonant pressure microsensors were successfully fabricated by MEMS bulk micromachining (Luo et al. 2014a). Fig. 7(a) and (b) show the SEM images of the top view of the fabricated SOI wafer and the “H”-type doubly clamped beams, respectively. As shown in Fig. 7c, a Pyrex 7740 glass cap wafer with through-glass via holes was anodically bonded to the patterned SOI wafer, forming a hermetic sealing for the resonators. Fig. 7d shows the SEM image of the bonding interface of a sensor unit after wafer dicing.

## Electrostatically Driven and Capacitive Sensed Resonant Pressure Sensors

Resonators driven by electrostatic forces have been most widely used (Hirano et al. 1992; Yoneoka et al. 2010; Wen et al. 2016) because of the simplified fabrication



**Fig. 7** The SEM images of (a) the top view of the fabricated SOI wafer and (b) the suspended “H”-type doubly clamped beam; (c) the photograph of wafer-level anodic bonding between the Pyrex 7740 glass wafer with through-glass vias and the patterned SOI wafer; (d) the SEM image of the bonding interface of a sensor unit after wafer dicing (Luo et al. 2014a, c)

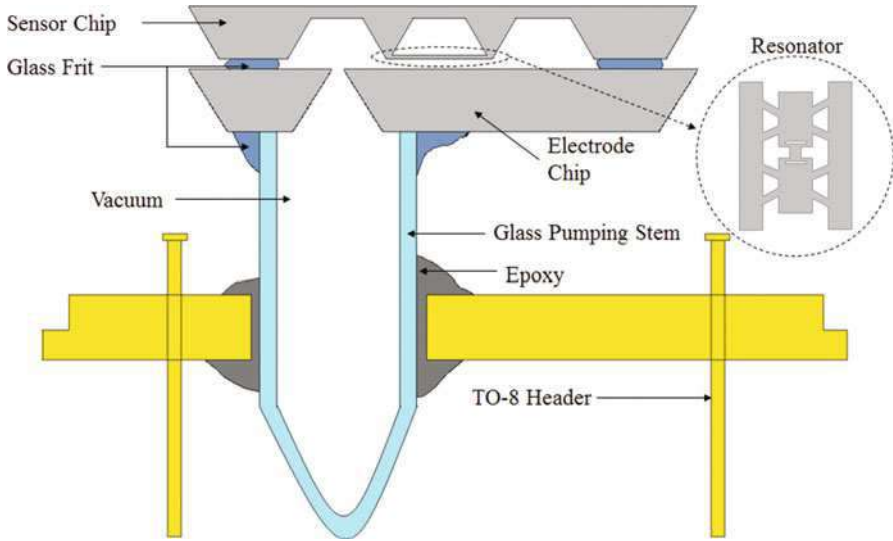
processes and easy integration with circuits. In case of electrostatic excitation, the driving load is simply the attractive force between the two plates of a capacitor, where one electrode plate is formed by (a part of) the resonator and the other by a (stationary) surface located at a close distance from the resonator. Capacitance detection is based on the fact that an AC current will flow through a DC-biased capacitor if the distance between the capacitor plates, and therefore also the capacitance, fluctuates.

## Design

Generally, the resonators are driven to vibrate into the in-plane or out-of-plane oscillations due to the different sensor configurations, which are further introduced in detail as follows.

### Out-of-Plane Vibration

The first silicon resonant pressure sensor was reported by Greenwood J C and Satchell D W in 1988, which was later commercialized by DRUCK Co. As shown in Fig. 8, the butterfly-like resonators anchored on a membrane were formed by

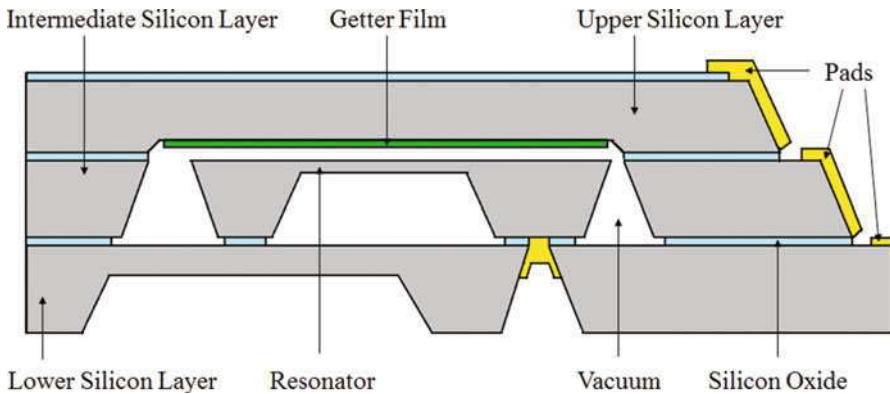


**Fig. 8** The diagram of the microchip by Greenwood J C

doping and selective wet etching (Borenstein and Gerrish 1997; Steinsland et al. 1995; Charavel et al. 2003) in one silicon wafer. The sensing element was then bonded to an electrode chip by glass frit, driving the resonator into vibration. The capacitance change between the resonator and the sensing electrode on the chip was used to detect the resonator's vibration. A chip-level vacuum package was achieved by glass pumping stem, completing the sensor function.

Honeywell designed a differential pressure sensor (Zook et al. 1991) using electrostatic drive and piezoresistive sense. The microsensor was based on the polysilicon processes (Guckel et al. 1990). The microbeam was fabricated by deposited scarified layer and polysilicon layer, which was released by etching the scarified layer. The vacuum was achieved through the deposited polysilicon shell using the reactive sealing technology (Guckel 1991). The driving electrodes on the microbeam and the sensing resistors were both formed by the ion implantation. The beam was driven into oscillation by applying sinusoidal voltage on the drive line which alternately attracted to the biased top electrode on the shell. The vibration was sensed by the resistance changes of the implanted resistor on the microbeam. When the pressure difference deflected the membrane, the stress would change the beam's frequency which was detected by the oscillation circuit.

The commercial P90 Series pressure sensor (Mandle et al. 1995; Longoni et al. 2006) designed by the Thales Group also worked in the out-of-plane vibration, as shown in Fig. 9. It was composed of three layers of silicon. The upper one was a cap layer which also integrated as the detection layer. The resonator was formed on the intermediate silicon layer by wet etching, leaving a thin film for connection. The lower layer functioned as the pressure sensing membrane by patterning and wet etching. The intermediate layer and lower layer were bonded together with silicon



**Fig. 9** P90 Series resonant pressure sensor

oxide for isolation. The bonded wafer needed some more etching to further finalize the resonator. After that, the upper cap wafer functioned as the detection layer, and the cap layer was fusion bonded with the microchip to maintain the vacuum (Petersen et al. 1988; Klaassen et al. 1996).

### In-Plane Mode

The resonators with an in-plane vibration mode were typically based on the high aspect ratio dry etching processes (Marty et al. 2005; Abdolvand and Ayazi 2008). The in-plane vibration usually benefits the high capacitive value, much more complicated and balanced structures, and high Q-factors.

Welham et al. 1999 improved Greenwood's resonant pressure sensor using fusion bonding and trench etching technologies. The sensing elements were comprised of a "double-shuttle" resonator, suspending on two pedestals which were the integral parts of the membrane. Two comb capacitors were formed at the ends of the resonator, driving the resonator into a lateral motion with a "push-pull" mode. The piezoresistors connected as an on-chip Wheatstone bridge functioned as the vibration detectors. The lateral working mode, which was insensitive to the silicon thickness, helped separate the vibration frequency in a wide pressure range.

The chip was also based on a three-wafer-stack configuration. The substrate wafer was patterned and wet etched to form both the membrane and the pedestals which were the anchors of the resonators. The wafer was fusion bonded with a second wafer with silicon oxide for isolation. After thinning and doping, the resonator was patterned and formed by trench etching using DRIE (Bhardwaj and Ashraf 1995). The third cap wafer leaving a hole was bonded to the chip wafer with glass frit to encapsulate the resonator. Finally, glass tube bonded to the chip stack sealed the vacuum on chip level. Furthermore, GE improved the vacuum packaging by silicon fusion bonding (SFB) in wafer level and released the resonant pressure transducers.

Different from above work, D. Chen and J. Wang's team developed a kind of resonant pressure sensors with an in-plane mode, employing silicon-glass anodic bonding for vacuum packaging (Xie et al. 2013, 2015a, b; Jiao et al. 2013).

## Design

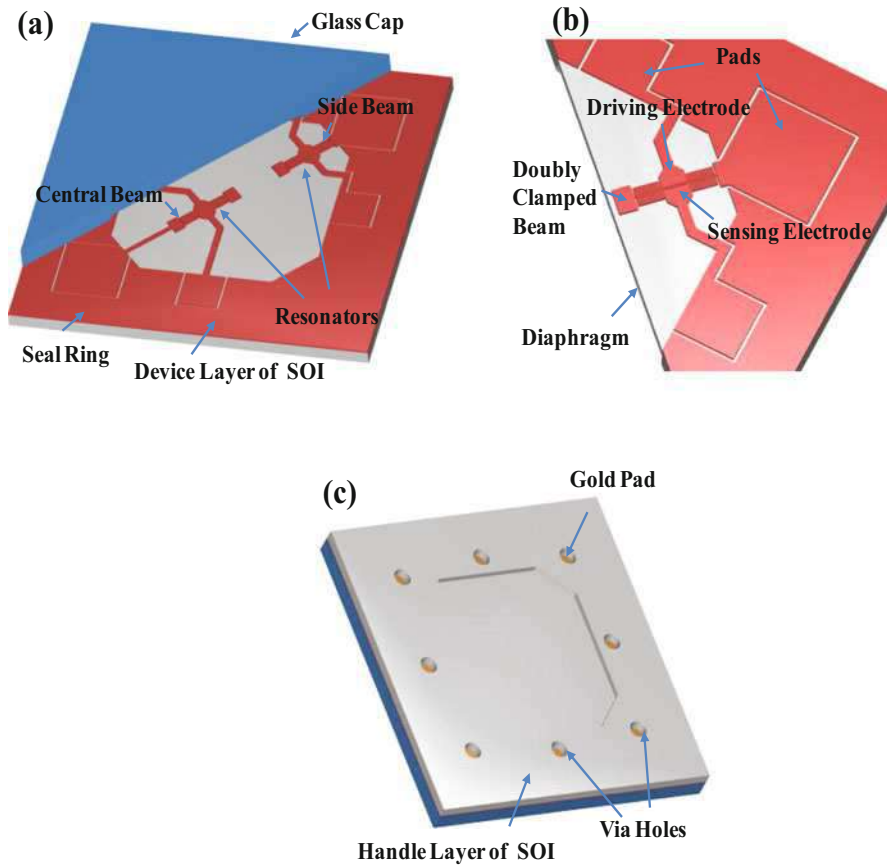
The proposed microsensor is based on a 4-in. SOI wafer for the convenience of the movable resonators (Xie et al. 2015b). The triple layers of SOI are defined as different functions. The handle layer functioning as a supporting layer is etched to form a pressure-sensitive diaphragm. Two resonators on the device layer are clamped on the diaphragm along the diagonal direction through the buried oxide layer. Each resonator is made of a doubly clamped resonant beam, a driving electrode, and a sensing electrode, which form two capacitors with 2  $\mu\text{m}$  gaps of parallel plates. The beam is driven into lateral vibration by the electrostatic force between the plates of the driving capacitor. In order to reduce the squeeze film damping and achieve a high Q-factor of the resonator, the resonators are vacuum packaged in a micro-chamber with a Pyrex 7740 glass cap based on anodic bonding. Via holes are etched on the handle layer for the wire interconnections of the resonators. The schematic of the micro sensor is shown in Fig. 10.

Two resonant beams are designed as a differential structure. The beam positioned besides the diaphragm center is named as "middle beam," in which a tensile stress is observed when pressure is applied to the diaphragm, while the other beam named as "side beam" is subjected to a compressive stress in response to applied pressure. The tensile stress leads to an increasing frequency shift of the middle beam, while the compressive stress results in a decreasing frequency shift of the side beam. Thus, when the dual resonators have the same pressure absolute sensitivities, the sensitivity of the sensor is doubled with the frequency difference indicating the pressure under measurement. Besides, the nonlinear error of resonators is effectively reduced using the differential frequency as the output signal. This method both improves the pressure sensitivity and reduces the nonlinear error without thinning the diaphragm, overcoming the contradictions between sensitivity and linearity.

Based on the SOI technology, the vibration modes of resonant beams were separated from membrane by properly choosing the thickness of the resonators. Note that the intrinsic resonant frequency of the in-plane vibration was thickness independent. The first six vibration modes of the microsensor including the beam and the membrane were shown in Fig. 11, where the base lateral mode (mode 1) was the beam's working mode. To avoid mode interference, the working mode of microbeam was separated from out-of-plane modes under the condition that the thickness of the beam was twice more than the width, while the interference between the beam and the membrane over the whole pressure range was eliminated by optimized resonant geometries using ANSYS simulation, as shown modes in Fig. 12.

## Fabrication

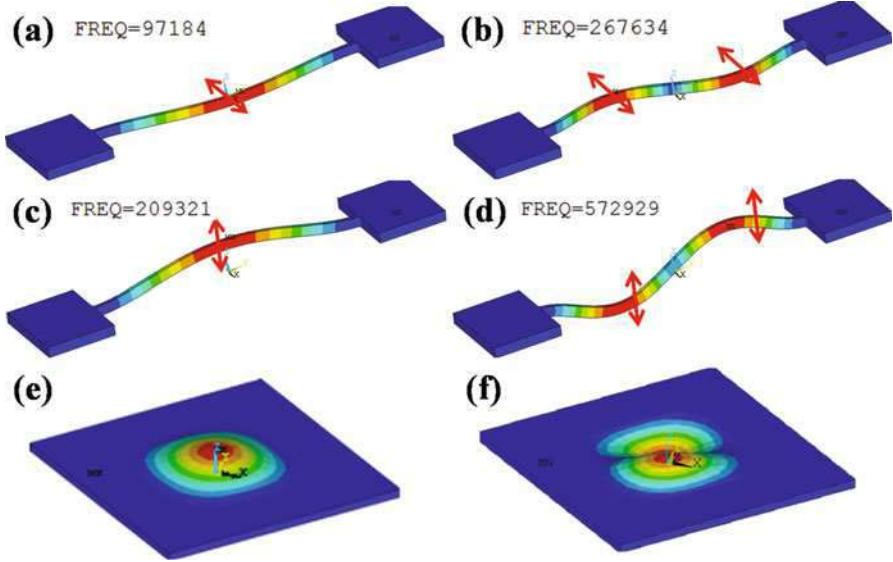
The device fabrication was started from a 4-in. SOI wafer with a 40  $\mu\text{m}$  (100) device layer, a 2  $\mu\text{m}$   $\text{SiO}_2$  sacrificed layer, and a 300  $\mu\text{m}$  substrate layer. The resonators were



**Fig. 10** Resonant pressure sensor based on SOI-glass anodic bonding (Xie et al. 2015b)

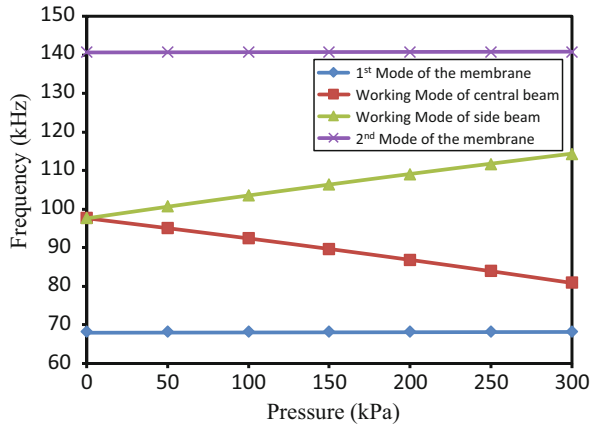
fabricated on the device layer using deep reactive ion etching (see Fig. 13a). Then, membranes and via holes were fabricated on the handle layer, where silicon oxides were firstly deposited and patterned. As the next step, photoresist was patterned on top of the silicon oxide layer, which functions as the etch mask to form via holes in the handle layer using deep reactive ion etching. After that, the patterned photoresist was removed, and the patterned silicon oxide layer was used as the second mask to form the pressure-sensitive membrane (see Fig. 13b). Then, the resonators were released in vapor HF where the oxide layer was removed (see Fig. 13c). Finally, vacuum packaging was realized by bonding the patterned SOI wafer with a glass cap (see Fig. 13d) using a bonding voltage of 800 Volt in an anodic bonding machine (SBe6, SUSS, German).

During the fabrication, a new process of metallization and wire interconnection was developed using an electrochemical etching approach (Xie et al. 2015a) enabling the patterning of the Au film on highly topographic surfaces without



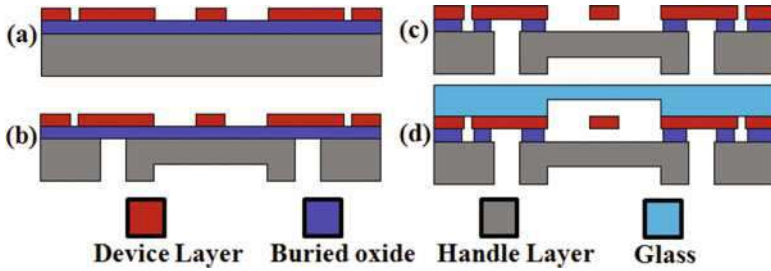
**Fig. 11** Six lowest vibration modes in the microsensor. (a) The fundamental in-plane mode (the working mode) of the microbeam; (b) the second in-plane harmonic mode of the microbeam; (c) and (d) the fundamental and second harmonic out-of-plane mode of the microbeam; (e) and (f) the first and second vibration mode of the membrane

**Fig. 12** The vibration frequencies changes in required pressure range without interference

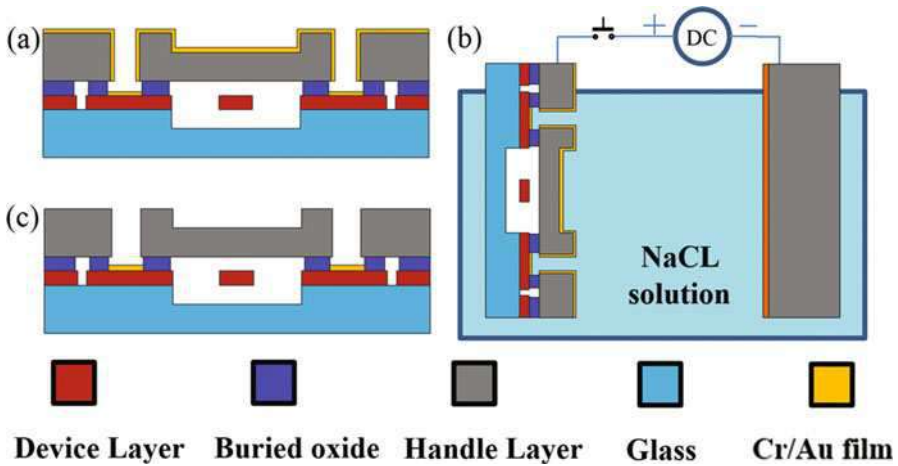


masks (see Fig. 14). More specifically, a Cr/Au film was sputtered on the handle layer as well as the exposed device layer of the patterned SOI wafer (see Fig. 14a). Then the Cr/Au film on the handle layer was connected to the anode of a DC source, and a second silicon substrate deposited with Pt was connected to the cathode. Both electrodes were immersed in a NaCl solution, enabling the selective removal of Au in the chloride ( $Cl^-$ ) solution at anode (see Fig. 14b). Since the Au





**Fig. 13** Device fabrication flow charts: (a) dry etching on the device layer to form resonators. (b) Fabrication of pressure-sensitive membranes and via holes on the handle layer based on deep reactive ion etching. (c) Removal of the oxide layer in vapor HF to release resonators. (d) Vacuum packaging between the fabricated silicon wafer and the glass cap wafer using anodic bonding (Xie et al. 2015b)

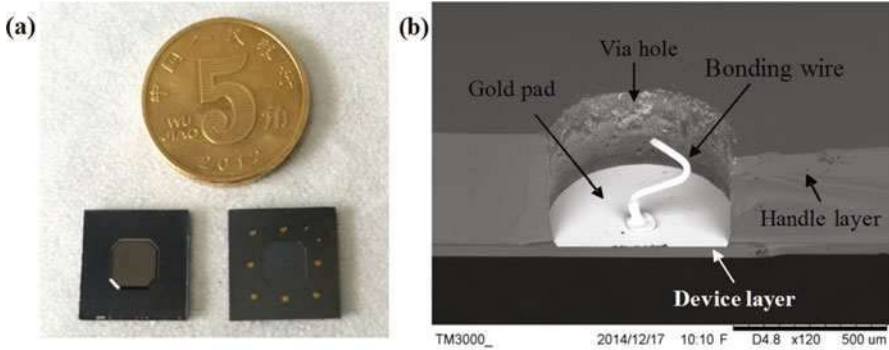


**Fig. 14** The schematic of selective pad patterning in via holes based on the electrochemical dissolution of gold (Xie et al. 2015a)

film on the device layer was not connected with the anode, the gold film on the device layer was left intact, which can be further used for wire interconnections (see Fig. 14c).

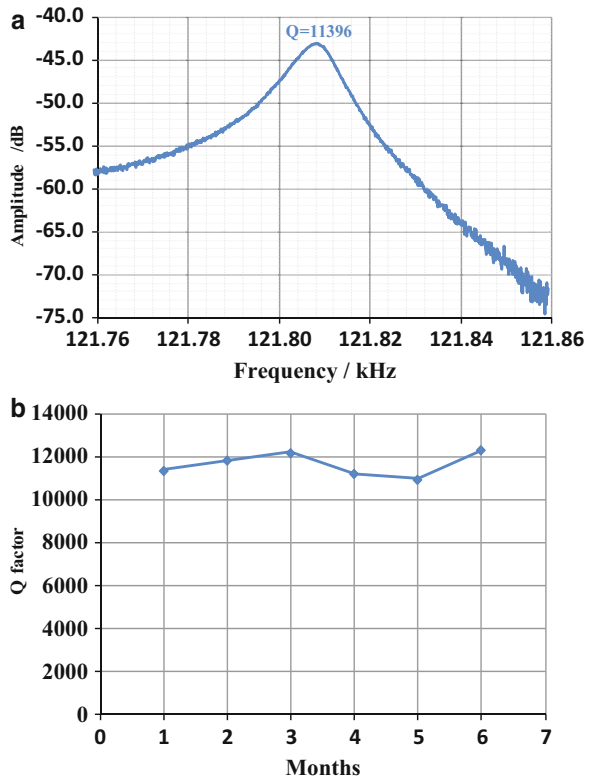
The fabricated device was shown in Fig. 15a, and the SEM picture of a via hole after wire bonding was shown in Fig. 15b. Experimental results confirmed that the gold portion on the handle layer was thoroughly removed with the gold portion only left on the device layer for wire bonding. It is confirmed that the proposed method enables the selective metallization on the device layer and the successful formation of electrical interconnections with the surrounding world.





**Fig. 15** The picture of the microsensor. (a) Sensor chip; (b) via hole with gold pad fabricated with electrochemical selective etching technology (Xie et al. 2015a)

**Fig. 16** Test results of vacuum reliability based on SOI-glass anodic bonding (Xie et al. 2015a, b)



The microsensor was characterized using the quality factor as an indicator of degree of the vacuum packaging. Figure 16 showed the quantified quality factor of 11,396, which was shown to be higher than 10,000 within 6 months of continuous working, validating the reliability of the SOI-glass wafer-level vacuum packaging.

## Self-Temperature Compensation

Temperature disturbance is a key concern in pressure microsensors. More specifically, in resonant pressure microsensors, the key parameters such as thermal stresses and Young's modulus vary with temperature, and thus the temperature variations can cause stress changes in resonators, which lead to frequency drift in response to temperature disturbances (Li et al. 2015; Zhu et al. 2016). Thus, temperature is a critical parameter in resonant pressure microsensors, and various approaches have been proposed to reduce the side effect of temperature variations as temperature compensation. Approaches of temperature compensation can be classified into two types, which are hardware and software compensations, respectively.

**Hardware compensations** are usually realized by including additional temperature-sensitive components within the sensor module or adding temperature compensating elements in the detection circuit.

Relying on the internal temperature-compensated structures or special materials, the issue of the temperature drift can be reduced, including micromachined Si-SiO<sub>2</sub> composite resonators (Melamud et al. 2009); differential structures (Chen et al. 2005), and heavy doping (Pensala et al. 2011).

Melamud et al. (2009) proposed a temperature-insensitive composite micro-mechanical resonator. According to the opposing temperature coefficients of Young's modulus of the Si and SiO<sub>2</sub>, composite resonators by Si-SiO<sub>2</sub> with zero linear temperature coefficient of frequency were fabricated. Experimental results show that resonators have a quadratic temperature coefficient of frequency of approximately  $-20 \text{ ppb}/^\circ\text{C}$  and a tunable turnover temperature in the range of  $-55 \text{ }^\circ\text{C}$  to  $125 \text{ }^\circ\text{C}$ .

The advantage of the hardware compensation is that the compensation principle is straightforward and compensated timeliness is strong, which, however, suffers from the issues of complicated sensor structures and low compensated precisions, which may require further temperature compensation.

**Software compensations.** To improve the accuracy of compensation, software compensations are proposed where microprocessors were used to process raw data obtained from additional temperature sensors to address the issue the temperature drift using appropriate compensation algorithms (Ji et al. 2000; Sun et al. 2004; Liang et al. 2007). However, these compensation approaches rely on external temperature sensors for monitoring the surrounding temperature changes. Due to the uneven temperature distribution and the delay of thermal conductions, these methods suffer from limited accuracy (Yang et al. 2013).

**Self-temperature compensation.** The concept of self-temperature compensation is based on the temperature characteristics of the sensors themselves, where additional temperature sensing elements are not requested to perform temperature compensations (Wang et al. 2013). Wang et al. (1997) proposed a self-temperature compensation method which utilized the different temperature characteristics between the fundamental frequency and the third harmonic frequency of the same resonator. Although this method used a single resonator to achieve temperature self-compensation, the resonator cannot function in a closed-loop manner, which results in low accuracy and long response time.

With the differential structure design of the micromachined silicon resonant pressure sensor, D. Chen and J. Wang's team developed a self-temperature compensation method to improve the performance of the sensor (Luo et al. 2015; Li et al. 2015; Xie et al. 2015a; Zhu et al. 2016).

## The Self-Temperature Compensation Algorithm

Based on the elasticity theory, the natural frequency of the resonant beam in the first-order vibration mode is

$$f_0 = f_1(0) = \frac{4.73^2}{2\pi l^2} \sqrt{\frac{EI}{A\rho}} = 1.028 \sqrt{\frac{E}{\rho}} \cdot \frac{b}{l^2} \quad (18)$$

$$f_1(\sigma) = f_1(0) \sqrt{1 + \frac{\sigma}{\sigma_c}} \quad (19)$$

where  $E$  is Young's modulus of the material,  $I$  is moment of inertia,  $A$  is the cross-sectional area of the beam,  $\rho$  is the material density,  $b$  is the width of the beam, and  $l$  is the length of the beam (Li et al. 2015). Among these parameters,  $f_1$  is related to the stress state of the beam (tensile or compressive), dominated by pressure ( $P$ ) under measurement, while  $E$ ,  $\rho$ ,  $b$ ,  $l$  are all related to temperature ( $T$ ). Thus, the resonant frequencies of the two beams can be expressed as a binary function of pressure and temperature, which is

$$\begin{cases} f_1 = F_1(p, T) \\ f_2 = F_2(p, T) \end{cases} \quad (20)$$

where  $f_1$  is the resonant frequency of the central beam,  $f_2$  is the resonant frequency of the side beam,  $p$  is the pressure load, and  $T$  is the temperature. According to Eq. 2, pressure and temperature can be expressed as a binary function of  $f_1$  and  $f_2$  based on mathematical translation as follows:

$$\begin{cases} p = G_1(f_1, f_2) \\ T = G_2(f_1, f_2) \end{cases} \quad (21)$$

where function  $G_1$  is the self-temperature compensation function.

According to Eq. 21, the pressure  $p$  is determined by the two resonant frequencies,  $f_1$  and  $f_2$ , which form a 3D surface at the coordinates  $(f_1, f_2)$ . This function can be expressed by a polynomial surface fitting, which is

$$\begin{aligned} p = & a_0 + a_1 f_1 + a_2 f_2 + a_3 f_1^2 + a_4 f_1 f_2 + a_5 f_2^2 + a_6 f_1^3 + a_7 f_1^2 f_2 \\ & + a_8 f_1 f_2^2 + a_9 f_2^3 + \varepsilon \end{aligned} \quad (22)$$

where  $a_0, \dots, a_9$  are coefficients and  $\varepsilon$  is higher-order infinitesimal.

## Calibration for Self-Compensation of the Resonant Pressure Sensor

In order to determine the coefficients in Eq. 22, a calibration process for the sensor is required. The  $m$  temperature points were selected for temperature calibration in the full temperature range of  $-40\text{ }^{\circ}\text{C} \sim 70\text{ }^{\circ}\text{C}$ , and  $n$  pressure points were selected for pressure calibration in the full pressure scale of  $50\text{ kPa} \sim 110\text{ kPa}$  (atmospheric pressure in this work). In total,  $m \times n$  arrays of temperature and pressure couples were used.

After calibration,  $a_0, \dots, a_9$  were obtained with calibration data based on the least square method. The mean square error  $R$  can be expressed as

$$R = \frac{1}{m \times n} \sum_{k=1}^{m \times n} [p_k - p(f_{1k}, f_{2k})]^2 \quad (23)$$

where  $p_k, f_{1k}$ , and  $f_{2k}$  are calibration data,  $k = 1, 2, 3, \dots, m \times n$ . In order to minimize  $R$ , partial derivatives of  $R$  with respect to  $a_0, \dots, a_9$  should be equal to zero, which is

$$\frac{\partial R}{\partial a_i} = 0, i = 0, 1, \dots, 9 \quad (24)$$

Thus, the coefficients,  $a_0, \dots, a_9$ , can be solved with Eq. 24. By substituting these coefficients into Eq. 21, the self-temperature compensation is realized.

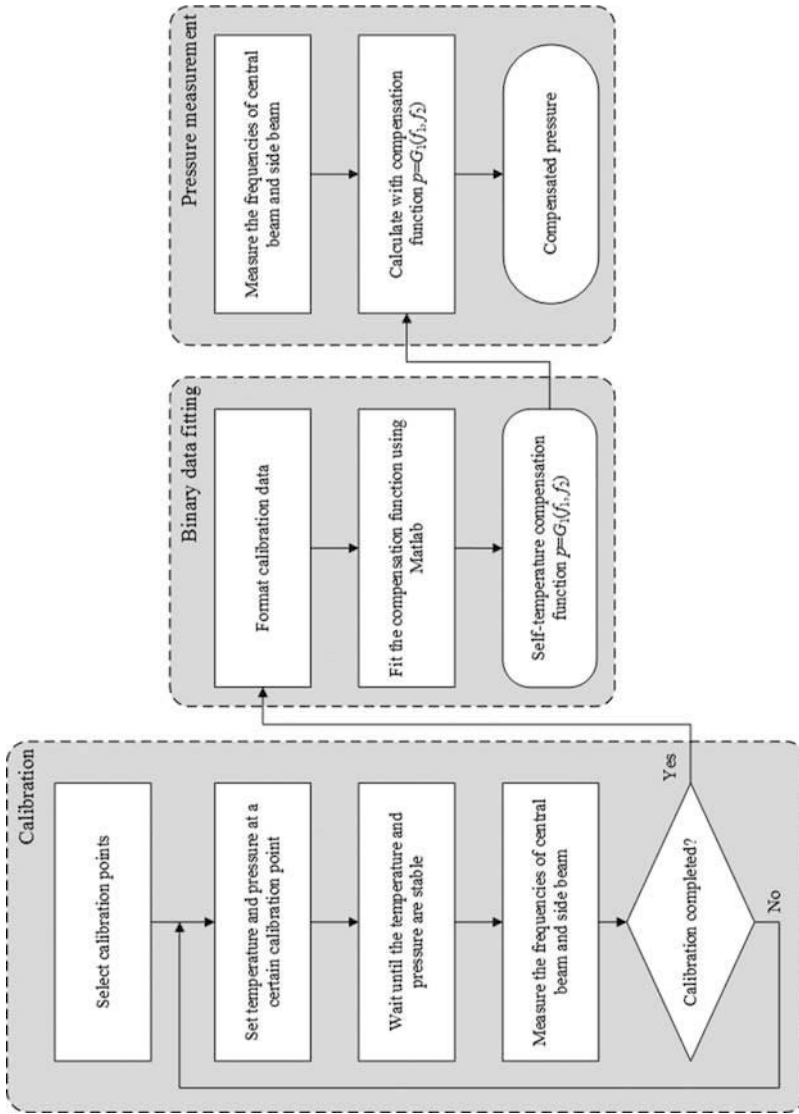
The completed self-temperature compensation processes are summarized in Fig. 17. As the first step, certain calibration points in the full temperature and pressure scales were selected with two resonant frequencies quantified. Secondly, polynomial surface fitting with the calibration data was conducted to obtain the self-temperature compensation function. Third, this compensation function was programmed into the post-processing unit of the sensor, and thus the pressure micro-sensor can function in the self-temperature compensation mode to calculate the real-time pressure.

In the calibration processes, the surrounding temperature was controlled by a temperature and humidity chamber, and the pressure was controlled by a pressure controller/calibrator. In addition, the resonant frequency was measured by a digital multimeter. The 3D surface plot of pressure in the coordinates  $(f_1, f_2)$  is shown in Fig. 18.

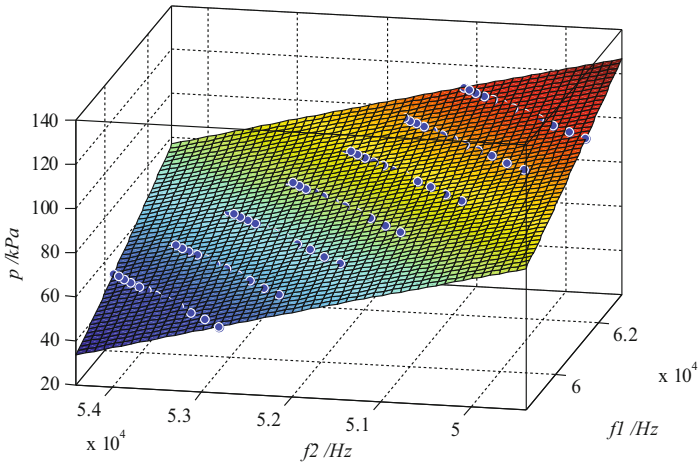
## Compensation Results and Analysis

The difference (error) between the experimental results of  $p_k$  and the compensated pressure value  $p(f_{1k}, f_{2k})$  is shown in Fig. 19. The compensation error was less than  $\pm 0.01\%$  F.S. ( $50 \sim 110\text{ kPa}$ ) in temperature range of  $-40\text{ }^{\circ}\text{C} \sim 70\text{ }^{\circ}\text{C}$ .

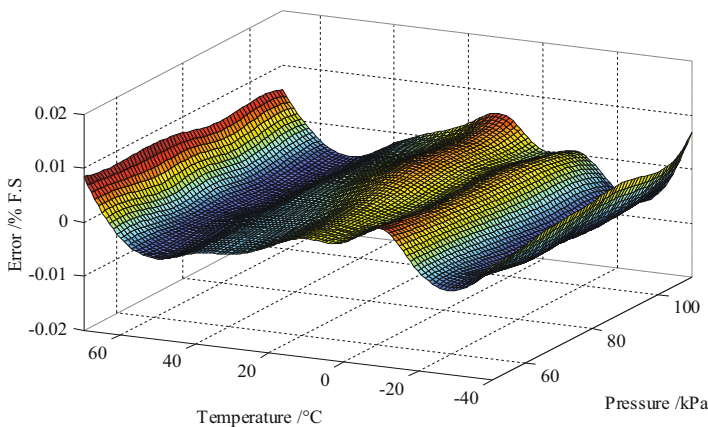
The errors of non-compensation and the mentioned self-temperature calibration are compared in Fig. 20, where the temperature was changed from  $-40\text{ }^{\circ}\text{C}$  to  $70\text{ }^{\circ}\text{C}$  at  $50\text{ kPa}$  (Fig. 20a),  $80\text{ kPa}$  (Fig. 20b), and  $110\text{ kPa}$  (Fig. 20c), respectively. Without



**Fig. 17** The self-temperature compensation flow chart includes three main steps: calibration, binary data fitting, and pressure measurement. In the calibration step, the frequencies of the central beam and side beam were measured at each temperature and pressure point (Li et al. 2015)



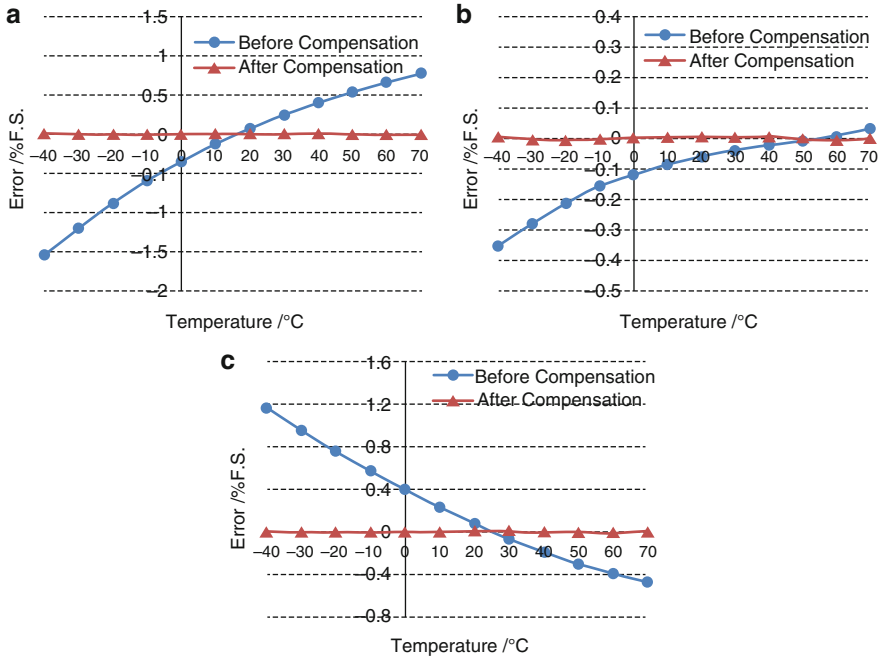
**Fig. 18** Surface plot of calculated pressure values obtained after cubic polynomial surface fitting as a function of two resonant frequencies  $f_1$  and  $f_2$ ; the blue dots indicate the calibration points (Li et al. 2015)



**Fig. 19** Surface plot of the pressure compensation error as a function of temperature and pressure, which is less than  $\pm 0.01\%$  F.S. (50 kPa to 110 kPa) ranged from  $(-40\text{ }^\circ\text{C}$  to  $70\text{ }^\circ\text{C})$  (Li et al. 2015)

temperature compensation, the maximal error was higher than 1.5%, even the pressure microsensors worked in the differential mode. After applying the self-temperature compensation algorithm, the errors were significantly decreased to less than 0.01%.

In addition, the sensors were tested at additional pressure and temperature values besides the calibration points to further validate the functionality of the self-temperature compensation method.  $-15\text{ }^\circ\text{C}$ ,  $10\text{ }^\circ\text{C}$ ,  $25\text{ }^\circ\text{C}$ , and  $50\text{ }^\circ\text{C}$  were chosen as temperature test points, and 50 kPa, 65 kPa, 80 kPa, 95 kPa, and 110 kPa were picked as pressure test points. The test results are shown in Table 1 where the quantified maximal error was  $-3.7\text{ Pa}$ , which was lower than  $\pm 0.01\%$  F.S.



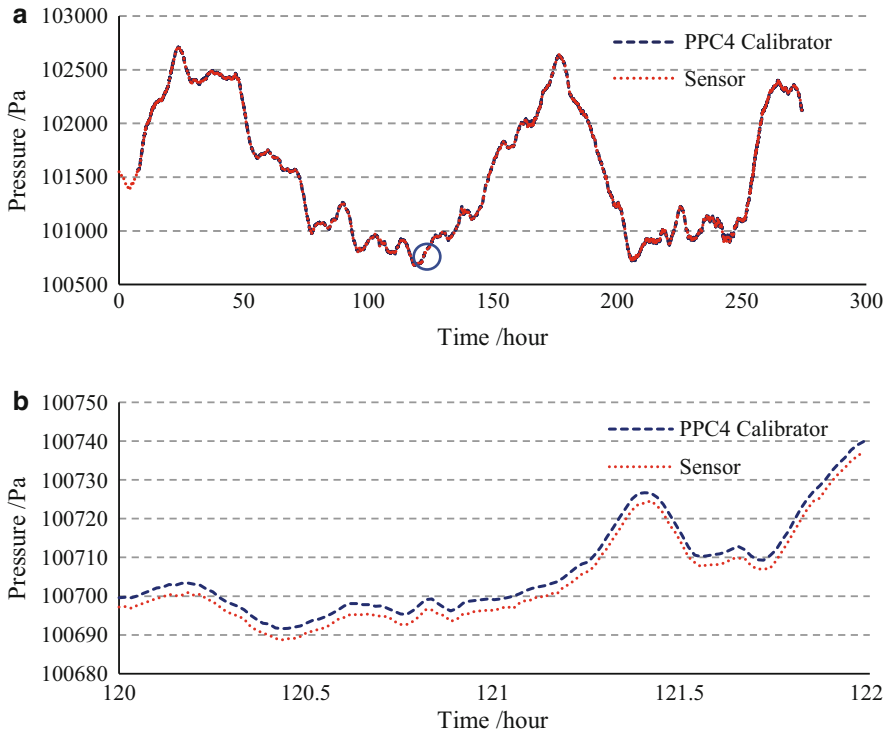
**Fig. 20** Comparison of the errors before and after compensation at the pressure of (a) 50 kPa, (b) 80 kPa, and (c) 110 kPa. Before compensation, the max error was higher than 1.5%, while after compensation, the errors were lower than 0.01% (Li et al. 2015)

**Table 1** The quantified errors at different pressure and temperature points. The maximal error measured was  $-3.7$  Pa, which was lower than  $\pm 0.01\%$  F.S. (Li et al. 2015)

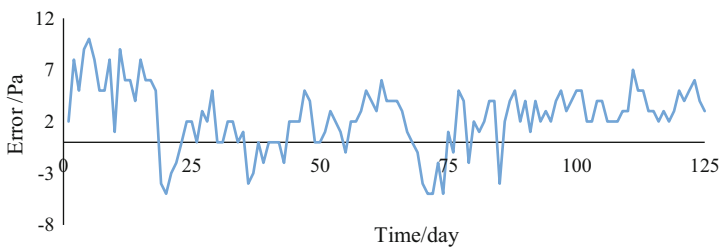
Error/Pa		Testing pressure/kPa				
		50	65	80	95	110
Testing temperature /°C	-15	1.7	1	-0.7	-0.8	-3.7
	10	0	-0.5	-0.2	-1.2	-1.9
	25	0.9	1.1	1.1	-0.2	-0.4
	50	1.3	1.2	0.8	-0.4	-0.7

Furthermore, the sensors were tested under actual atmospheric conditions for short-term and long-term stability quantifications (a total of 270 h and 125 days, respectively).

Figure 21a shows that the device developed in this study demonstrated a comparable result (after temperature compensation) as the pressure controller/calibrator during 270 h testing in the actual atmosphere. Figure 21b shows the detailed results in the time duration from 120 to 122 h, where no deviation larger than 8 Pa between the pressure controller/calibrator and the pressure microsensors demonstrated in this study was observed.



**Fig. 21** The actual atmospheric test within a duration of 270 h. (a) The device introduced in this chapter demonstrated comparable results with the pressure controller/calibrator over 270 h in atmosphere; (b) the detailed experiment results where the deviations lower than 8 Pa were observed (Li et al. 2015)



**Fig. 22** Atmospheric pressure reading deviation between the demonstrated microsensor and the calibrator in room during 4 months (Li et al. 2015)

Figure 22 records the reading deviation between the microsensor and the calibrator during 125 days in room, where the biggest deviation is no larger than 10 Pa between the pressure calibrator and the pressure sensor.



## Conclusions

The intrinsic resonant frequency, geometrical parameters, and sensitivities of the resonant pressure microsensors are discussed in this chapter, which provides references for future design of microsensors.

The micromachined silicon resonant pressure sensor based on electromagnetic driving is presented, which uses SOI-MEMS processes, through-glass vias (TGV) electrical interconnections, silicon-to-glass anodic bonding, and getter material preparation to achieve high-Q performance. The FEM simulations for the sensor design are also discussed.

At the same time, the micromachined silicon resonant pressure sensor based on electrostatic driving is also presented with its design and fabrication discussed in detail. Device fabrication also includes key steps of SOI-MEMS processes, silicon-to-glass anodic bonding, and getter material preparation. Differently, to achieve wafer-level vacuum packaging and wire interconnections, through-silicon-vias (TSV) and maskless gold etching based on the electrochemical principle were adopted.

To improve the performance of the micromachined silicon resonant pressure sensor when working under wide temperature ranges, the self-temperature compensation method is included. Based on the micromachined silicon resonant pressure sensors with two “H”-type doubly clamped resonant beams, the algorithm enabling the calculation of temperature and pressure based on two resonant frequencies is proposed and optimized. After compensation, the errors were lower than 0.01% F.S. (50–110 kPa) in the temperature range of  $-40\text{ }^{\circ}\text{C}$  to  $70\text{ }^{\circ}\text{C}$ , and therefore the micromachined silicon resonant pressure sensor can work continuously in the self-temperature compensation mode with longtime stability.

**Acknowledgments** We thank the National Basic Research Program of China (973 Program, Grant No. 2014CB744600), the National Natural Science Foundation of China (Grant Nos. 61431019, 61372054 and 61671430), the Beijing Municipal Natural Science Foundation (Grant No. 4152056), the Beijing Municipal Science and Technology Commission (Grant No. D11110100160000), and the Beijing NOVA Program of Science and Technology for financial support.

---

## References

- Abdolvand R, Ayazi F (2008) An advanced reactive ion etching process for very high aspect-ratio sub-micron wide trenches in silicon. *Sens Actuators A Phys* 144(1):109–116
- Beeby SP, Ensell G, Baker BR et al (2000) Micromachined silicon resonant stain gauges fabricated using SOI wafer technology. *IEEE J MEMS* 9:104–111
- Bhardwaj JK, Ashraf H (1995) Advanced silicon etching using high-density plasmas. In: *Proceedings of SPIE – the international society for optical engineering, Orlando*
- Borenstein JT, Gerrish ND (1997) Etch selectivity of novel epitaxial layers for bulk micro-machining. *MRS Online Proc Lib* 546
- Cabuz C, Shoji S, Fukatsu K et al (1994) Fabrication and packaging of a resonant infrared sensor integrated in silicon. *Sens Actuators A Phys* 43:92–99

- Charavel R, Lanconte J, Raskin JP (2003) Advantages of p++ polysilicon etch stop layer versus p++ silicon. *Proc SPIE Smart Sensors, Actuators and MEMS* 5116:699–709
- Chen DY (2002) Research on micromachined resonant beam pressure sensors. Dissertation, University of Chinese Academy of Sciences
- Chen DY, Cui DF, Xia SH et al (2005) Design and modeling of a silicon nitride beam resonant pressure sensor for temperature compensation. In: *Proceedings of the 2nd IEEE international conference on information acquisition*, Hong Kong
- Chen DY, Wu ZW, Shi XJ et al (2009) Design and modeling of an electromagnetically excited silicon nitride beam resonant pressure sensor. In: *Proceedings of the 4th IEEE international conference of nano/micro engineered and molecular systems*, Shenzhen
- Chen DY, Li YX, Liu M et al (2010a) Design and experiment of a laterally driven micromachined resonant pressure sensor for barometers. *Procedia Eng* 5:1490–1493
- Chen DY, Li YX, Liu M et al (2010b) A novel laterally driven micromachined resonant pressure sensor. In: *Proceedings of the 9th IEEE international conference on sensors*, Hawaii
- Choa SH (2005) Reliability of MEMS packaging: vacuum maintenance and packaging induced stress. *Microsyst Technol* 11:1187–1196
- Damjanovic D (1998) Materials for high-temperature piezoelectric transducers. *Curr Opin Solid State Mater Sci* 3:469–473
- Eaton WP, Smith JH (1997) Micromachined pressure sensors: review and recent developments. *Smart Mater Struct* 7(6):530–539
- Esashi M (2008) Wafer level packaging of MEMS. *J Micromech Microeng* 18:01–13
- Greenwood JC, Satchell DW (1988) Miniature silicon resonant pressure sensor. *IEE Proc D Control Theory Appl* 135(5):369–372
- Greenwood JC, Wray T (1993) High accuracy pressure measurement with a silicon resonant sensor. *Sens Actuators A Phys* 37:82–85
- Guckel H (1991) Surface micromachined pressure transducers. *Sens Actuators A Phys* 28(28):133–146
- Guckel H, Sniegowski JJ, Christenson TR et al (1990) The application of fine-grained, tensile polysilicon to mechanically resonant transducers. *Sens Actuators A Phys* 21(1–3):346–351
- Harada K, Ikeda K, Kuwayama H et al (1996) Various applications of resonant pressure sensor chip based on 3D micromachining. *Sens Actuators A Phys* 73:261–266
- Hirano T, Furuhashi T, Gabriel KJ et al (1992) Design, fabrication, and operation of submicron gap comb-drive microactuators. *IEEE J MEMS* 1:52–59
- Ikeda K, Kuwayama H, Kobayashi T et al (1990) Silicon pressure sensor integrates resonant strain gauge on diaphragm. *Sens Actuators A Phys* 21:146–150
- Ikeda K, Kuwayama H, Kobayashi T et al (1991) Three-dimensional micromachining of silicon pressure sensor integrating resonant strain gauge on diaphragm. *Sens Actuators A Phys* 23:1007–1010
- Ji HX, Yu P, Liang XJ (2000) Research of sensor nonlinear compensation. *Machinery* 10:1127–1130
- Jiao HL, Xie B, Wang JB et al (2013) Electrostatically driven and capacitively detected differential lateral resonant pressure microsensor. *Micro Nano Lett* 8(10):650–653
- Kasten K, Amelung J, Mokwa W (2000) CMOS-compatible capacitive high-temperature pressure sensors. *Sens Actuators A Phys* 85:147–152
- Kinnell PK, Craddock R (2009) Advances in silicon resonant pressure transducers. *Procedia Chem* 1:104–107
- Klaassen EH, Petersen K, Noworolski JM et al (1996) Silicon fusion bonding and deep reactive ion etching: a new technology for microstructures. *Sens Actuators A Phys* 52(1–3):132–139
- Li YX, Chen DY, Wang JB (2011) Stress isolation used in MEMS resonant pressure sensor package. *Procedia Eng* 25:455–458
- Li YX, Chen DY, Wang JB (2012) Low temperature wafer level adhesive bonding using BCB for resonant pressure sensor. *Key Eng Mater* 503:75–80
- Li YX, Chen DY, Wang JB (2013) A new stress isolation method in the packaging of resonant pressure micro sensors. *Sens Lett* 11(2):264–269

- Li YN, Wang JB, Luo ZY et al (2015) A resonant pressure microsensors capable of self-temperature compensation. *Sensors* 15(5):10048–10058
- Liang WF, Wang XD, Liang PE (2007) Pressure sensor temperature compensation based on least squares support vector machine. *Chin J Sci Instrum* 12:2235–2238
- Longoni G, Conte A, Moraja M et al (2006) Stable and reliable Q-factor in resonant MEMS with getter film. In: Proceedings of the 44th IEEE annual international reliability physics symposium, San Jose
- Luo ZY, Chen DY, Wang JB (2013) A differential resonant barometric pressure sensor using SOI-MEMS technology. In: Proceedings of the 12th IEEE international conference on sensors, Baltimore
- Luo ZY, Chen DY, Wang JB et al (2014a) A high-Q resonant pressure microsensors with through-glass electrical interconnections based on wafer-level MEMS vacuum packaging. *Sensors* 14:24244–24257
- Luo ZY, Chen DY, Wang JB (2014b) A SOI-MEMS based resonant barometric pressure sensor with differential output. *Key Eng Mater* 609(610):1033–1039
- Luo ZY, Chen DY, Wang JB (2014c) Resonant pressure sensor with Through-glass electrical interconnect based on SOI wafer technology. In: Proceedings of IEEE NEMS, Hawaii
- Luo ZY, Li YN, Xie B et al (2015) A self-temperature compensating barometer with dual doubly-clamped resonators. In: Proceedings of the 18th international conference on solid-state sensors, actuators & microsystems, Anchorage
- Mandle J, Lefort O, Migeon A (1995) A new micromachined silicon high-accuracy pressure sensor. *Sens Actuators A Phys* 46:129–132
- Marty F, Rousseau L, Saadany B et al (2005) Advanced etching of silicon based on deep reactive ion etching for silicon high aspect ratio microstructures and three-dimensional micro- and nanostructures. *Microelectron J* 36(7):673–677
- Melamud R, Chandorkar SA, Kim B et al (2009) Temperature-insensitive composite micro-mechanical resonators. *J MEMS* 18(6):1409–1419
- Mitchell J, Lahiji GR, Najafi K (2005) Encapsulation of vacuum sensors in a wafer level package using a gold-silicon eutectic. In: Proceedings of the 13th international conference on solid-state sensors, actuators & microsystems, Seoul
- Moraja M, Amiotti M (2003) Getters films at wafer level for wafer to wafer bonded MEMS. In: Proceedings of symposium on design, test, integration and packaging of MEMS/MOEMS, Mandelieu-Lanapoule
- O'Mahony C, Hill M, Olszewski Z et al (2009) Wafer-level thin-film encapsulation for MEMS. *Microelectron* 86:1311–1313
- Otmani R, Benmoussa N, Benyoucef B (2011) The thermal drift characteristics of piezoresistive pressure sensor. *Phys Procedia* 21:47–52
- Pensala T, Jaakkola A et al (2011) Temperature compensation of silicon MEMS resonators by heavy doping. In: Proceedings of the IEEE international ultrasonics symposium, Orlando
- Petersen K, Barth P, Poydock J et al (1988) Silicon fusion bonding for pressure sensor. In: Proceedings of the solid-state sensor and actuator workshop, Hilton Head Island
- Pramanik C, Islam T, Saha H (2006) Temperature compensation of piezoresistive micro-machined porous silicon pressure sensor by Ann. *Microelectron Reliab* 46:343–351
- Santo ZM, Mozek M, Macek S et al (2010) An LTCC-based capacitive pressure sensor with a digital output. *Inf MIDEM* 40:74–81
- Schulz M, Sauerwald J, Richter D et al (2009) Electromechanical properties and defect chemistry of high-temperature piezoelectric materials. *Ionics* 15:157–161
- Steinsland E, Nese M, Hanneborg A et al (1995) Boron etch-stop in TMAH solutions. *Sens Actuators A Phys* 1(97):190–193
- Sun YC, Chen ZY, Wang J (2004) Normalizing the polynomial-match for a non-linear function in sensors and solid electronics. *J Electron Devices* 1:1–7
- Wang JS, Dong YG, Feng GP et al (1997) Temperature characteristics of quartz resonant force sensors and self-temperature-compensation. *J Tsinghua Univ (Sci Technol)* 8:12–14

- Wang JB, Chen DY, Xia SH et al (2008a) A novel method to eliminate the co-channel interference of micromachined diffused silicon resonant pressure sensor. In: Proceedings of the 7th IEEE international conference on sensors, Lecce
- Wang JB, Shi XJ, Liu L et al (2008b) A novel resonant pressure sensor with boron diffused silicon resonator. In: Proceedings of SPIE international conference on optical instrument and technology-MEMS/NEMS technology and applications, Beijing
- Wang JB, Chen DY, Liu L et al (2009) A micromachined resonant pressure sensor with DETFs resonator and differential structure. In: Proceedings of the 8th IEEE international conference on sensors, Chirstchurch
- Wang JC, Xia XY et al (2013) Piezoresistive pressure sensor with dual-unit configuration for on-chip self-compensation and suppression of temperature drift. In: Proceedings of the international conference on solid-state sensors, actuators & microsystems, Barcelona
- Welham CJ, Greenwood J, Bertioli MM (1999) A high accuracy resonant pressure sensor by fusion bonding and trench etching. *Sens Actuators A Phys* 76(1–3):298–304
- Wen H, Daruwalla A, Mirjalili R et al (2016) HARPSS-fabricated nano-gap comb-drive for efficient linear actuation of high frequency BAW resonators. In: Proceedings of IEEE MEMS, Shanghai
- Xie B, Jiao HL, Wang JB et al (2013) An electrostatically-driven and capacitively-sensed differential lateral resonant pressure microsensor. In: Proceedings of the 8th annual IEEE international conference on nano/micro engineered and molecular systems, Suzhou
- Xie B, Xing Y, Wang Y et al (2015a) A lateral differential resonant pressure microsensor based on SOI-glass wafer-level vacuum packaging. *Sensors* 15(9):24257–24268
- Xie B, Xing Y, Wang Y et al (2015b) Vacuum-packaged resonant pressure sensor with dual resonators for high sensitivity and linearity. *Process Eng* 120:194–199
- Yang L, Su Y, Qiu AP et al (2013) Self-temperature compensation for high quality factor micro-machined gyroscope. *Opt Precis Eng* 11:2870–2876
- Yoneoka S, Salvia JC, Bahl G et al (2010) Active electrostatic compensation of micromechanical resonators under random vibrations. *IEEE J MEMS* 19(5):1270–1272
- Zhu L, Xie B, Xing YH et al (2016) A resonant pressure sensor capable of temperature compensation with least squares support vector machine. *Procedia Eng* 168:1731–1734
- Zook JD, Burns DW, Guckel H et al (1991) Resonant microbeam strain transducers. In: Proceedings of the 6th international conference on solid-state sensors and actuators, San Francisco



# A Micromachined Vibratory Gyroscope

Qiancheng Zhao, Chunhua He, Jian Cui, and Guizhen Yan

## Contents

Overview .....	422
Structure and Fabrication .....	423
Dynamics Theory .....	425
Drive and Detection Principles .....	430
Drive and Detection with Slide-Film Combs .....	430
Drive and Detection with Squeeze-Film Combs .....	432
Transimpedance Amplifier Detection Method .....	433
Closed-Loop Design for the Drive Mode .....	435
AGC Self-Oscillation Scheme .....	435
AGC + PLL Scheme .....	438
Experimental Tests .....	440
Mode-Matching Control .....	442
Temperature Characteristics Analysis .....	442
Automatic Mode-Matching Control System .....	443
Real-Time Mode-Matching Control System .....	445
Experimental Tests .....	447
Closed-Loop Design for the Sense Mode .....	449
Force Rebalance Control System .....	449
Analysis in the Frequency Domain of Angular Rate .....	451
Analysis in the Frequency Domain of Force .....	452
Experimental Tests .....	453
Temperature Compensation .....	456
Temperature Compensation Principle .....	456
Experimental Test .....	457
Conclusion .....	457
References .....	458

Q. Zhao (✉) · C. He · J. Cui · G. Yan

National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Beijing, China

e-mail: [zqc@pku.edu.cn](mailto:zqc@pku.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,  
[https://doi.org/10.1007/978-981-10-5945-2\\_12](https://doi.org/10.1007/978-981-10-5945-2_12)

421

---

**Abstract**

MEMS vibratory gyroscopes are widely used in various fields such as consumer electronics, automotive industry, and navigation systems owing to their merits of low power, low cost, and small volume. A MEMS gyroscope is one of the most important inertial devices, which is researched all over the world.

This chapter will introduce a MEMS vibratory gyroscope, aiming to make more people understand what it is and how it works. Firstly, structure design and fabrication, as well as its dynamics theory, are illustrated. Secondly, closed-loop control for the drive mode and mode-matching control are demonstrated. Finally, closed-loop control for the sense mode, followed by temperature compensation method, is presented in detail. In addition, theoretical deduction, simulation, and experimental tests are combined tightly to make the proposed techniques understand easily.

---

**Keywords**

MEMS gyroscope · Closed-loop control · Mode-matching · Temperature compensation

---

**Overview**

MEMS gyroscopes, which are sensitive to external angular rate, are paid more and more attention from 1980s. According to the performance indice, MEMS gyroscopes can be classified into rate grade, tactical grade, and inertial grade (Yazdi et al. 1998). Rate-grade gyroscopes can be applied to consumer electronics, mobile phones, digital camera, game machines, and wireless mouses. Tactical-grade gyroscopes can be applied to industry control, intelligent automobile, trains, and steamship. Inertial-grade gyroscopes can be applied to satellite, navigation, aerospace, and missile guidance.

There are many kinds of gyroscopes, such as ring laser gyro (RLG), dynamically tuned gyro (DTG), interferometric fiber-optic gyro (IFOG), and MEMS gyro (Yole Developpement 2015, 2016). The performances of RLG or IFOG are far superior to those of MEMS gyro, so there is much work to do for MEMS gyros.

According to incomplete statistics, the universities, companies, and institutes, that research on MEMS gyroscopes, are as follows: Georgia institute of Technology, Standford University, University of Michigan, University of California Berkeley, Irvine, Los Angeles, Middle East Technical University, University of Freiburg, University of Southampton, Seoul National University, Ghent University, Tsinghua University, Peking University, Southeast University, Shanghai Jiaotong University, Zhejiang University, Bosch, ST, InvenSense, NXP, ADI, TI, and so on. Nowadays, the researches on MEMS gyroscopes are focused on the following techniques:

- (1) New material, fabrication, and process (Franssila 2004; Liu et al. 2006; Cho et al. 2013; Shao et al. 2015).
- (2) ASIC monolithic integration (Rombach et al. 2015, 2016).

- (3) High vacuum packaging (Askari et al. 2006; Senkal et al. 2014, 2015; Torunbalci et al. 2015), such as wafer-level packaging.
- (4) New structure and operation principle (Askari et al. 2016; Ahn et al. 2015; Vafanejad and Kim 2015), such as ring, disk, and hemisphere gyroscopes.
- (5) Mode-matching control (Sonmezoglu et al. 2014; He et al. 2013a; Flader et al. 2016), noise suppressing (Lajimi et al. 2017; Sharma 2007; Tatar 2010; Cao et al. 2016; Zihajehzadeh et al. 2015), and coupling signal suppressing (Lee et al. 2008; Acar and Shkel 2005; Cui et al. 2011a).
- (6) Closed-loop control for drive mode (Cui et al. 2011b; Yang and Li 2015; Liu et al. 2014) and force rebalance control for the sense mode (Chen et al. 2016; Rombacha et al. 2015; He et al. 2013b; Tatar et al. 2012).
- (7) Self-calibration (Trusov et al. 2013; Zhang et al. 2015; Aktakka and Najafi 2016) and temperature compensation (Prihodko et al. 2013; Casinovi et al. 2016; Fontanella et al. 2016).
- (8) Reliability test, failure analysis, and reliability design (Li et al. 2014, 2015; Makkonen et al. 2012; Pasquale and Soma 2011; Pierron et al. 2006; Chen et al. 2011).

Aiming at some techniques mentioned above, this chapter will introduce a micromachined vibratory gyroscope.

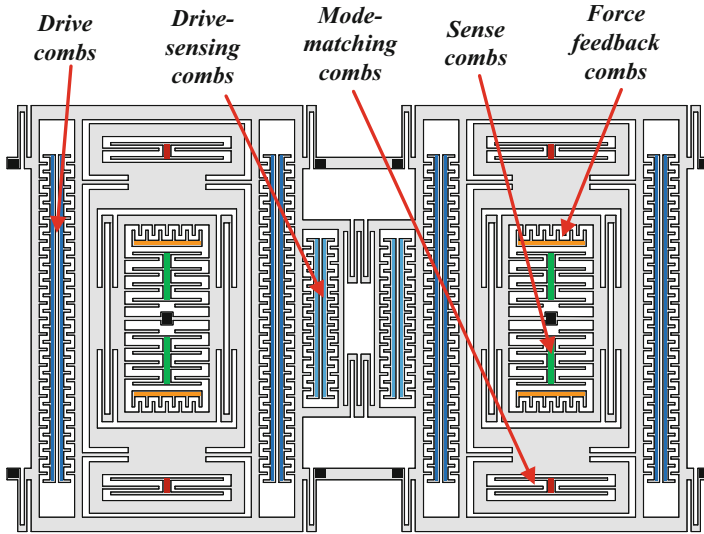
---

## Structure and Fabrication

The structure design aims to make a gyroscope achieve 1 deg/h bias instability as well as bandwidth larger than 90 Hz. Mode-matching loop will be adopted to reduce the white noise, and closed-loop control for the sense mode will be applied to extend the bandwidth and improve the nonlinearity.

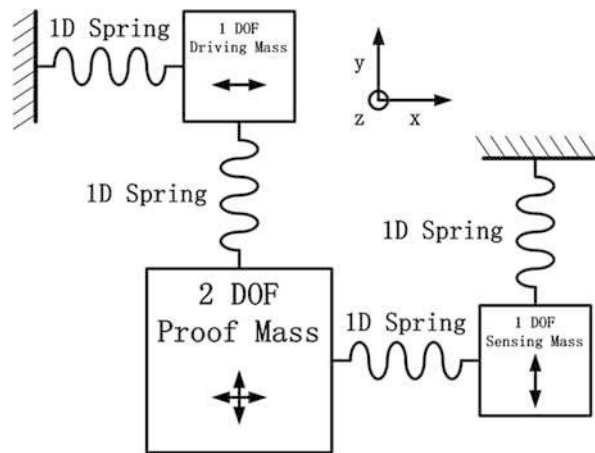
Figure 1 shows the simplified schematic of a Z-axis doubly decoupled tuning fork MEMS vibratory gyroscope, and the chip size is about 5 mm × 3 mm. There are two groups of combs in the drive mode, namely, slide-film drive-sensing combs and slide-film drive combs, which are designed for the closed-loop control of the drive mode. On the other hand, there are three groups of combs in the sense mode, namely, slide-film force feedback combs, squeeze-film sense combs, and squeeze-film mode-matching combs. The sense combs and the force feedback combs are designed for the closed-loop control of the sense mode.

Since slide-film combs can produce a good linear force without inducing the electrostatic stiffness effect, they are suitable for generating the drive force and feedback force. If squeeze-film combs are chosen as the force feedback combs, the resonant frequency of the sense mode would be changed as external Coriolis force, which will increase the complexity of the mode-matching control. Due to a large drive displacement in the drive mode, slide-film combs are applied as the drive-sensing combs to advance the linearity of sense capacitances. Besides, it is beneficial to enhance the quality factor. On the other hand, squeeze-film combs are utilized as the sense combs to improve the detection sensitivity. Meanwhile, thanks to the force



**Fig. 1** The schematic of a Z-axis doubly decoupled MEMS tuning fork gyroscope

**Fig. 2** The diagram of a doubly decoupled model



rebalance control, the nonlinearity issue induced by squeeze-film combs can be solved. Based on the electrostatic stiffness effect, mode-matching combs can be used to tune the resonant frequency of the drive mode ( $f_d$ ) to equal that of the sense mode ( $f_s$ ).

The diagram of a doubly decoupled model is illustrated in Fig. 2. It figures out that the driving mass and sensing mass have only one degree of freedom (DOF); however, the proof mass has two DOFs. That is to say, the driving mass and proof mass will vibrate along the x-axis when the gyroscope is working, and then the sensing mass and proof mass will vibrate along y-axis when angular rate appears in



z-axis. The Coriolis force is proportional to the displacement and the differential comb capacitance, so transimpedance amplifier can be adopted to fulfill the angular rate detection.

SOG (silicon on glass) process using a silicon/glass wafer bonding and DRIE (deep reactive ion etching) is applied to the fabrication. The initial materials are 4 inch highly doped silicon wafers with the resistivity of 0.01–0.03  $\Omega$  cm and Pyrex 7740 glass whose thermal expansion coefficient matches that of silicon. First, the anchor areas and mechanical moving parts are defined in the silicon substrate by a 20- $\mu$ m deep DRIE etch using silicon dioxide as a mask. A 200-nm Ti/Pt/Au layer is sputtered on the glass wafer and patterned by a lift-off process to form interconnects. Afterward, the two wafers are anodically bonded together, and then the silicon substrate is thinned to about 100  $\mu$ m by KOH etching for 80- $\mu$ m thick device structures. Finally, the gyroscope structure is released by ICP DRIE. More details of the fabrication process can be found in the literature (Li et al. 2000).

## Dynamics Theory

Neglecting the influences of coupling signals, the dynamics equations of the two modes of a MEMS gyroscope can be derived as

$$m_d \ddot{x}(t) + c_d \dot{x}(t) + k_d x(t) = F_0 \sin(\omega t) \quad (1)$$

$$m_s \ddot{y}(t) + c_s \dot{y}(t) + k_s y(t) = -2m_p \Omega v \quad (2)$$

where  $F_d = F_0 \sin(\omega t)$  and  $F_c = -2m_p \Omega v$  are the drive force and Coriolis force, respectively.  $v$  is the drive velocity, and  $\Omega$  is angular rate.  $m_d$ ,  $m_s$ , and  $m_p$  are the drive mass, sense mass, and proof mass, respectively.  $c_d$ ,  $k_d$ , and  $x$  are the damping coefficient, stiffness coefficient, and the displacement of the drive mode, respectively.  $c_s$ ,  $k_s$ , and  $y$  are the damping coefficient, stiffness coefficient, and the displacement of the sense mode, respectively.

Assume that  $\omega_d = \sqrt{\frac{k_d}{m_d}}$  and  $\omega_s = \sqrt{\frac{k_s}{m_s}}$  are the intrinsic radian frequencies of the two modes,  $\xi_d = \frac{c_d}{2m_d \omega_d}$  and  $\xi_s = \frac{c_s}{2m_s \omega_s}$  are the damping ratios of the two modes, and  $Q_d = \frac{1}{2\xi_d}$  and  $Q_s = \frac{1}{2\xi_s}$  are the quality factors of the two modes. Thus, the transfer functions of the two equations above can be expressed as

$$G_d(s) = \frac{x(s)}{F_d(s)} = \frac{1}{m_d s^2 + c_d s + k_d} = \frac{1/m_d}{s^2 + 2\xi_d \omega_d s + \omega_d^2} \quad (3)$$

$$G_s(s) = \frac{y(s)}{F_c(s)} = \frac{1}{m_s s^2 + c_s s + k_s} = \frac{1/m_s}{s^2 + 2\xi_s \omega_s s + \omega_s^2} \quad (4)$$

(1) Since the initial displacement and velocity of the drive mode are both zero, the initial conditions can be written as

$$x(t) = 0, \quad x'(t) = 0 \quad (5)$$

The displacement of the drive mode can be obtained, after solving (1) and (5), as follows

$$x(t) = \frac{\omega A_d}{\omega_d \sqrt{1 - \xi_d^2}} e^{-\frac{\omega_d}{2Q_d} t} \sin \left( \sqrt{1 - \xi_d^2} \omega_d t + \theta_d \right) + A_d \sin (\omega t - \varphi_d)$$

$$A_d = \frac{F_0}{m_d \omega_d^2 \sqrt{\left(1 - \frac{\omega^2}{\omega_d^2}\right)^2 + \left(\frac{\omega}{Q_d \omega_d}\right)^2}} = \frac{F_0/m_d}{\sqrt{\omega^4 + 2(2\xi_d^2 - 1)\omega_d^2 \omega^2 + \omega_d^4}} \quad (6)$$

$$\varphi_d = \operatorname{tg}^{-1} \frac{\omega_d \omega}{Q_d(\omega_d^2 - \omega^2)}, \quad \theta_d = \operatorname{tg}^{-1} \frac{\sqrt{1 - \xi_d^2}}{\xi_d - Q_d(\omega_d^2 - \omega^2)/\omega_d^2}$$

It figures out that the drive displacement includes two parts. One is transient response, while another one is stable response. The former decays exponentially, and the smaller is  $Q_d$ , the larger is the decay speed. The amplitude and phase of the latter are determined by  $\omega_d$  and  $Q_d$ . Closed-loop control of the drive mode aims to obtain stable amplitude of the drive velocity and make the resonant frequency equal the intrinsic frequency. Therefore, it is necessary to detailedly analyze the drive velocity. After taking a derivative of  $x$  with respect to  $t$ , the drive velocity can be yielded as

$$v(t) = \frac{-\omega A_d}{\sqrt{1 - \xi_d^2}} e^{-\frac{\omega_d}{2Q_d} t} \sin \left( \sqrt{1 - \xi_d^2} \omega_d t + \theta_d - \alpha_d \right) + \omega A_d \cos (\omega t - \varphi_d)$$

$$\alpha_d = \operatorname{tg}^{-1} \frac{\sqrt{1 - \xi_d^2}}{\xi_d}$$

Thus, the stable amplitude  $v_d$  can be obtained as

$$v_d = \omega A_d = \frac{F_0/m_d}{\sqrt{[\omega^4 + 2(2\xi_d^2 - 1)\omega_d^2 \omega^2 + \omega_d^4]/\omega^2}} \quad (8)$$

If the denominator reaches minimum,  $v_d$  will be maximum. Assume that

$$h(\omega) = \frac{\omega^4 + 2(2\xi_d^2 - 1)\omega_d^2 \omega^2 + \omega_d^4}{\omega^2} \quad (9)$$

Taking a derivative of  $h(\omega)$  with respect to  $\omega$  yields

$$h'(\omega) = \frac{4\omega^4 - 4\omega_d^4}{\omega^3} \tag{10}$$

Set  $h'(\omega)$  to be zero, we get

$$2(\omega^4 - \omega_d^4)/\omega^3 = 0 \tag{11}$$

Because  $\omega$  is larger than 0, thus

$$\omega = \omega_d \tag{12}$$

$$h(\omega_d) = 4\xi_d^2\omega_d^2 \tag{13}$$

Taking the second derivative of  $h(\omega)$  with respect to  $\omega$  yields

$$h''(\omega) = \frac{2\omega^4 + 6\omega_d^4}{\omega^4} \tag{14}$$

Because

$$h''(\omega_d) = 8 > 0 \tag{15}$$

$h(\omega_d)$  is the minimum when  $\omega > 0$ . Now, the maximum of  $v_d$  is yielded as

$$A_v = v_{dmax} = \frac{F_0}{2\xi_d m_d \omega_d} = \frac{F_0 Q_d}{m_d \omega_d} \tag{16}$$

If  $\omega$  equals  $\omega_d$ ,  $\varphi_d$  is  $90^\circ$ , and the stable drive velocity is

$$v(t) = A_v \sin(\omega_d t) = \frac{F_0 Q_d}{m_d \omega_d} \sin(\omega_d t) \tag{17}$$

(2) As for the sense mode, the angular rate can be defined as

$$-\Omega(t) = -\Omega_R \cos(\omega_R t) \tag{18}$$

Then the Coriolis force can be derived as

$$F_c = -2m_p \Omega v = -2m_p \Omega_R \cos(\omega_R t) A_v \sin(\omega_d t) = -F_{cm} \cos(\omega_R t) \sin(\omega_d t) \tag{19}$$

where  $\omega_R$  is the radian frequency of angular rate and  $F_{cm} = 2m_p A_v \Omega_R$  is the amplitude of  $F_c$ .

(a) If  $\omega_R = 0$ , then  $-\Omega = -\Omega_R$ . Similarly,  $y$  can be deducted as

$$y(t) = \frac{-\omega_d A_s}{\omega_s \sqrt{1 - \xi_s^2}} e^{-\frac{\omega_s}{2Q_s} t} \sin \left( \sqrt{1 - \xi_s^2} \omega_s t + \theta_s \right) - A_s \sin (\omega_d t - \varphi_s)$$

$$A_s = \frac{F_{cm}}{m_s \omega_s^2 \sqrt{\left(1 - \frac{\omega_d^2}{\omega_s^2}\right)^2 + \left(\frac{\omega_d}{Q_s \omega_s}\right)^2}} = \frac{F_{cm}/m_s}{\sqrt{\omega_d^4 + 2(2\xi_s^2 - 1)\omega_s^2 \omega_d^2 + \omega_s^4}} \quad (20)$$

$$\varphi_s = tg^{-1} \frac{\omega_s \omega_d}{Q_s (\omega_s^2 - \omega_d^2)}, \theta_s = tg^{-1} \frac{\sqrt{1 - \xi_s^2}}{\xi_s - Q_s (\omega_s^2 - \omega_d^2)/\omega_s^2}$$

It figures out that the sense displacement includes two parts. One is transient response, while another one is stable response. The former decays exponentially, and the smaller is  $Q_s$ , the larger is the decay speed. The amplitude and phase of the latter are determined by  $\omega_s$  and  $Q_s$ . The larger is  $A_s$ , the bigger is the mechanical sensitivity. Hence, the denominator of  $A_s$  should be minimum. Assume that

$$f(\omega_d) = \omega_d^4 + 2(2\xi_s^2 - 1)\omega_s^2 \omega_d^2 + \omega_s^4 \quad (21)$$

Taking a derivative of  $f(\omega_d)$  with respect to  $\omega_d$  yields

$$f'(\omega_d) = 4\omega_d^3 + 4(2\xi_s^2 - 1)\omega_s^2 \omega_d \quad (22)$$

Setting  $f'(\omega_d)$  to be zero, we get

$$4\omega_d [\omega_d^2 - (1 - 2\xi_s^2)\omega_s^2] = 0 \quad (23)$$

Because  $\omega_d$  is larger than 0, thus

$$\omega_d = \sqrt{1 - 2\xi_s^2} \omega_s = \sqrt{1 - \frac{1}{2Q_s^2}} \omega_s \quad (24)$$

Hence

$$f \left( \sqrt{1 - 2\xi_s^2} \omega_s \right) = 4\xi_s^2 \omega_s^4 (1 - \xi_s^2) \quad (25)$$

Taking the second derivative of  $f(\omega_d)$  with respect to  $\omega_d$  yields

$$f''(\omega_d) = 12\omega_d^2 + 4(2\xi_s^2 - 1)\omega_s^2 \quad (26)$$

Since  $\xi_s$  is often smaller than 0.707,

$$f'' \left( \sqrt{1 - 2\xi_s^2} \omega_s \right) = 8(1 - 2\xi_s^2)\omega_s^2 > 0 \quad (27)$$

Therefore,  $f\left(\sqrt{1-2\xi_s^2}\omega_s\right)$  is the minimum, and the maximum of  $A_s$  is as follows

$$A_{s\max} = \frac{F_{cm}Q_s}{m_s\omega_s^2\sqrt{1-\xi_s^2}} = \frac{F_{cm}Q_s}{m_s\omega_s^2\sqrt{1-1/(4Q_s^2)}} \tag{28}$$

If  $Q_s$  is large enough, we can yield

$$A_{imax} \approx \frac{F_{cm}Q_s}{m_s\omega_s^2} \tag{29}$$

On the other hand, taking a derivative of  $y$  with respect to  $t$ , the sense velocity can be yielded as

$$v_s(t) = \frac{\omega_d A_s}{\sqrt{1-\xi_s^2}} e^{-\frac{\omega_s}{2Q_s}t} \sin\left(\sqrt{1-\xi_s^2}\omega_s t + \theta_s - \alpha_s\right) - \omega_d A_s \cos(\omega_d t - \varphi_s) \tag{30}$$

$$\alpha_s = tg^{-1} \frac{\sqrt{1-\xi_s^2}}{\xi_s}$$

Based on the stable response of sense velocity, the mechanical sensitivity can be defined as

$$S_m = \frac{-\omega_d A_s}{-\Omega_R} = \frac{2m_p A_v \omega_d}{m_s \omega_s^2 \sqrt{\left(1 - \frac{\omega_d^2}{\omega_s^2}\right)^2 + \left(\frac{\omega_d}{Q_s \omega_s}\right)^2}} \tag{31}$$

Similar to the analysis above, if the following condition is met

$$\omega_s = \omega_d, \tag{32}$$

$S_m$  will be largest. That is, the mechanical sensitivity will be maximum after mode-matching control. Thus  $\varphi_s = 90^\circ$ , and

$$A_{vs} = v_{s\max} = \frac{F_{cm}}{2\xi_s m_s \omega_s} = \frac{F_{cm}Q_s}{m_s \omega_s} \tag{33}$$

$$S_{m\max} = \frac{2F_0 m_p Q_d Q_s}{m_d \omega_d m_s \omega_s} \tag{34}$$

$$v_{s\max}(t) = -A_{vs} \sin(\omega_s t) = \frac{-F_{cm}Q_s}{m_s \omega_s} \sin(\omega_s t) = \frac{-2F_0 m_p \Omega_R Q_d Q_s}{m_d \omega_d m_s \omega_s} \sin(\omega_d t) \tag{35}$$

Hence,  $S_m$  can be enlarged by magnifying  $F_0$ ,  $m_p$ ,  $Q_d$ ,  $Q_s$ , and reducing  $m_d$ ,  $m_s$ ,  $\omega_d$ ,  $\omega_s$ . To sum up, as for the stable response of the sense mode, the resonant frequency is  $\sqrt{1 - 2\xi_s^2}\omega_s$  when displacement detection scheme is adopted. However, the resonant frequency is  $\omega_s$ , rather than  $\sqrt{1 - 2\xi_s^2}\omega_s$ , when velocity detection scheme is adopted. Besides, the intrinsic frequency of the two modes can be acquired accurately with velocity detection scheme, and it is immune to the influence of air damping. Hence, in order to achieve better mode-matching, velocity detection scheme should be applied to both drive mode and sense mode.

- (b) The static characteristics of a MEMS gyroscope are depicted in detail when  $\omega_R$  is equal to 0. As for the dynamic characteristics, which means  $\omega_R > 0$ , it will be analyzed in the following section.

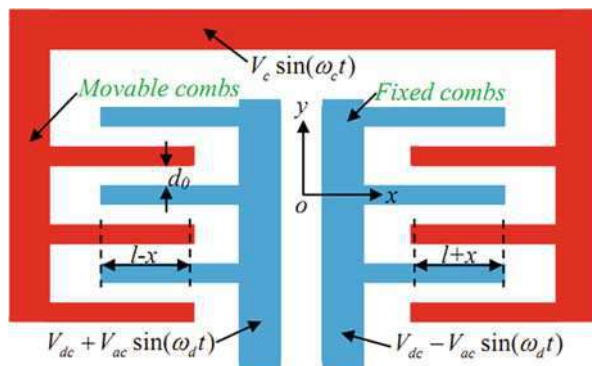
## Drive and Detection Principles

### Drive and Detection with Slide-Film Combs

#### (1) Drive principle

Figure 3 shows the diagram of the slide-film combs. The red parts are fixed and connected to the high-frequency carrier  $V_c \sin(\omega_c t)$ . On the other hand, the blue parts are movable combs and connected to the driving signals  $V_{dc} + V_{ac} \sin(\omega_d t)$  and  $V_{dc} - V_{ac} \sin(\omega_d t)$ .  $V_{dc}$ ,  $V_{ac}$ , and  $V_c$  are the amplitudes of DC, AC, and carrier signals, respectively.  $\omega_c$  often ranges from  $2\pi \times 1$  MHz to  $2\pi \times 10$  MHz, and  $\omega_d$  often ranges from  $2\pi \times 8$  kHz to  $2\pi \times 10$  kHz.  $h$ ,  $l$ ,  $d_0$ ,  $\epsilon$ , and  $N$  are the comb thickness, overlapped length, gap, permittivity, and the number of the comb fingers, respectively.

**Fig. 3** The diagram of the slide-film combs



Thus, the differential voltage between the fixed combs and movable combs are

$$\begin{aligned} V_1 &= V_{dc} + V_{ac} \sin(\omega_d t) - V_c \sin(\omega_c t), V_2 \\ &= V_{dc} - V_{ac} \sin(\omega_d t) - V_c \sin(\omega_c t) \end{aligned} \quad (36)$$

Then we get

$$\begin{aligned} V_1^2 + V_2^2 &= 2V_{dc}^2 + V_{ac}^2 + V_c^2 - [V_{ac}^2 \cos(2\omega_d t) + V_c^2 \cos(2\omega_c t) + 4V_{dc}V_c \sin(\omega_c t)]V_1^2 \\ &- V_2^2 = 4V_{dc}V_{ac} \sin(\omega_d t) - 4V_{ac}V_c \sin(\omega_d t) \sin(\omega_c t) \end{aligned} \quad (37)$$

The potential energies can be expressed as

$$E_1 = \frac{2NV_1^2 \epsilon h(l-x)}{2d_0}, E_2 = \frac{2NV_2^2 \epsilon h(l+x)}{2d_0} \quad (38)$$

Hence, neglecting the high-frequency terms since they will be filtered by the gyroscope, the total driving force is obtained as

$$\begin{aligned} F_{tot} &= F_1 + F_2 = \frac{-\partial E_1}{\partial x} + \frac{-\partial E_2}{\partial x} \\ &= \frac{N\epsilon h(V_1^2 - V_2^2)}{d_0} \underline{\underline{gyro}} \frac{4NV_{dc}V_{ac}\epsilon h \sin(\omega_d t)}{d_0} \end{aligned} \quad (39)$$

Then the electrostatic stiffness, resulted from the derivative of  $F_{tot}$  with respect to  $x$ , can be described as

$$k_x = \frac{-\partial F_{tot}}{\partial x} = 0 \quad (40)$$

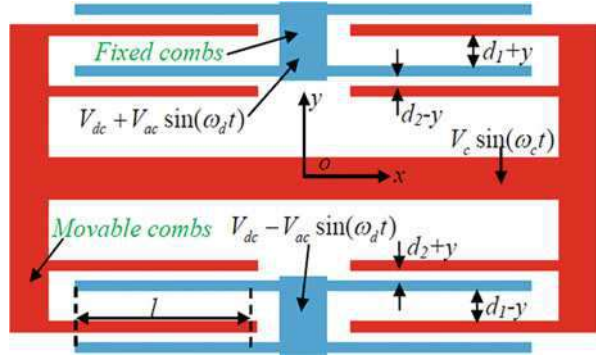
Equation (40) demonstrates that electrostatic stiffness cannot be generated by slide-film combs, so that slide-film combs cannot be used for mode-matching control. However, the driving force possesses good drive linearity since  $F_{tot}$  has nothing to do with  $x$ . Thereby, slide-film combs can be utilized to driving for the two modes.

## (2) Detection principle

Slide-film combs can be also used for displacement or velocity detection. As shown in Fig. 3, removing the drive signal exerted to the fixed combs, they can be utilized to differential capacitance detection. The total differential capacitance  $C_{tot}$  can be derived as

$$C_{tot} = C_{right} - C_{left} = \frac{2N\epsilon h(l+x)}{d_0} - \frac{2N\epsilon h(l-x)}{d_0} = \frac{4N\epsilon h x}{d_0} \quad (41)$$

**Fig. 4** The diagram of the squeeze-film combs



It indicates that  $C_{tot}$  is proportional to the displacement  $x$ , and it possesses good sense linearity. Since the displacement of the drive mode is often large to advance the performances, it requires good sense linearity, and slide-film combs are applicable.

**Drive and Detection with Squeeze-Film Combs**

(1) Drive principle

Figure 4 shows the diagram of the squeeze-film combs.  $d_1$  and  $d_2$  are the initial gaps of a comb unit. Similarly, the potential energies can be expressed as

$$E_1 = N \left[ \frac{\epsilon h l V_1^2}{2(d_1 + y)} + \frac{\epsilon h l V_1^2}{2(d_2 - y)} \right], E_2 = N \left[ \frac{\epsilon h l V_2^2}{2(d_1 - y)} + \frac{\epsilon h l V_2^2}{2(d_2 + y)} \right] \quad (42)$$

Hence, neglecting the high-frequency terms since they will be filtered by the gyroscope, the total driving force is obtained as

$$\begin{aligned} F_{tot} &= F_1 + F_2 = \frac{-\partial E_1}{\partial y} + \frac{-\partial E_2}{\partial y} \\ &= N \left[ \frac{\epsilon h l V_1^2}{2(d_1 + y)^2} - \frac{\epsilon h l V_1^2}{2(d_2 - y)^2} - \frac{\epsilon h l V_2^2}{2(d_1 - y)^2} + \frac{\epsilon h l V_2^2}{2(d_2 + y)^2} \right] \\ &\approx N \epsilon h l \left( \frac{1}{2d_1^2} - \frac{1}{2d_2^2} \right) (V_1^2 - V_2^2) \underline{\underline{gyro}} 2N \epsilon h l V_{dc} V_{ac} \left( \frac{1}{d_1^2} - \frac{1}{d_2^2} \right) \sin(\omega_d t) \quad (43) \end{aligned}$$

Then the electrostatic stiffness, resulted from the derivative of  $F_{tot}$  with respect to  $y$ , can be described as



$$\begin{aligned}
k_y &= \frac{-\partial F_{tot}}{\partial y} = N \left[ \frac{\varepsilon h l V_1^2}{(d_1 + y)^3} + \frac{\varepsilon h l V_1^2}{(d_2 - y)^3} + \frac{\varepsilon h l V_2^2}{(d_1 - y)^3} + \frac{\varepsilon h l V_2^2}{(d_2 + y)^3} \right] \\
&\approx N \left[ \frac{\varepsilon h l (V_1^2 + V_2^2)}{d_1^3} + \frac{\varepsilon h l (V_1^2 + V_2^2)}{d_2^3} \right] = N \varepsilon h l \left( \frac{1}{d_1^3} + \frac{1}{d_2^3} \right) (V_1^2 + V_2^2) = N \varepsilon h l \left( \frac{1}{d_1^3} + \frac{1}{d_2^3} \right) \\
&\quad \{ 2V_{dc}^2 + V_{ac}^2 + V_c^2 - [V_{ac}^2 \cos(2\omega_d t) + V_c^2 \cos(2\omega_c t) + 4V_{dc} V_c \sin(\omega_c t)] \}
\end{aligned} \tag{44}$$

Equation (43) indicates that  $F_{tot}$  has nothing to do with  $y$ , and the linearity will be affected if  $y$  is large. In addition, Eq. (44) demonstrates that electrostatic stiffness can be generated by squeeze-film combs, so that squeeze-film combs can be used for mode-matching control. However, if they are used to driving, the mode-matching control loop will be affected. Therefore, compared with slide-film combs, squeeze-film combs are not suitable to be used to generate driving force.

The frequency-tuned range is closely related to  $V_{dc}$ ,  $V_{ac}$ ,  $V_c$ ,  $d_1$ ,  $d_2$ ,  $N$ ,  $h$ ,  $l$ , and  $\varepsilon$ , as shown in Eq. (44). If only taking  $V_{dc}$  into account, the electrostatic stiffness can be simplified as

$$k_y = 2V_{dc}^2 N \varepsilon h l \left( \frac{1}{d_1^3} + \frac{1}{d_2^3} \right) \tag{45}$$

Hence, in order to simplify the control scheme, we can only adopt DC voltage to achieve mode-matching control.

## (2) Detection principle

Squeeze-film combs can be also used for displacement or velocity detection. As shown in Fig. 4, removing the drive signal exerted to the fixed combs, they can be utilized to differential capacitance detection. The total differential capacitance  $C_{tot}$  can be derived as

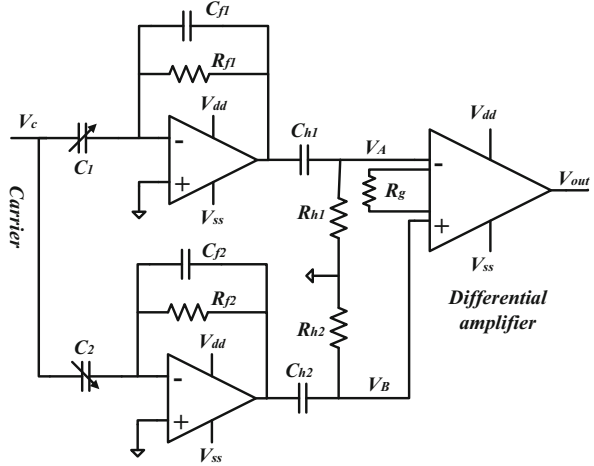
$$\begin{aligned}
C_{tot} &= C_{up} - C_{down} = \left( \frac{N \varepsilon h l}{d_1 + y} + \frac{N \varepsilon h l}{d_2 - y} \right) - \left( \frac{N \varepsilon h l}{d_1 - y} + \frac{N \varepsilon h l}{d_2 + y} \right) \\
&= 2N \varepsilon h l y \left( \frac{1}{d_2^2 - y^2} - \frac{1}{d_1^2 - y^2} \right) \approx 2N \varepsilon h l y \left( \frac{1}{d_2^2} - \frac{1}{d_1^2} \right)
\end{aligned} \tag{46}$$

It indicates that  $C_{tot}$  is proportional to the displacement  $y$ , and it possesses good sense linearity only if  $y$  is small enough. Since the displacement of the sense mode is often small, it requires big sensitivity, and squeeze-film combs are applicable.

## Transimpedance Amplifier Detection Method

The schematic of the readout circuit with transimpedance amplifiers is shown in Fig. 5.  $V_c$  is the carrier signal, which may be a DC or a high-frequency signal.  $V_A$  and

**Fig. 5** The schematic of the readout circuit with transimpedance amplifiers



$V_B$  are the inputs of the differential amplifier.  $C_1$  and  $C_2$  are the differential detection capacitances of a MEMS gyroscope,  $R_{f1}$  and  $R_{f2}$  are the feedback resistors,  $C_{f1}$  and  $C_{f2}$  are the feedback capacitors,  $C_{h1}$  and  $C_{h2}$  are the capacitors of a first-order high pass filter (HPF), and  $R_{h1}$  and  $R_{h2}$  are the resistors of a HPF. According to Fig. 5, we get

$$C_1 = C_{01} + \Delta C_1 \sin(\omega_d t + \alpha) \quad (47)$$

$$C_2 = C_{02} - \Delta C_2 \sin(\omega_d t + \alpha) \quad (48)$$

$$V_A(s) = \frac{-sR_{f1}C_1}{1 + sR_{f1}C_{f1}} \times \frac{sR_{h1}C_{h1}}{1 + sR_{h1}C_{h1}} V_C \quad (49)$$

$$V_B(s) = \frac{-sR_{f2}C_2}{1 + sR_{f2}C_{f2}} \times \frac{sR_{h2}C_{h2}}{1 + sR_{h2}C_{h2}} V_C \quad (50)$$

$$\begin{aligned} V_{out}(s) &= k_g(V_B - V_A) \\ &= k_g V_C \left( \frac{-sR_{f2}C_2}{1 + sR_{f2}C_{f2}} \times \frac{sR_{h1}C_{h1}}{1 + sR_{h1}C_{h1}} - \frac{-sR_{f1}C_1}{1 + sR_{f1}C_{f1}} \times \frac{sR_{h2}C_{h2}}{1 + sR_{h2}C_{h2}} \right) \end{aligned} \quad (51)$$

If  $\Delta C_1 = \Delta C_2 = \Delta C$ ,  $C_{01} = C_{02} = C_0$ ,  $C_{f1} = C_{f2} = C_f$ ,  $R_{f1} = R_{f2} = R_f$ ,  $C_{h1} = C_{h2} = C_h$ ,  $R_{h1} = R_{h2} = R_h$ , and  $1 \ll \omega_d C_h R_h$ , the output can be derived as

$$\begin{aligned} V_{out}(s) &= \frac{2k_g V_C s R_f \Delta C \sin(\omega_d t + \alpha)}{1 + sR_f C_f} \times \frac{sR_h C_h}{1 + sR_h C_h} \\ &\approx \frac{2k_g V_C s R_f \Delta C \sin(\omega_d t + \alpha)}{1 + sR_f C_f} \end{aligned} \quad (52)$$

If  $1 \gg \omega_d C_f R_f$ ,  $V_{out}$  can be simplified as

$$V_{out} = \frac{2k_g V_C s R_f \Delta C \sin(\omega_d t + \alpha)}{1 + s R_f C_f} \approx 2k_g V_C \omega_d R_f \Delta C \cos(\omega_d t + \alpha) \quad (53)$$

It figures out that the phase difference between  $V_{out}$  and  $\Delta C \sin(\omega_d t + \alpha)$  is  $90^\circ$ , and the amplitude has to do with  $\omega_d$ ,  $k_g$ ,  $R_f$ , and  $V_C$ .

### Closed-Loop Design for the Drive Mode

There are two prevalent closed-loop control schemes for the drive mode: AGC (automatic gain control) self-oscillation scheme and AGC + PLL (phase-locked loop) scheme. The stability of these two control system will be deduced below.

#### AGC Self-Oscillation Scheme

Figure 6 illustrates the AGC self-oscillation control loop for the drive mode of a MEMS gyroscope.  $R$  is the setting target,  $A$  is the amplitude of the drive velocity,  $e$  is the control error,  $G$  is the control voltage,  $V_{dc}$  is the driving DC voltage, and  $C(s)$  is the PI (proportion-integration) controller.  $K_{vf}$ ,  $K_{dc}$ , and  $K_{cv}$  are the gains of voltage to force module, displacement to capacitance module, and capacitance to voltage module, respectively.

The key control parameter of the AGC loop is the error between the measured amplitude of the drive velocity and the target. Thus, the aim is to minimize this error and make the system have a swift, accurate, and robust transient response performance. PID (proportion-integration-differentiation) control determines the reaction to the past, current, and future trend of the error and produces a weighted sum of these three actions to adjust the objectives. Usually, the PID controller may be substituted by PI controller since the differentiator circuit may introduce considerable noises. Thereby, PI controller is preferable to employ in this design. The

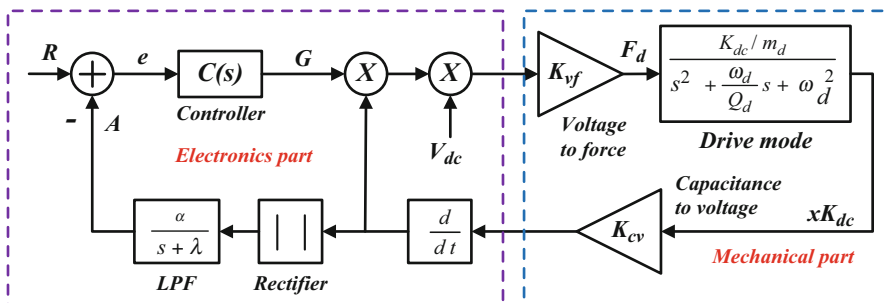


Fig. 6 The AGC self-oscillation control loop for the drive mode of a gyroscope

specific algorithm will be deduced as follows. The transfer function of  $C(s)$  can be written as

$$C(s) = K_p + K_i/s \quad (54)$$

Thus, the plant control equation of the whole loop is formulated as

$$\begin{cases} \ddot{x} + \frac{\omega_d}{Q_d}\dot{x} + \omega_d^2x = \frac{F_d}{m_d} \\ F_d = GK_{vf}V_{dc}K_{dc}K_{cv}\dot{x} \\ \dot{G} = -K_p\dot{A} + K_i(R - A) \\ \dot{A} = |\alpha K_{dc}K_{cv}\dot{x}| - \lambda A \end{cases} \quad (55)$$

From Eq. (55), the dynamic vibration equation can be rewritten as

$$\ddot{x} + \left( \frac{\omega_d}{Q_d} - \frac{GK_{vf}V_{dc}K_{dc}K_{cv}}{m_d} \right) \dot{x} + \omega_d^2x = 0 \quad (56)$$

The aim of the control system is to make the system damping approximate to zero, depending on which the self-oscillation can be realized but not utilizing PLL or other signal generator. Equation (56) illuminated that the AGC control allows the gyroscope to self-oscillate by dynamically adjusting the system's damping. The condition of self-oscillation is  $m_d\omega_d = Q_dGK_{vf}V_{dc}K_{dc}K_{cv}$ . If the measured amplitude is smaller than the desired reference, the AGC system will enlarge the loop gain to reduce the system damping. Conversely, if the oscillation amplitude is large than the reference, the AGC system will add damping to the gyroscope to restrain the amplitude. The gyroscope adjusts itself until the damping of the system is equal to zero, and the system will maintain constant amplitude oscillation. Thus,

$$\ddot{x} + \omega_d^2x = 0. \quad (57)$$

Equation (57) is a linear system with the analytic solution

$$x = u \cos(\omega_d t + \theta) \quad (58)$$

As mentioned above, the controller will minimize the damping value when there is a disturbance in the system. Thus, the damping can be limited to a certain extreme small range, which makes the nonlinear system a weak nonlinear system. In this case, the amplitude and phase of the vibration are both slowly varying parameters with respect to time. From a physical viewpoint, it means that the variation rates of the amplitude and phase are much slower with respect to time. Therefore, the averaging method is feasible to tackle this issue. The essential of the averaging method is to make  $x$  in the nonlinear system of the same form as that in the linear system, such as

$$x = u(t) \cos(\omega_d t + \theta(t)) \quad (59)$$

where  $\theta(t)$  is the phase delay referring to the drive signal and  $u(t)$  is the amplitude. From a mathematical viewpoint, expressing the amplitude and phase as functions of

time is actually equivalent to considering the disturbance caused by the nonlinear terms as a revision for the linear system's solution. Differentiating  $x$  with respect to time yields the velocity

$$\dot{x} = -u\omega_d \sin(\omega_d t + \theta) + \dot{u} \cos(\omega_d t + \theta) - u\dot{\theta} \sin(\omega_d t + \theta) \quad (60)$$

According to the averaging method, the velocity is also expected to have the same form as the position signal. Thereby, the sum of the last two terms can be assumed to be zero (Closkey et al. 2001), namely

$$\dot{u} \cos(\omega_d t + \theta) - u\dot{\theta} \sin(\omega_d t + \theta) \equiv 0. \quad (61)$$

Hence, the velocity can be rewritten as

$$\dot{x} = -u\omega_d \sin(\omega_d t + \theta). \quad (62)$$

The acceleration along the x-axis can be given by differentiating the velocity with respect to time

$$\ddot{x} = -\dot{u}\omega_d \sin(\omega_d t + \theta) - u\omega_d(\omega_d + \dot{\theta}) \cos(\omega_d t + \theta) \quad (63)$$

Applying the averaging method to solve Eq. (55), we can yield the simplified equations

$$\begin{cases} \dot{\bar{u}} = \frac{1}{2}\bar{u} \left( \frac{GK_{vf}V_{dc}K_{dc}K_{cv}}{m_d} - \frac{\omega_d}{Q_d} \right) \\ \dot{\bar{\theta}} = 0 \\ \dot{\bar{A}} = \frac{2}{\pi} \alpha K_{dc}K_{cv}\omega_d \bar{u} - \lambda \bar{A} \\ \dot{\bar{G}} = -K_p \dot{\bar{A}} + K_i (R - \bar{A}). \end{cases} \quad (64)$$

Thus, the equilibrium point of the AGC system can be calculated as

$$\bar{u} = \frac{\pi R \lambda}{2\alpha K_{dc}K_{cv}\omega_d}, \quad \bar{A} = R, \quad \bar{G} = \frac{m_d \omega_d}{K_{dc}K_{cv}K_{vf}V_{dc}Q_d} \quad (65)$$

It shows that the system has a unique equilibrium point and the steady-state error is removed entirely after introducing an integral term. The Jacobian matrix of the nonlinear system at the equilibrium point is represented by

$$J_F(\bar{u}, \bar{A}, \bar{G}) = \begin{bmatrix} 0 & 0 & \frac{\pi R \lambda K_{vf} V_{dc}}{4\alpha m_d \omega_d} \\ \frac{2}{\pi} \alpha K_{dc} K_{cv} \omega_d & -\lambda & 0 \\ -\frac{2}{\pi} \alpha K_{dc} K_{cv} K_p \omega_d & \lambda K_p - K_i & 0 \end{bmatrix} \quad (66)$$

Its characteristic equation is

$$s^3 + \lambda s^2 + \frac{R\lambda K_{cv}K_{dc}K_pK_{vf}V_{dc}}{2m_d}s + \frac{R\lambda K_{cv}K_{dc}K_iK_{vf}V_{dc}}{2m_d} = 0 \tag{67}$$

Based on the Routh–Hurwitz criterion, the linearized system is asymptotically stable only if each coefficient is positive, and the product of the inner two coefficients is larger than that of the outer two (Dorf and Bishop 2005), i.e.,

$$\lambda K_p > K_i \tag{68}$$

Equation (68) means that only if the product of  $\lambda$  and  $K_p$  is larger than  $K_i$ , is the exact solution of the original nonlinear equations convergent to the desired equilibrium point, and the steady-state error gradually reaches around zero.

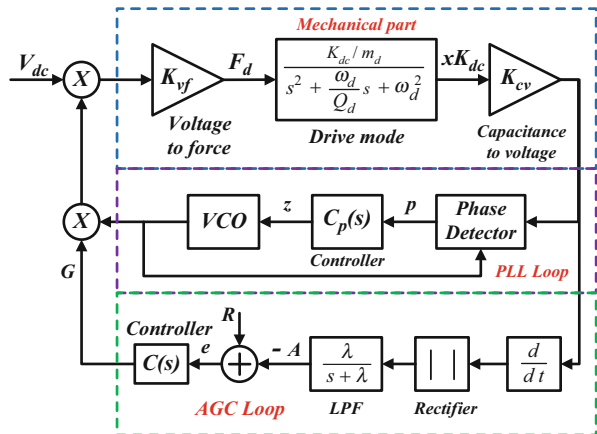
### AGC + PLL Scheme

Figure 7 illustrates the AGC + PLL control loop for the drive mode of a MEMS gyroscope. Different from the former control system, this system includes extra PLL loop, which comprises a phase detector (PD), a controller, and a VCO (voltage-controlled oscillator).  $Y\cos\theta$  is the output of VCO, and the equation of VCO can be expressed as

$$\dot{\theta} = \omega_0 + K_{VCO}z \tag{69}$$

where  $\omega_0, K_{vco}$  are the presetting frequency and tuning gain. The transfer function of PI controller  $C_p(s)$  is

**Fig. 7** The AGC + PLL control loop for the drive mode of a gyroscope



$$C_p(s) = \tau_p + \tau_i/s \tag{70}$$

Assume that the displacement is expressed as

$$x = u(t) \cos(\theta(t) + \varphi(t)) \tag{71}$$

where  $\theta(t) + \varphi(t)$  is the transient phase. Thus, the control equations of the AGC + PLL system can be yielded as

$$\begin{cases} \ddot{x} + \frac{\omega_d}{Q_d} \dot{x} + \omega_d^2 x = \frac{F_d}{m_d} \\ F_d = GK_{vf}V_{dc}y \cos \theta \\ \dot{G} = -K_p \dot{A} + K_i(R - A) \\ \dot{A} = \lambda(|K_{dc}K_{cv}\dot{x}| - A) \\ \dot{\theta} = \omega_0 + K_{VCO}z \\ \dot{z} = \tau_p \dot{p} + \tau_i p \\ \dot{p} = \sigma(K_{dc}K_{cv}xy \cos \theta - p) \end{cases} \tag{72}$$

Applying the averaging method to solve Eq. (72), we can yield the simplified equations as follows:

$$\begin{cases} \dot{\bar{u}} = -\frac{1}{\omega_0 + K_{VCO}\bar{z}} \left[ \frac{\overline{GK_{vf}V_{dc}y}}{2m_d} \sin \bar{\varphi} + \frac{1}{2}\bar{u} \left( K_{VCO}\dot{\bar{z}} + \frac{\omega_d}{Q_d}(\omega_0 + K_{VCO}\bar{z}) \right) \right] \\ \dot{\bar{\varphi}} = -\frac{1}{\omega_0 + K_{VCO}\bar{z}} \left[ \frac{\overline{GK_{vf}V_{dc}y}}{2m_d\bar{u}} \cos \bar{\varphi} + \frac{1}{2} \left( (\omega_0 + K_{VCO}\bar{z})^2 - \omega_d^2 \right) \right] \\ \dot{\bar{G}} = -K_p \dot{\bar{A}} + K_i(R - \bar{A}) \\ \dot{\bar{A}} = \lambda \left[ \frac{2}{\pi} K_{dc}K_{cv}\bar{u}(\omega_0 + K_{VCO}\bar{z}) - \bar{A} \right] \\ \dot{\bar{z}} = \tau_p \dot{\bar{p}} + \tau_i \bar{p} \\ \dot{\bar{p}} = \sigma \left( \frac{1}{2} K_{dc}K_{cv}\bar{u}y \cos \bar{\varphi} - \bar{p} \right) \end{cases} \tag{73}$$

Hence, the equilibrium point of the AGC + PLL system can be calculated as

$$\begin{aligned} \bar{p} = 0, \bar{\varphi} = -\frac{\pi}{2}, \bar{A} = R, \bar{z} = \frac{\omega_d - \omega_0}{K_{VCO}}, \bar{u} = \frac{\pi R}{2K_{dc}K_{cv}\omega_d}, \bar{G} \\ = \frac{\pi R m_d \omega_d}{2K_{dc}K_{cv}K_{vf}V_{dc}Q_d y} \end{aligned} \tag{74}$$

It shows that the system has a unique equilibrium point and the steady-state error is removed entirely after introducing an integral term. The Jacobian matrix of the nonlinear dynamic system at the equilibrium point is represented by

$$\begin{aligned}
 & J(\bar{u}, \bar{\varphi}, \bar{G}, \bar{A}, \bar{p}, \bar{z}) \\
 & = \begin{pmatrix} -\frac{\omega_d}{2Q} & -\frac{\pi^2 R^2 y \sigma K_{vco} \tau_p}{16 K_{cv} K_{dc} \omega_d^3} & \frac{y K_{vf} V_{dc}}{2 m_d \omega_d} & 0 & -\frac{\pi R K_{vco}}{4 Q K_{cv} K_{dc} \omega_d} & -\left( \frac{\pi R K_{vco} (\tau_i - \sigma \tau_p)}{4 K_{cv} K_{dc} \omega_d^2} \right) \\
 0 & -\frac{\omega_d}{2Q} & 0 & 0 & -K_{vco} & 0 \\
 -\frac{2 \lambda K_{cv} K_{dc} K_p \omega_d}{\pi} & 0 & 0 & -K_i + \lambda K_p & -\frac{R \lambda K_p K_{vco}}{\omega_d} & 0 \\
 \frac{2 \lambda K_{cv} K_{dc} \omega_d}{\pi} & 0 & 0 & -\lambda & \frac{R \lambda K_{vco}}{\omega_d} & 0 \\
 0 & \frac{\pi R y \sigma \tau_p}{4 \omega_d} & 0 & 0 & 0 & \tau_i - \sigma \tau_p \\
 0 & \frac{\pi R y \sigma}{4 \omega_d} & 0 & 0 & 0 & -\sigma \end{pmatrix} \quad (75)
 \end{aligned}$$

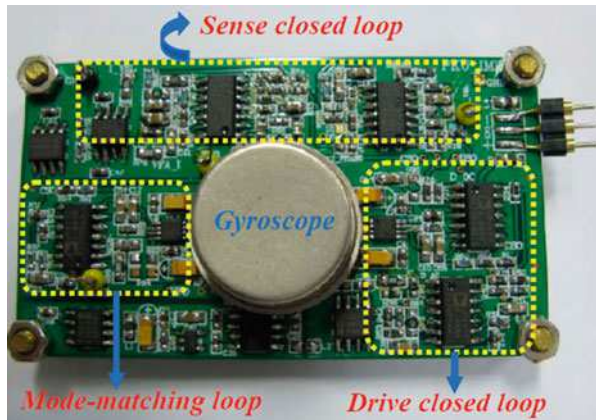
Equation (75) indicates that its characteristic equation is a sixth-order system, which complicates the analysis. Similar to the solving method above, the stable condition can be calculated. Reader can try to solve it, and we no longer illustrate here.

### Experimental Tests

The AGC control circuit for a MEMS gyroscope is shown in Fig. 8, and it is a pure analog circuit. On the other hand, the AGC + PLL control circuit is illustrated in Fig. 9. It is a mixed signal circuit, and the detection and control algorithms are implemented in FPGA (field-programmable gate array) device. Actually, except for closed-loop control for the drive mode, these two circuits can also accomplish the mode-matching control and force rebalance control. Here, we only take AGC control circuit, for example, to validate the theoretical analysis above.

The transient response of the drive velocity is snatched by a Tektronix TDS2000B oscilloscope, as shown in Fig. 10. The DC component of the drive signal is fixed at

**Fig. 8** The AGC control circuit for a MEMS gyroscope

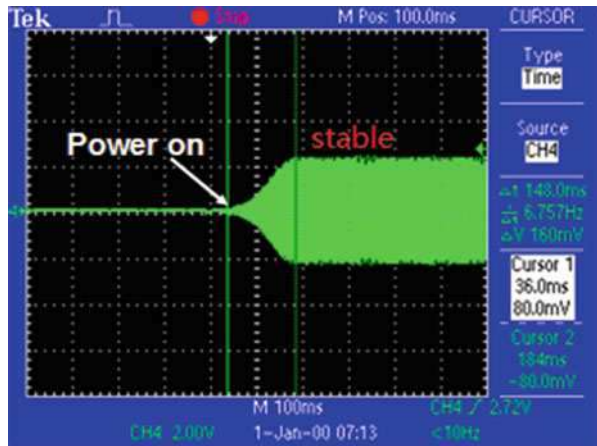




**Fig. 9** The AGC + PLL control circuit for a MEMS gyroscope



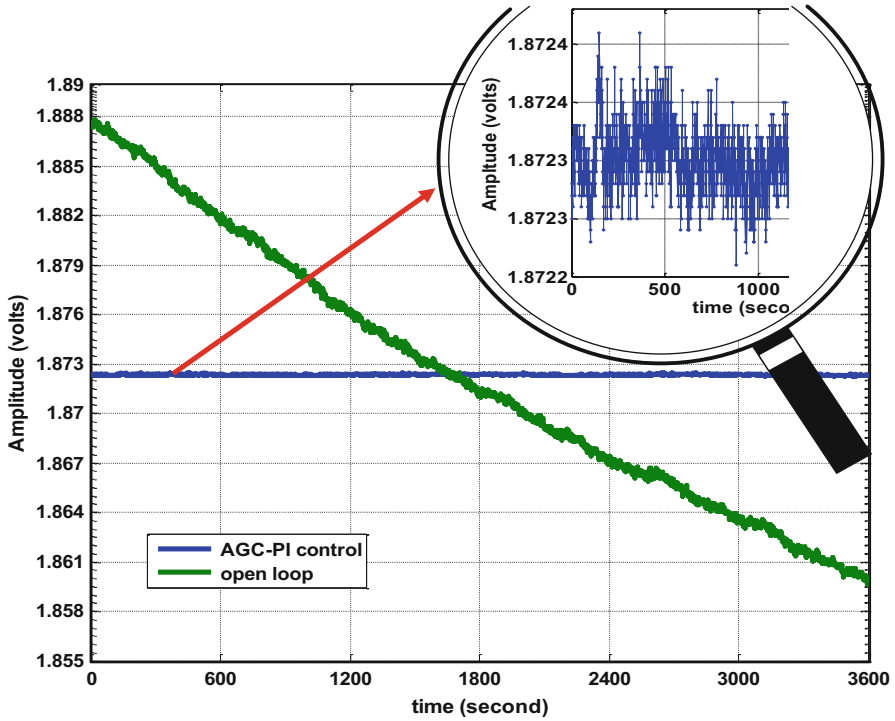
**Fig. 10** The velocity output waveform snatched by a Tek oscilloscope



5 V with automatically adjusting the AC component around 2.5 V. To restrain the circuit noise and filter the harmonic, a first-order LPF is designed with the cutoff frequency 100 Hz. The PI control parameters are set at  $K_i = 90$ ,  $K_p = 16$ , and  $R = 0.2$ . It takes about 150 ms for the gyroscope to self-oscillate to the desired target without obvious overshoot from power-on stage.

Figure 11 shows the drift characteristic of the amplitude of drive velocity for the AGC loop and open loop. Agilent 34401 is used to acquire the amplitude of the drive velocity at one second intervals within an hour. The experiment results demonstrate that the drift range of the open loop is 28.2 mV, which is about 140 times that of AGC-PI control 0.2 mV. The relative fluctuation of velocity amplitude with AGC-PI control is calculated to be about 16.6 ppm, which is reduced by two orders of magnitude compared with the drift of around 4300 ppm in the open loop.

The large fluctuation in the open loop resulted from the slight drift of the resonant frequency of the drive mode and some integrated circuits caused by the temperature rising. During our experiments, we observe that the temperature is not even on the



**Fig. 11** Stability of the drive velocity for the AGC loop and open loop

PCB by using a thermal imager since the power consumption is considerable, about 900 mW with  $\pm 15$  V power supply. The frequency drift alters the mechanical gain at the drive frequency point and temperature gradient can affect the performances of the integrated circuits. With AGC control, the resonant frequency and vibration amplitude can be precisely controlled, so the drift of the gyro can be improved substantially. Thus, it is concluded that the proposed AGC control improved the precision and stability of the gyroscope velocity remarkably, which is the fundamental prerequisite of the stability for the angular rate synchronous demodulation in the sense mode.

## Mode-Matching Control

### Temperature Characteristics Analysis

When DC voltage  $V_p$  is exerted to the mode-matching combs, according to Eq. (45), the electrostatic stiffness can be described as:

$$k_e = 2V_p^2 N \epsilon h l \left( \frac{1}{d_1^3} + \frac{1}{d_2^3} \right) \quad (76)$$

Due to mode-matching, the relationship between the two resonant frequencies and  $k_e$  can be expressed as:

$$k_e = 4\pi^2 m_s (f_s^2 - f_d^2) \quad (77)$$

In order to simplify the analysis, here, only take the effect of Young's modulus into account, thus the two resonant frequencies can be written as:

$$f_d(T) = f_{d0} \sqrt{1 - k_{ET}(T - T_0)} \approx f_{d0} \left( 1 - \frac{1}{2} k_{ET}(T - T_0) \right) \quad (78)$$

$$f_s(T) = f_{s0} \sqrt{1 - k_{ET}(T - T_0)} \approx f_{s0} \left( 1 - \frac{1}{2} k_{ET}(T - T_0) \right) \quad (79)$$

where  $f_{d0}$  and  $f_{s0}$  are the corresponding resonant frequencies of the two modes at room temperature  $T_0$ . The temperature coefficient of Young's modulus  $k_{ET}$  is between 25 ppm/ $^{\circ}$ C and 75 ppm/ $^{\circ}$ C. Thus, the frequency difference of the two modes can be derived as:

$$f_s(T) - f_d(T) = (f_{s0} - f_{d0}) \left( 1 - \frac{1}{2} k_{ET}(T - T_0) \right) \quad (80)$$

Since the resonant frequencies vary with ambient temperature  $T$ , it is important to make a real-time mode-matching control. Solving the equations above, the relationship between  $V_p$  and  $f_d$  can be yielded as:

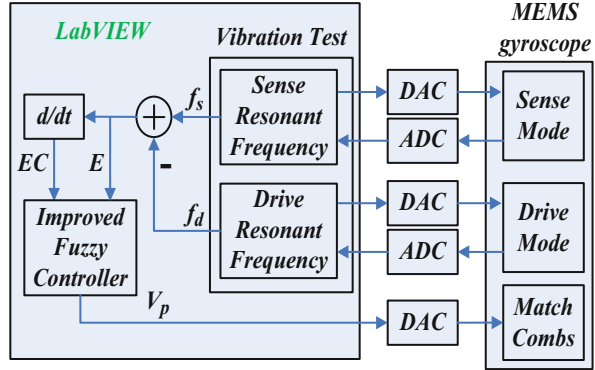
$$V_p(T) = \sqrt{\frac{4(f_{s0}^2 - f_{d0}^2)\pi^2 m_s}{f_{d0}^2 N \epsilon h l \left( \frac{1}{d_1^3} + \frac{1}{d_2^3} \right)}} f_d(T) \quad (81)$$

Equation (81) indicates that if only the temperature effect of the Young's modulus is considered,  $V_p$  is proportional to  $f_d$  and also has an approximately linear dependence on the ambient temperature. In reality,  $V_p$  is not absolutely proportional to  $f_d$ . Other nonideal factors may affect the tuning voltage, such as the variations of gap or overlap areas of the capacitors due to the structural thermal distortion.

## Automatic Mode-Matching Control System

Based on the analysis above, an automatic mode-matching control with improved fuzzy algorithm is proposed, which is illustrated in Fig. 12. It mainly comprises

**Fig. 12** The schematic of the automatic mode-matching control system with improved fuzzy controller



a gyroscope, NI data acquisition card, and LabVIEW software. ADC and DAC represent analog to digital converter and digital to analog converter, respectively. Since the resonant frequencies and quality factors can be measured by vibration test method rapidly and accurately (Yan et al. 2011), the latest match error  $E$  and its differential  $EC$  can be obtained simultaneously. The controlled plant is a complex nonlinear system, so a 2D fuzzy controller can be applied to adjust  $V_p$  according to  $E$  and  $EC$  to achieve mode-matching intelligently.

Fuzzy algorithm is one of the most important intelligent control approaches, which is competent for any nonlinear control system. Owing to the control rapidity and strong robustness with expert experience, it is widely utilized in some gyroscope control systems, such as displacement control and angular rate estimation (Xiao et al. 2009; Fei and Xin 2012). The formula of the traditional incremental fuzzy controller can be expressed as:

$$V_n = V_{n-1} + \Delta V_n = V_{n-1} + fuzzy(k_p E, k_d EC) \tag{82}$$

where  $V_n$  and  $V_{n-1}$  are the current and the last output of  $V_p$ , and  $\Delta V_n$  is the output increment determined by the fuzzy controller.  $k_p$  and  $k_d$  are the weights of  $E$  and  $EC$ .  $E$ ,  $EC$ , and  $\Delta V_n$  consist of seven language values  $\{nb, nm, ns, ze, ps, pm, pb\}$  on the discourse of  $[-3, 3]$ .  $nb$ ,  $nm$ ,  $ns$ ,  $ze$ ,  $ps$ ,  $pm$ , and  $pb$  represent negative big, negative medium, negative small, zero, positive small, positive medium, and positive big, respectively. The rule table is the core which is set based on expert experience. The rules are established referring to the following principles: (1) If  $E$  and  $EC$  are large,  $\Delta V_n$  should be adjusted to rapidly reduce the error. (2) If  $E$  and  $EC$  are small,  $\Delta V_n$  should be tuned to restrain overshoot and eliminate static error. For example, if  $E$  is  $nb$  and  $EC$  is  $nb$ , then  $\Delta V_n$  is  $pb$ . If  $E$  is  $nb$  and  $EC$  is  $pb$ , then  $\Delta V_n$  is  $ze$ .

The traditional fuzzy controller can be regarded as a nonlinear PD controller, as shown in Eq. (82), which will result in a nonideal control precision due to the lack of an integration term. According to classical control theory, the control static error cannot be eliminated without an integration term. Therefore, traditional fuzzy

algorithm should be improved by introducing an integration term  $k_i E$  to achieve a better control effect, as depicted below.

$$V_n = V_{n-1} + \Delta V_n = V_{n-1} + \text{fuzzy}(k_p E, k_d EC) + k_i E \quad (83)$$

According to the fuzzy rules and Mamdani algorithm, a look-up table (LUT) for fuzzy control can be yielded with the following formula.

$$\Delta V_n' = \bigcup_{i=1}^{49} (E' \text{ and } EC') \circ [(E_i \text{ and } EC_i) \rightarrow \Delta V_{ni}] \quad (84)$$

where  $E'$  and  $EC'$  are the typical inputs of  $E$  and  $EC$ .  $E_i$ ,  $EC_i$ , and  $\Delta V_{ni}$  are the fuzzy rules of fuzzy sets. Afterward, centroid method is used for defuzzification of  $\Delta V_n'$  to obtain  $\Delta V_n$ . Finally, the control LUT can be stored in the memory to save computing time, which simplifies the realization in LabVIEW or other devices.

The improved fuzzy controller can be considered as the combination of traditional fuzzy controller and PID controller, which makes it possess good control speed, strong robustness, and high control precision.

This novel automatic mode-matching method scarcely depends on the  $Q_s$  and is hardly affected by the uncertain parasitic phase shift induced by the readout circuits. Besides, improved fuzzy algorithm and vibration test theory are adopted to the feedback control loop, which will enhance the control precision and rapidity. Furthermore, this method can be easily realized.

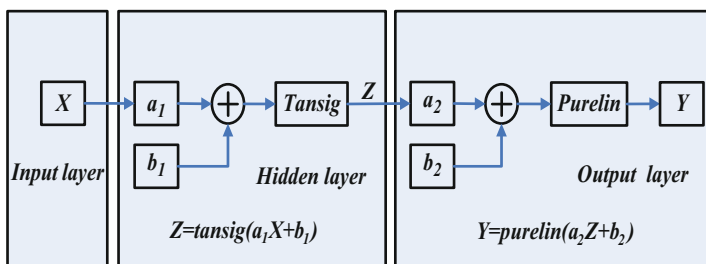
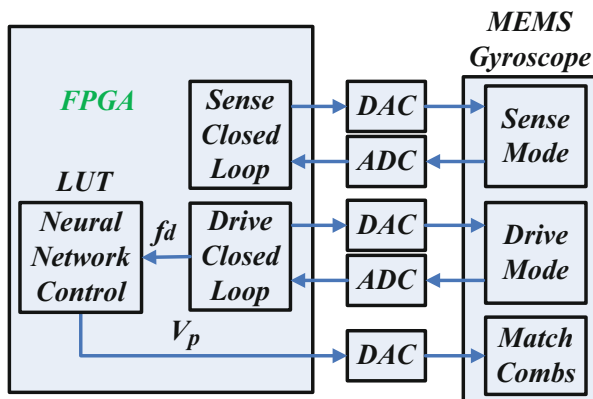
## Real-Time Mode-Matching Control System

The automatic mode-matching method depicted in Fig. 12 is still a once-time control method, and it will fail to work when Coriolis force comes into being since vibration test must work offline. However, it can be applied to obtain the corresponding  $f_d$  and  $V_p$  rapidly under a few different temperature points. Then BP (back propagation) neural network algorithm is adopted to study the relationship between  $f_d$  and  $V_p$  based on those measured data and make a prediction afterwards. Since  $f_d$  changes with ambient temperature and can be monitored in the drive closed-loop, it can be used as a temperature feedback variable and applied to real-time mode-matching control and online temperature compensation. Thus,  $V_p$  varies with  $f_d$  when ambient temperature changes, which achieves a real-time mode-matching control, as shown in Fig. 13.

Neural network algorithm is another important intelligent control approach, which can acquire the experience from some training data and then accomplish the control or prediction well. Since it possesses the strong ability of generalization, self-adaption, and nonlinear mapping for the nonlinear or uncertain system, it is usually adopted to the temperature compensation for some MEMS devices (Zhang et al. 2009; Yang et al. 2010).

In this chapter, a three-layer BP neural network will be utilized for the real-time prediction control, as shown in Fig. 14.  $X$  is the input matrix of input layer, and  $Y, Z$

**Fig. 13** The schematic of the real-time mode-matching control system with neural network controller



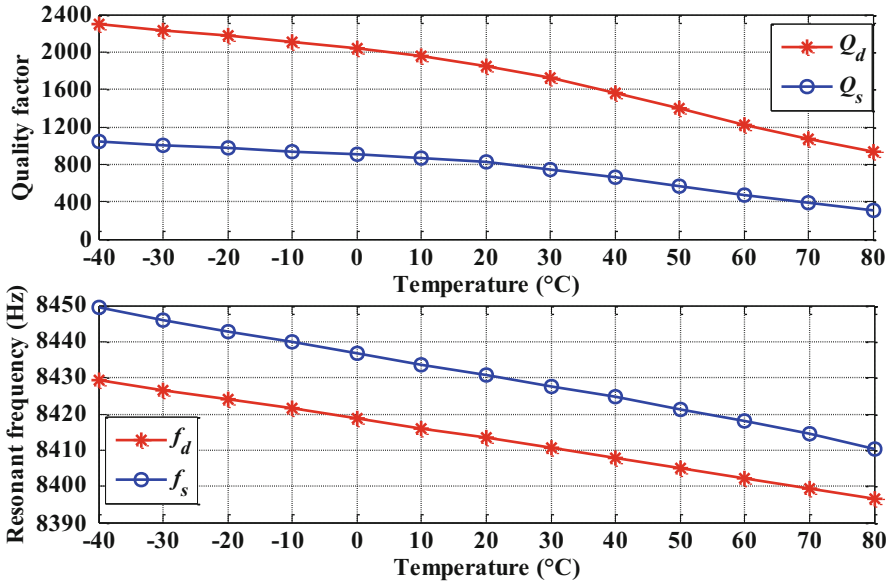
**Fig. 14** The architecture of a three-layer BP neural network controller

are the output matrices of hidden layer and output layer, respectively. Tansig and pureline functions are utilized as the transfer functions of the hidden layer and output layer, respectively. Trainlm (Levenberg–Marquardt algorithm) function is utilized for network training, and mean square error is used as the performance estimation function. In the study process, the weights matrices  $a_1$ ,  $a_2$  and the threshold matrices  $b_1$ ,  $b_2$  are adjusted with steepest descent method to make the training error ( $ER$ ) meet the expected target.

$$ER = (Y - Tar)' \times (Y - Tar) / 2 \tag{85}$$

where  $Tar$  is the expected output matrix used for the training. After the training with the acquired data,  $a_1$ ,  $a_2$  and  $b_1$ ,  $b_2$  are confirmed. Then the BP neural network can be used to make a prediction according to the inputs.

Neural network algorithm is so complex that it will occupy much calculating time and vast hardware resources, but this calculation can be accomplished by Matlab software beforehand. Similarly, a LUT of the relationship between  $f_d$  and  $V_p$  is generated by Matlab and stored in the ROM of FPGA device, which will simplify the realization.



**Fig. 15** The resonant frequencies and quality factors of the two modes both vary with the ambient temperature

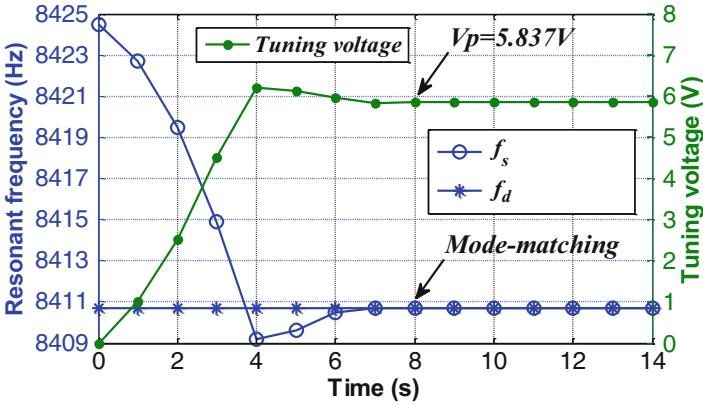
### Experimental Tests

The test circuit for a vacuum packaging gyroscope is mounted on the rate table in an oven chamber. The temperature of the oven chamber is configured from  $-40\text{ }^{\circ}\text{C}$  to  $80\text{ }^{\circ}\text{C}$ . We conduct the tests per  $10\text{ }^{\circ}\text{C}$ , and every temperature point is kept for about 30 min to make the gyroscope’s temperature identical with ambient temperature.

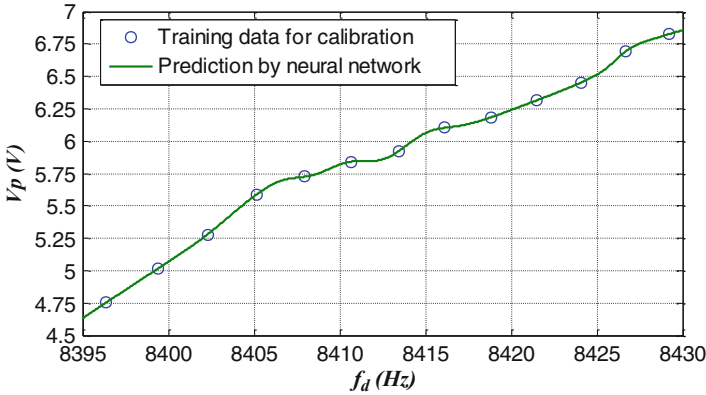
Experimental results indicate that both  $f_s$  and  $f_d$  vary with ambient temperature nearly linearly, and the change is about 35 Hz over the full temperature range, as shown in Fig. 15. But their slopes are different, which will induce some match error in the once-time mode-matching control system. Thus, mode-matching voltage should trace with ambient temperature real-time. On the other hand, due to the increasing thermal damping, both  $Q_s$  and  $Q_d$  drop fast and nonlinearly as temperature increases, as shown in Fig. 15.  $Q_s$  is always smaller than  $Q_d$  because there are so many squeeze-film combs in the sense mode.

Figure 16 demonstrates that  $f_s$  changes with tuning voltage as time goes on. It only needs about 8 s to achieve mode-matching automatically in the improved fuzzy control system, which testifies the automatic mode-matching control system is feasible, rapid, and accurate. The mode-matching voltage is measured to be 5.837 V when ambient temperature is  $30\text{ }^{\circ}\text{C}$ . After mode-matching control, the resonant frequencies of the two modes are both 8410.69 Hz.

Therefore, the automatic mode-matching control system is proved effective to obtain the corresponding  $f_d$  and  $V_p$  when ambient temperature is constant. Thirteen groups of data about  $f_d$  and  $V_p$  can be yielded with the fuzzy control system under 13



**Fig. 16** Improved fuzzy control for automatic mode matching control when ambient temperature is set 30 °C

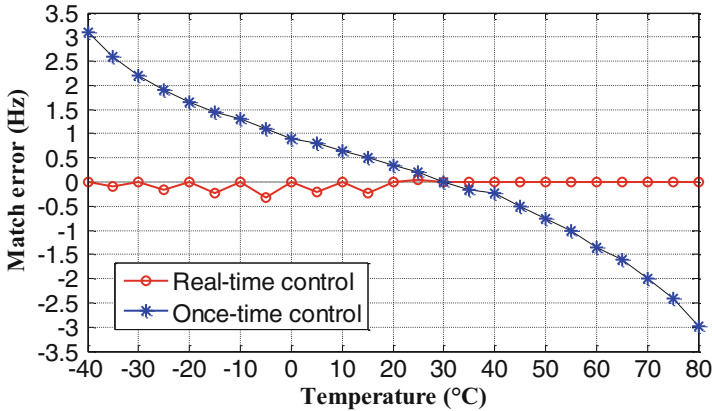


**Fig. 17** The relationship between the drive resonant frequency and the mode-matching voltage

different temperatures configured from  $-40\text{ }^{\circ}\text{C}$  to  $80\text{ }^{\circ}\text{C}$ , as marked with 13 blue circles in Fig. 17. Then neural network algorithm is applied to study the relationship between  $f_d$  and  $V_p$  based on those data, and make a prediction, as shown in the green curve in Fig. 11. It indicates that the relationship between  $f_d$  and  $V_p$  is nonlinear, which will be used for the real-time mode-matching control.

In the once-time mode-matching control system,  $V_p$  is a constant, for instance,  $V_p$  is set to be 5.837 V, thus the gyroscope can maintain mode-matching only if the ambient temperature is  $30\text{ }^{\circ}\text{C}$ . The matching error is larger than 6 Hz over the temperature range from  $-40\text{ }^{\circ}\text{C}$  to  $80\text{ }^{\circ}\text{C}$ , as depicted in the blue curve in Fig. 18. On the other hand, the matching error can be restrained to be smaller than 0.32 Hz in the full temperature range in the real-time mode-matching control system, as illustrated in the red curve in Fig. 18. The control precision is improved by more than one order of magnitude, which is enough for a gyroscope with  $Q_s$  less than 1200.





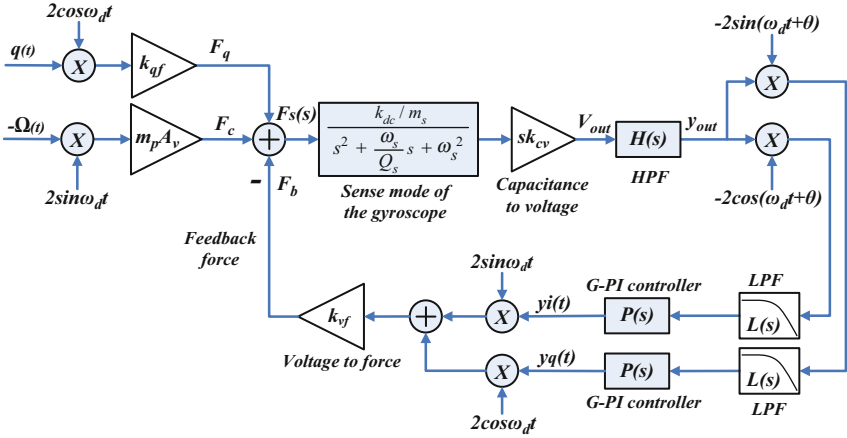
**Fig. 18** The relationships between the mismatched frequency and temperature in two control systems

## Closed-Loop Design for the Sense Mode

### Force Rebalance Control System

The force rebalance control system for the sense mode of a gyroscope is shown in Fig. 19. There are two closed loops in the system due to the demodulation and remodulation components, namely Coriolis force rebalance control loop and quadrature rebalance control loop. Given the completeness, the demodulation and remodulation signals of the two loops should be orthogonal. A HPF is adopted to eliminate the DC signal, and a LPF is used to suppress the high frequency signal. After G-PI control, a remodulation process is required to construct the feedback force to balance the external forces. Then  $y_i(t)$  and  $y_q(t)$  are used as the final outputs of the closed-loop system.  $-\Omega(t)$  and  $q(t)$  are the input angular rate and equivalent coupling signal, respectively.  $\theta$  is the demodulation phase.  $F_s$  is the total force exerted to the sense mode.  $F_c$ ,  $F_q$ , and  $F_b$  represent Coriolis force, quadrature force, and feedback force, respectively.  $m_p A_v$ ,  $k_{qf}$ , and  $k_{vf}$  are the corresponding force coefficients of them. If  $m_p A_v = k_{qf} = k_{vf}$ ,  $y_i(t)$  and  $y_q(t)$  will equal  $-\Omega(t)$  and  $q(t)$  after force rebalance, respectively.

In literature (Sonmezoglu et al. 2012; Tatar et al. 2012), the closed-loop analysis is mainly conducted in the frequency domain of angular rate ( $f_r$  domain). However, it is still not enough without the analysis in the frequency domain of force ( $f$  domain). Strictly, the closed loop is a force rebalance control system, instead of an angular rate feedback system, thus the stability analysis results in  $f$  domain will be more interesting and accurate. Through the demodulation or modulation process, signals can be transferred between  $f$  domain and  $f_r$  domain. Here, the stability analysis of the control system in both  $f_r$  domain or  $f$  domain will be conducted.



**Fig. 19** Block diagram of the force rebalance control system for the sense mode of a MEMS gyroscope

The transfer functions of sense mode of the gyroscope  $G_s(s)$ , HPF  $H(s)$ , LPF  $L(s)$ , and G-PI controller  $P(s)$  are expressed as follows

$$G_s(s) = \frac{k_{dc}}{m_s s^2 + c_s s + k_s} = \frac{k_{dc}/m_s}{s^2 + s\omega_s/Q_s + \omega_s^2} \tag{86}$$

$$H(s) = \frac{s}{s + \omega_h} \tag{87}$$

$$L(s) = \frac{\omega_l^2}{s^2 + 2\xi_l \omega_l s + \omega_l^2} \tag{88}$$

$$P(s) = \frac{k_p(s + \omega_2)}{(s + \omega_1)(s + \omega_3)} \tag{89}$$

where  $\omega_h$  is the cutoff frequency of the first-order HPF.  $\xi_l$  and  $\omega_l$  are the damping ratio and center frequency of the second-order LPF, respectively.  $(s + \omega_2)/(s + \omega_1)$  can be considered as a PI controller or a lag compensator, and  $k_p/(s + \omega_3)$  is a first-order LPF applied to eliminate the high-frequency components. Because the useful force signals locate within a narrow band around  $\omega_d$ , there is no need to fulfill a wide-band force control for making the controller design easier. Thereby, narrow-band force feedback control is adopted.

In  $f_r$  domain,  $s = j\omega_R = j2\pi f_r$ , while in  $f$  domain,  $s = j\omega = j2\pi f$ . The deductions in the two domains will be conducted.

### Analysis in the Frequency Domain of Angular Rate

The two loops in Fig. 19 are similar to each other, which can not only guarantee the same bandwidth of them but also simplify the design of the controllers. Here, only take the Coriolis force control loop, for example, to fulfill the deduction. Block diagram of the open loop system for angular rate is shown in Fig. 20. After a series of deductions in  $f_r$  domain and filtering the high-frequency terms, the transfer function of the open loop system  $D_r(s)$  can be derived as

$$D_r(s) = \frac{y_i(s)}{-\Omega(s)} = -jm_p A_v k_{cv} \left[ \frac{(s - j\omega_d)e^{j\theta} G_s(s - j\omega_d) H(s - j\omega_d) - (s + j\omega_d)e^{-j\theta} G_s(s + j\omega_d) H(s + j\omega_d)}{(s - j\omega_d)e^{j\theta} G_s(s - j\omega_d) H(s - j\omega_d) - (s + j\omega_d)e^{-j\theta} G_s(s + j\omega_d) H(s + j\omega_d)} \right] L(s)P(s) \quad (90)$$

Thus, the closed-loop transfer function  $T_r(s)$  and sensitivity transfer function  $S_r(s)$  can be defined as follows

$$T_r(s) = \frac{D_r(s)}{1 + D_r(s)} \quad (91)$$

$$S_r(s) = \frac{1}{1 + D_r(s)} \quad (92)$$

When mode-matching,  $\omega_d = \omega_s$ , which is often larger than  $2\pi \times 8$  kHz.  $\omega_h$  is set to be  $2\pi \times 80$  Hz, so  $\omega_d \gg \omega_h$ . The bandwidth (BW) of a MEMS gyroscope is often smaller than 200 Hz, thus  $2\omega_s \gg \omega_R$ . In mode-matching,  $\theta$  is set to be  $90^\circ$ . Assume that  $\omega_t = \omega_s/(2Q_s)$ , then the original mechanical response  $SF_m(s)$  can be derived from  $D_r(s)$  when the term  $L(s)P(s)$  is neglected, as follows

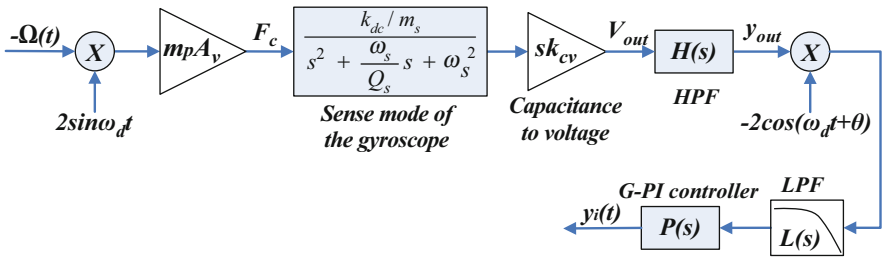


Fig. 20 Block diagram of the open loop system for angular rate

$$\begin{aligned}
SF_m(s) &= -jm_p A_v k_{cv} \left[ (s-j\omega_d) e^{j\theta} G_s(s-j\omega_d) H(s-j\omega_d) - (s+j\omega_d) e^{-j\theta} G_s(s+j\omega_d) H(s+j\omega_d) \right] \\
&= -jm_p A_v k_{cv} \left[ j(j\omega_R-j\omega_d) \times \frac{k_{dc}/m_s}{\omega_s^2 - (\omega_R-\omega_d)^2 + j(\omega_R-\omega_d)\omega_s/Q_s} \times \frac{j(\omega_R-\omega_d)}{j(\omega_R-\omega_d) + \omega_h} \right. \\
&\quad \left. + j(j\omega_R+j\omega_d) \times \frac{k_{dc}/m_s}{\omega_s^2 - (\omega_R+\omega_d)^2 + j(\omega_R+\omega_d)\omega_s/Q_s} \times \frac{j(\omega_R+\omega_d)}{j(\omega_R+\omega_d) + \omega_h} \right] \\
&\approx \frac{m_p A_v k_{cv} k_{dc}}{m_s} \left[ \frac{j(\omega_R-j\omega_d)}{(2\omega_s-\omega_R)\omega_R - j(2\omega_s-2\omega_R)\omega_t} - \frac{j(\omega_R+j\omega_d)}{(2\omega_s+\omega_R)\omega_R - j(2\omega_s+2\omega_R)\omega_t} \right] \\
&\approx \frac{m_p A_v k_{cv} k_{dc}}{m_s} \left[ \frac{j(\omega_R-j\omega_s)}{2\omega_s(\omega_R-j\omega_t)} - \frac{j(\omega_R+j\omega_s)}{2\omega_s(\omega_R-j\omega_t)} \right] = \frac{m_p A_v k_{cv} k_{dc}}{m_s(j\omega_R + \omega_t)} = \frac{m_p A_v k_{cv} k_{dc}}{m_s(s + \omega_t)}
\end{aligned} \tag{93}$$

It demonstrates that  $SF_m(s)$  is a first-order system whose cutoff frequency is  $\omega_t$ . Meanwhile, the original bandwidth is  $BW = \omega_t = \omega_s/(2Q_s)$ . After vacuum packaging,  $Q_s$  is larger than 500, which means BW will be smaller than 10 Hz. Thereby, in order to meet more application demands, BW of a mode-matching gyroscope should be extended by the force rebalance control. The scale factor of a MEMS gyroscope with closed-loop controlled sense mode can be expressed as

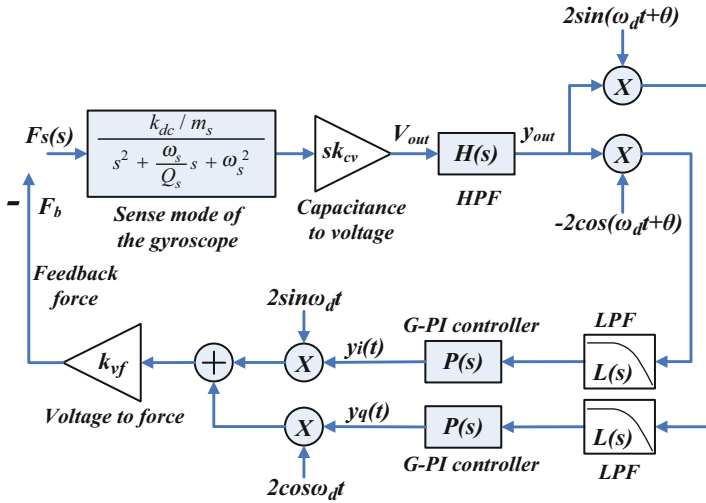
$$SF_{closed} = \frac{y_i(t)}{-\Omega(t)} = \frac{m_p A_v}{k_{vf}} \tag{94}$$

Since  $k_{vf}$ ,  $A_v$ , and  $m_p$  hardly vary with ambient temperature, compared with Eq. (31) of the open-loop system, the temperature sensitivity of the scale factor of the closed-loop system is improved. In addition, after force rebalance, the sense combs are kept in the equilibrium position, which will improve the nonlinearity and enhance the measured range. The bandwidth can be also changed flexibly by tuning the parameters of the G-PI controller. In general, the performances of a force rebalance control system are superior to those of an open-loop system.

## Analysis in the Frequency Domain of Force

Block diagram of the open-loop system for external force is illustrated in Fig. 21. After a series of deductions in  $f$  domain and filtering the high-frequency terms, the transfer function of the open-loop system  $D(s)$  can be easily derived as

$$\begin{aligned}
D(s) &= \frac{F_b(s)}{F_s(s)} \\
&= 2j s k_{cv} k_{vf} \left[ e^{-j\theta} L(s-j\omega_d) P(s-j\omega_d) - e^{j\theta} L(s+j\omega_d) P(s+j\omega_d) \right] H(s) G_s(s)
\end{aligned} \tag{95}$$



**Fig. 21** Block diagram of the open loop system for external force

Thus, the closed-loop transfer function  $T(s)$  and sensitivity transfer function  $S(s)$  can be defined as follows

$$T(s) = \frac{D(s)}{1 + D(s)} \tag{96}$$

$$S(s) = \frac{1}{1 + D(s)} \tag{97}$$

Hence, the frequency responses in  $f_r$  domain and  $f$  domain can be simulated by Matlab with Eqs. (86)–(97). The key parameters of filters and G-PI controller are finally set according to simulation results and stability indices. In the experimental tests, a virtual rate-table method (Cui et al. 2012) can be adopted to obtain the frequency responses and validate the theoretical analysis.

### Experimental Tests

Here, we take a mode-matching MEMS gyroscope, for example, to realize the force rebalance control. After mode-matching control, the resonant frequencies of the two modes are both 8410.6 Hz at room temperature, and  $Q_s$  is about 750.

The experimental frequency responses of  $D_r(s)$ ,  $T_r(s)$ ,  $S_r(s)$ , and  $SF_m(s)$  are depicted in Figs. 22 and 23. From these two diagrams, the phase margin ( $PM$ ), gain margin ( $GM$ ), cutoff frequencies ( $f_c$ ), sensitivity margins ( $M_s$ ), and bandwidth ( $BW$ ) can be calculated, as listed in Table 1.

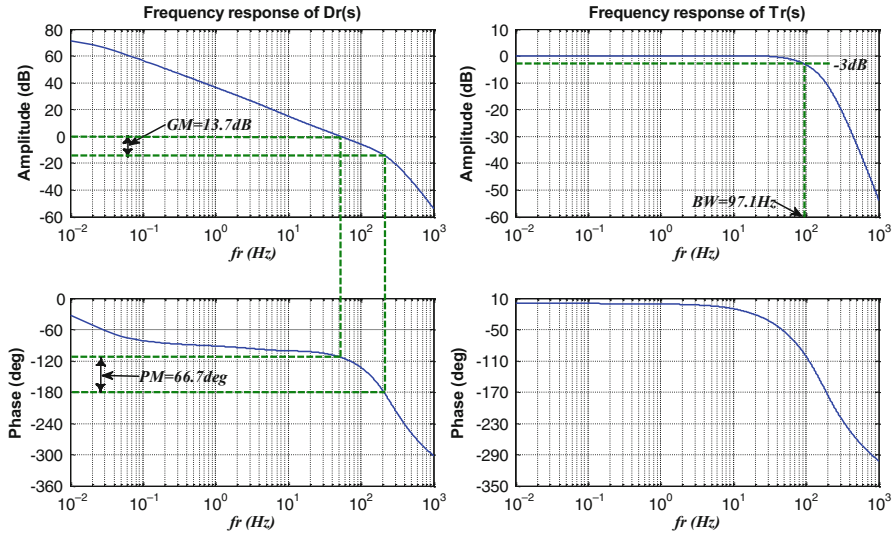


Fig. 22 Frequency responses of  $D_r(s)$  and  $T_r(s)$

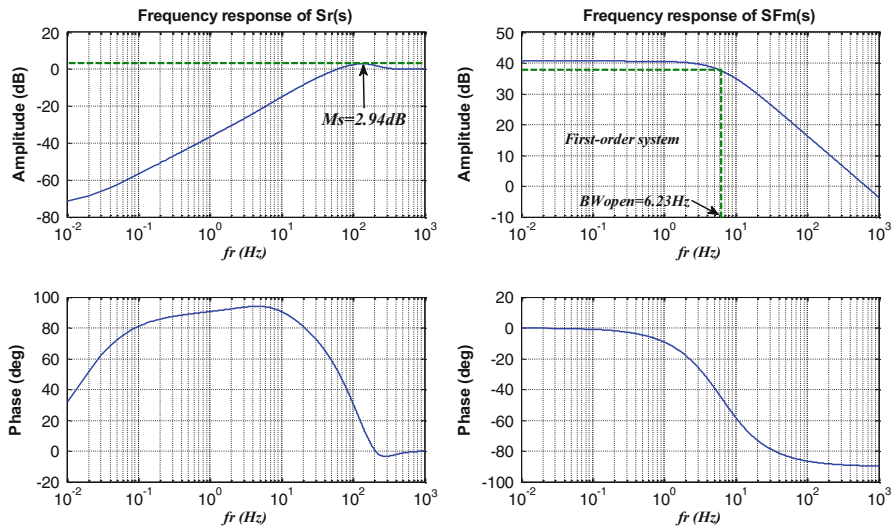
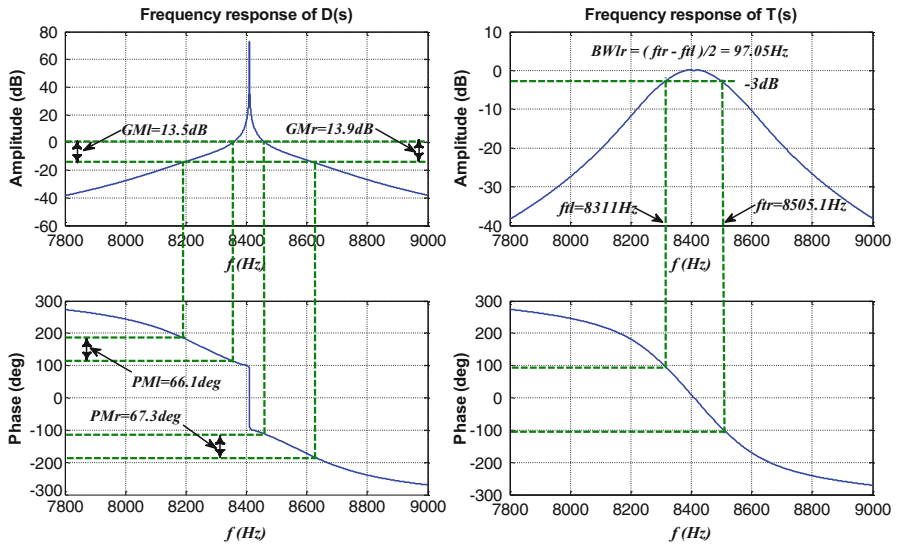


Fig. 23 Frequency responses of  $S_r(s)$  and  $SF_m(s)$

Similarly, the experimental frequency responses of  $D(s)$ ,  $T(s)$ , and  $S(s)$  are depicted in Figs. 24 and 25. Owing to narrow-band force rebalance control, there are two crossover frequencies appearing in the frequency-amplitude response of  $D(s)$  in  $f$  domain. Thus, according to classical control theory, there are two corresponding phase margins ( $PM_l$  and  $PM_r$ ), gain margins ( $GM_l$  and  $GM_r$ ), sensitivity margins ( $M_{sl}$  and  $M_{sr}$ ), and cutoff frequencies ( $f_{tl}$  and  $f_{tr}$ ), as marked in Fig. 24 and Fig. 25. They are interesting since we scarcely meet this case. After all, many control theory books

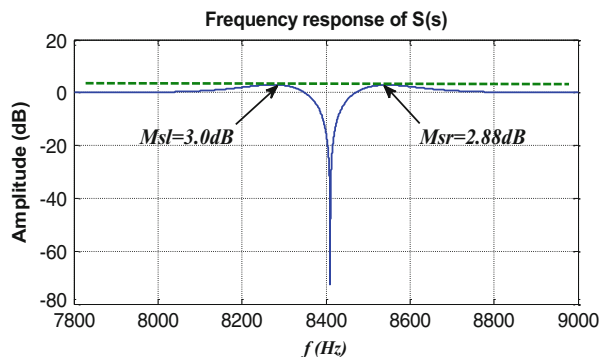
**Table 1** Stability indices of the gyroscope with closed-loop controlled sense mode

Type	$f$ domain		$f_r$ domain
Phase margin (deg)	$PM_l = 66.1$	$PM_r = 67.3$	$PM = 66.7$
Gain margin (dB)	$GM_l = 13.5$	$GM_r = 13.9$	$GM = 13.7$
Sensitivity margin (dB)	$M_{sl} = 3.0$	$M_{sr} = 2.88$	$M_s = 2.94$
Cutoff frequency (Hz)	$f_{il} = 8311$	$f_{ir} = 8505.1$	$f_i = 97.1$
Bandwidth (Hz)	$BW_r = 97.05$		$BW = 97.1$



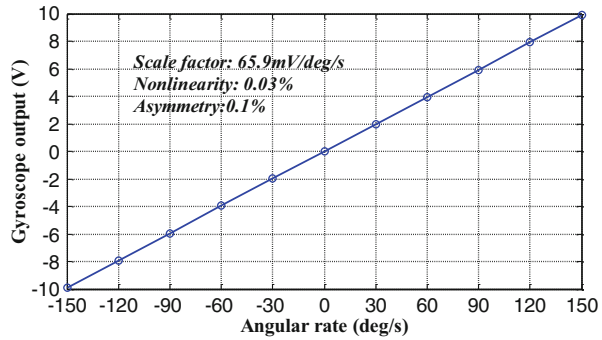
**Fig. 24** Frequency responses of  $D(s)$  and  $T(s)$

**Fig. 25** Frequency responses of  $S(s)$



usually show us that there is only one crossover frequency, one phase margin, one gain margin, one sensitivity margin, and one cutoff frequency in the frequency responses. Whatever, classical linear control theory is still suitable to analyze the stability. Although there is only one crossover frequency, one phase margin, one gain

**Fig. 26** The scale factor of the gyroscope with closed-loop controlled sense mode



margin, one sensitivity margin, and one cutoff frequency in  $f_r$  domain, the control loop is a force rebalance system, it had better study the stability in  $f$  domain, which helps us understand the system's robustness better.

The key stability indices listed in Table 1 indicate that the closed-loop system is stable and robust enough. Besides, the bandwidth  $BW_{lr}$  of  $T(s)$  is calculated to be about 97.05 Hz, which means that the force outside this narrow-band will never be balanced. Meanwhile, noises outside the bandwidth will be almost eliminated by the controller and filters. After closed-loop design,  $\omega_1$ ,  $\omega_2$ ,  $\omega_3$ , and  $k_p$  are finally set to be 0.1 rad/s, 50 rad/s, 5000 rad/s, and 400, respectively.

The mode-matching MEMS gyroscope with the closed-loop controlled sense mode achieves a scale factor of 65.9 mV/deg/s with nonlinearity about 0.03% and asymmetry about 0.1%, as shown in Fig. 26. From Fig. 27, the bias instability and angle random walk (ARW) are evaluated to be 0.68 deg/h and 0.028 deg/h<sup>1/2</sup>, respectively.

## Temperature Compensation

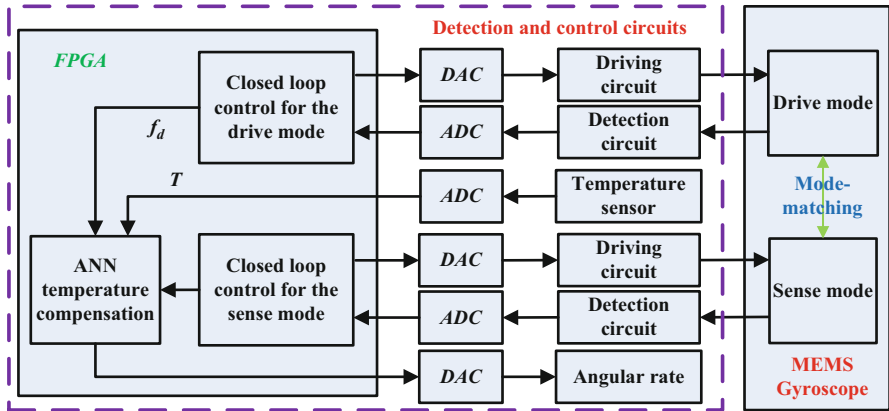
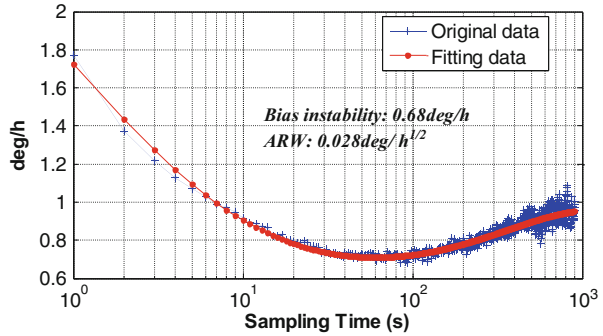
### Temperature Compensation Principle

As mentioned above,  $f_s$ ,  $f_d$ ,  $Q_s$ , and  $Q_d$  all vary with ambient temperature. Meanwhile, the parameters of capacitors, resistors, and amplifiers are also affected by ambient temperature. Hence, temperature compensation is very important to suppress the temperature sensitivity.

The schematic of an ANN temperature compensation system for a MEMS gyroscope is shown in Fig. 28. Given that  $f_d$  varies with temperature linearly, it can be utilized to temperature compensation for the drift induced by mechanical structure. Given that the temperatures of mechanical structure and circuit are a little different, a temperature sensor is adopted to measure the circuit's temperature  $T$ , which will be also applied to the compensation. Thus, an artificial neural network (ANN), such as BP neural network mentioned above, can be adopted to study the relationship between  $f_d$ ,  $T$ , and zero bias. Afterwards, the ANN temperature compensation model will be used to make a prediction and eliminate the bias



**Fig. 27** Allan variance graphs of the gyroscope with closed-loop controlled sense mode



**Fig. 28** The schematic of an ANN temperature compensation system for a MEMS gyroscope

drift. The algorithm of a BP neural network has been introduced above, so we no longer repeat here.

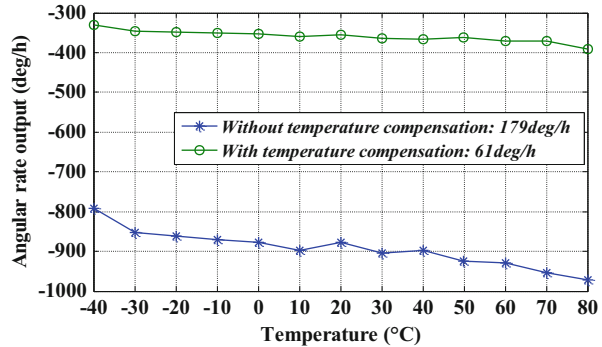
### Experimental Test

Figure 29 illustrates that zero bias varies with ambient temperature with or without temperature compensation. After compensation, the temperature sensitivity of bias drift over the temperature range from  $-40\text{ }^{\circ}\text{C}$  to  $80\text{ }^{\circ}\text{C}$  is improved to  $61\text{ deg/h}$  from  $179\text{ deg/h}$ , which indicates the proposed compensation method is feasible and effective.

### Conclusion

This chapter introduces the overview and fundamental operation principle of a MEMS vibratory gyroscope, meanwhile, some mature or novel methods of closed-loop control for the drive mode, mode-matching control, force rebalance control, and

**Fig. 29** Zero bias varies with ambient temperature with or without temperature compensation



temperature compensation are presented in detail. Besides, theoretical deduction, simulation, and experimental tests are combined tightly to make the proposed techniques understand easily. Actually, a MEMS gyroscope is a complicated system, and the relevant techniques are far more than those mentioned above.

## References

- Acar C, Shkel AM (2005) Structurally decoupled micromachined gyroscopes with post-release capacitance enhancement. *J Micromech Microeng* 15:1092–1101
- Ahn CH, Ng EJ, Hong VA et al (2015) Mode-matching of wineglass mode disk resonator gyroscope in (100) single crystal silicon. *J Microelectromech Syst* 24(2):343–350
- Aktakka E E and Najafi K (2016) A six-axis micro platform for in situ calibration of MEMS inertial sensors. *Proc MEMS 2016*, pp 243–246
- Askari S, Asadian MH, Kakavand K, Shkel AM (2006) Vacuum sealed and getter activated MEMS Quad Mass Gyroscope demonstrating better than 1.2 million quality factor. 2016 I.E. international symposium on inertial sensors and systems, pp 142–143
- Askari S, Asadian MH, Kakavand K, Shkel AM (2016) Vacuum sealed and getter activated MEMS Quad Mass Gyroscope demonstrating better than 1.2 million quality factor. 2016 I.E. international symposium on inertial sensors and systems, pp 142–143
- Cao HL, Li HS, Liu J et al (2016) An improved interface and noise analysis of a turning fork microgyroscope structure. *Mech Syst Signal Process* 70–71:1209–1220
- Casinovi G, Norouzpour-Shirazi A, Dalal M, Ayazi F (2016) Gyroscope sensing and self-calibration architecture based on signal phase shift. *Sensors Actuators A* 241:1–11
- Chen BD, Liu HZ, Wang HT et al (2011) Thermal shock induced nanocrack as high efficiency surface conduction electron emitter. *Appl Surf Sci* 257:9125–9128
- Chen F, Li XX, Kraft M (2016) Electromechanical sigma-delta modulators ( $\Sigma\Delta$ ) force feedback interfaces for capacitive MEMS inertial sensors: A review. *IEEE Sensors J* 16(17):6476–6495
- Cho J, Woo J K, et al. (2013) A high-Q birdbath resonator gyroscope (BRG). *Proc Transducers 2013*, pp 1847–1850
- Closkey RTM, Vakakis A, Gutierrez R (2001) Modelocalization induced by a nonlinear control loop. *Nonlinear Dyn* 3:221–236
- Cui J, Guo Z Y, Yang Z C, et al. (2011a) Electrical coupling suppressing for a microgyroscope using ascending frequency drive with 2-DOF PID controller. *Proc Transducers 2011*, pp 2002–2005
- Cui J, Guo ZY, Yang ZC et al (2011b) Electrical coupling suppression and transient response improvement for a microgyroscope using ascending frequency drive with a 2-DOF PID controller. *J Micromech Microeng* 21:095020

- Cui J, He CH et al (2012) Virtual rate-table method for characterization of microgyroscopes. *IEEE Sensors J* 12:2192–2198
- Dorf RC, Bishop RH (2005) *Modern control system*, 10th edn. Science Press, Beijing
- Fei JT, Xin MY (2012) An adaptive fuzzy sliding mode controller for MEMS triaxial gyroscope with angular velocity estimation. *Nonlinear Dyn* 70:97–109
- Flader IB, Ahn CH, Ng EJ et al. (2016) Stochastic method for disk resonating gyroscope mode matching and quadrature nulling. *Proc MEMS 2016*, pp 998–1001
- Fontanella R, Accardo D, Caricati E et al. (2016) An extensive analysis for the use of back propagation neural networks to perform the calibration of MEMS gyro bias thermal drift. *Proc Plans 2016*, pp 672–680
- Franssila S (2004) *Introduction to microfabrication*. Wiley, New York
- He CH, Zhao QC, Liu DC et al. (2013a) An automatic real-time mode-matching MEMS gyroscope with fuzzy and neural network control. *Proc Transducers 2013*, pp 54–57
- He CH, Zhao QC, Liu YX et al (2013b) Closed loop control design for the sense mode of micromachined vibratory gyroscopes. *Sci China Technol Sci* 56:1112–1118
- Lajimi SAM, Heppler GR, Abdel-Rahman EM (2017) A mechanical-thermal noise analysis of a nonlinear microgyroscope. *Mech Syst Signal Process* 83:163–175
- Lee A, Ko H et al (2008) Non-ideal behavior of a driving resonator loop in a vibratory capacitive microgyroscope. *Microelectron J* 39:1–6
- Li Z, Yang Z, Xiao Z et al (2000) A bulk micromachined vibratory lateral gyroscope fabricated with wafer bonding and deep trench etching. *Sensors Actuators A* 83:24–29
- Li J, Broas M, Makkonen J et al (2014) Shock impact reliability and failure analysis of a three-axis MEMS gyroscope. *J Microelectromech Syst* 23(2):347–355
- Li J, Mattila T, Vuorinen V (2015) MEMS reliability. Chapter 41 of *handbook of silicon based MEMS*. *Mater Technol* 744–763
- Liu XS, Yang ZC, Yan GZ et al. (2006) Design and fabrication of a lateral axis gyroscope with asymmetric comb-fingers as sensing capacitors. *Proc NEMS 2006*, pp 762–765
- Liu DC, He CH, Zhao QC et al (2014) Digital signal processing for a micromachined vibratory gyroscope based on a three dimensional adaptive filter demodulator. *Measurement* 50:198–202
- Makkonen J, Broas M, Li J et al. (2012) Reliability assessment of MEMS devices – a case study of a 3 axis gyroscope. *Proc ESTC 2012*, pp 1–8
- Pasquale GD, Soma A (2011) MEMS mechanical fatigue: effect of mean stress on gold microbeams. *J Microelectromech Syst* 20(4):1054–1063
- Pierron ON, Abnet CC, Muhlstein CL (2006) Methodology for low- and high-cycle fatigue characterization with kHz-frequency resonators. *Sensors Actuators A* 128:140–150
- Prikhodko IP, Trusov AA, Shkel AM (2013) Compensation of drifts in high-Q MEMS gyroscopes using temperature self-sensing. *Sensors Actuators A* 201:517–524
- Rombach S, Maurer M, Manoli Y (2015) Continuous-time lowpass and bandpass  $\Delta\Sigma$ -modulators for closed-loop readout circuits of capacitive MEMS gyroscopes. *Inertial Sensors and Systems*, pp 7–18
- Rombach S, Marx M, Nessler S et al (2016) An interface ASIC for MEMS vibratory gyroscopes with a power of 1.6 mW, 92 dB DR and  $0.007^{\circ}/s/\sqrt{\text{Hz}}$  noise floor over a 40 Hz band. *IEEE J Solid State Circuits* 51(8):1915–1927
- Rombacha S, Maurer M, Wendler D, Manoli Y (2015) Implementation of a modulated lowpass  $\Delta\Sigma$ -modulator for MEMS gyroscopes with low-power consumption and low sampling. *Procedia Eng* 120:832–835
- Senkal D, Askari S, Ahamed MJ et al (2014) 100K Q-factor toroidal ring gyroscope implemented in wafer-level epitaxial silicon encapsulation process. *Proc MEMS 2014*, pp 24–27
- Senkal D, Ahamed MJ, Askari S, Shkel AM (2015) Demonstration of 1 million Q-factor on microglassblown wineglass resonators with out-of-plane electrostatic transduction. *J Microelectromech Syst* 24(1):29–37
- Shao P, Tavassoli V, Mayberry CL, Ayazi F (2015) A 3D-HARPSS polysilicon microhemispherical shell resonating gyroscope: design, fabrication, and characterization. *IEEE Sensors J* 15(9):4974–4985

- Sharma A. (2007) CMOS systems and circuits for sub-degree per hour MEMS gyroscopes. Ph. D., Georgia Institute of Technology
- Sonmezoglu S, Alper SE, Akin T (2012) An automatically mode-matched MEMS gyroscope with 50Hz bandwidth. *Proc. MEMS 2012*:523–526
- Sonmezoglu S, Alper SE, Akin T (2014) An automatically mode-matched MEMS gyroscope with wide and tunable bandwidth. *J Microelectromech Syst* 23(2):284–297
- Tatar E (2010) Quadrature error compensation and its effects in the performance of fully decoupled MEMS gyroscope. MS. D., Middle East Technical University
- Tatar E et al (2012) Quadrature-error compensation and corresponding effects on the performance of fully decoupled MEMS gyroscopes. *J Microelectromech Syst* 21(3):656–667
- Torunbalci MM, Alper SE, Akin T (2015) Wafer level hermetic sealing of MEMS devices with vertical feedthroughs using anodic bonding. *Sensors Actuators A* 224:169–176
- Trusov AA, Prikhodko IP, et al. (2013) 1 PPM precision self-calibration of scale factor in MEMS Coriolis vibratory gyroscopes. *Proc Transducers 2013*, pp 2531–2534
- Vafanejad A, Kim ES (2015) Effect of diaphragm perforation on quality factor of hemispherical resonator gyroscope. *Proc Transducers 2015*, pp 27–30
- Xiao QJ, Li SY, Chen WY et al (2009) Fuzzy tuning PI control for initial levitation of micro-machined electrostatically levitated gyroscope. *Electron Lett* 45(16):818–819
- Yan JJ, He CH, Ge F et al (2011) Research of resonant frequency and quality factor test methods for high vacuum sealed microgyroscope based on LabVIEW. *Procedia Eng* 15:2566–2570
- Yang C, Li HS (2015) Digital control system for the MEMS tuning fork gyroscope based on synchronous integral demodulator. *IEEE Sensors J* 15(10):5755–5764
- Yang C, Li C, Zhang C (2010) The application of RBF neural network in the compensation for temperature drift of the silicon pressure sensor. *Proc Int ICCDA'10, Qinhuangdao*, 25–27 June 2010, vol 2, pp 434–437
- Yazdi N, Ayazi F, Najafi K (1998) Micromachined inertial sensors. *Proc IEEE* 86(8):1640–1659
- Yole Developpement 2015 (2016) <http://www.yole.fr/>
- Zhang QT, Tan ZF, Guo LD (2009) Compensation of temperature drift of MEMS gyroscope using BP neural network. *Proc Int ICIECS'09, Wuhan*, 19–20 Dec 2009, pp 1–4
- Zhang JM, He CH, Liu YX et al (2015) A novel scale-factor calibration method for a MEMS gyroscope based on virtual Coriolis force. *Proc NEMS 2015*, pp 58–62
- Zihajehzadeh S, Loh D, Lee TJ et al (2015) A cascaded Kalman filter-based GPS\_MEMS-IMU integration for sports applications. *Measurement* 73:200–210



# A Double Differential Torsional MEMS Accelerometer with Improved Temperature Robustness

Dingbang Xiao, Xuezhong Wu, Qingsong Li, and Zhanqiang Hou

## Contents

Introduction .....	463
Micromechanical Accelerometers and Their Applications .....	463
Micromechanical Accelerometers Classification .....	464
Challenge for Micromachined Accelerometers .....	465
Capacitive Torsional Accelerometers .....	465
Overview .....	466
Fundamentals of Micromechanical Capacitive Torsional Accelerometers .....	466
Linear Accelerometers and Torsional Accelerometers .....	466
Model of Torsional Accelerometers .....	468
Operational Principle of the Traditional Capacitive Torsional Accelerometer .....	471
Double Differential Capacitive Torsional Accelerometers .....	475
Mechanical Design of the Double Differential Capacitive Torsional Accelerometer .....	479
Design of the Torsional Beam .....	479
Design of the Unbalanced Structure .....	486
Design of the Stress-Released Frame .....	488
Electrode Design .....	490
Analysis of the Static Mechanical Sensitivity .....	491
Mode Analysis .....	492
Analysis of the Cross-Axis Error .....	494
Analysis of the Temperature Robustness .....	494
Analysis of the Shock Resistance .....	497
Fabrication Technique .....	498
Fabrication of the Silicon Structure .....	498
Fabrication of the Glass Substrate and Anodic Bonding .....	499
Trimming and Packaging .....	499
Readout Circuit of the Double Differential Capacitive Torsional Accelerometer .....	502
Electrical Model of the Accelerometer .....	502
Traditional Capacitance Readout Circuit .....	503
Readout Circuit with Integrated Reference Capacitor .....	504

---

D. Xiao (✉) · X. Wu · Q. Li · Z. Hou  
National University of Defense Technology, Changsha, China  
e-mail: [dingbangxiao@nudt.edu.cn](mailto:dingbangxiao@nudt.edu.cn)

Experiment and Results Discussion .....	506
Test of the Scale Factor .....	506
Test of the Nonlinearity .....	507
Test of Cross-Axis Error .....	507
Test of Natural Frequency and Q .....	508
Test of the Bandwidth .....	508
Test of Bias Stability .....	509
Noise of the Accelerometer .....	511
Test of the Temperature Performance .....	511
Conclusion .....	514
References .....	514

### Abstract

This chapter presents the fundamental theory, mechanical design, fabrication technique, detecting circuit, and characterization of a novel double differential capacitive torsional accelerometer. The accelerometer consists of a double differential sensing structure with four proof masses hanging on a common V-shaped torsional beam which mainly aims to improve the temperature robustness and long-term performance of the torsional accelerometer. This chapter supplies a new method for the accelerometer performances improvement and this method can also be used in the design of other sensors.

### Keywords

Torsional accelerometers · Double differential · High performance

Micromachined accelerometers are one of the most important types of silicon-based sensors and they have been used in massive products. Among kinds of micromachined accelerometers, the capacitive torsional accelerometer is one of the most popular structures for its simple design, low cost, high sensitivity, good dc response and noise performance, and other outstanding performances. However, significant challenges remain in increasing their environment robustness and long-term performance.

This chapter will present a novel double differential capacitive torsional accelerometer mainly aiming at improving its temperature robustness and long-term performance. It mainly consists of a double differential sensing structure with four proof masses hanging on a common V-shaped torsional beam and a new readout circuit with integrated on-chip reference capacitor.

Main topics of this chapter include the fundamental theory, mechanical design, fabrication technique, detecting circuit, and characterization of the double differential torsional accelerometer, forming a complete process of the accelerometer design. Testing results for the prototypes show that the accelerometer designed in this chapter has satisfactory performances and it demonstrates the feasibility and effectivity of the accelerometer design proposed in this chapter.

The goal of this chapter is to provide the reader with fundamental knowledge of the novel double differential accelerometer. Moreover, this chapter supplies a new method for the accelerometer performances improvement and this method can also be used in the design of other sensors.

---

## Introduction

In this section, we present a brief overview of the micromechanical accelerometers and their application, the main types and the challenge of micromechanical accelerometers, and one of the most popular types of mechanical accelerometer: capacitive torsional accelerometers.

## Micromechanical Accelerometers and Their Applications

In simple terms, accelerometers are sensors that measure the linear acceleration of a moving object, with the main type being mechanical. Micromachined accelerometers are silicon-based mechanical accelerometers and as their performance keeps constantly improving in time, they are becoming a viable alternative to expensive and bulky conventional mechanical accelerometers. With micro-machining processes that allow batch production of micro-electro-mechanical system (MEMS) on a chip similar to integrated circuit, unit costs unimaginable in any other technology are achieved. Moreover, advances in the fabrication techniques that allow electronics to be integrated on the same silicon chip together with the mechanical sensor elements provide an unmatched integration capability (Fischer et al. 2015). In the past years, they alone have the second largest sales volume after pressure sensors. They are widely used in automotive applications, such as the active safety systems, implement vehicle stability systems and electronic suspension; in biomedical applications for activity monitoring; in numerous consumer applications, such as active stabilization of picture in camcorders, head-mounted displays and virtual reality and sport equipment; in industrial applications such as robotics and machine and vibration monitoring; and in many other applications, such as tracking and monitoring mechanical shock and vibration during transportation and handling of a variety of equipment and goods. High-sensitivity accelerometers are crucial components in self-contained navigation and guidance systems, seismometry for oil exploration and earthquake prediction, and micro-gravity measurements and platform stabilization in space. The impact of low-cost, small, high-performance, micromachined accelerometers in these applications is not just limited to reducing overall size, cost, and weight. It opens up new market opportunities such as personal navigators for consumer applications, or it enhances the overall accuracy and performance of the systems by making formation of large arrays of devices feasible (Yazdi et al. 1998).

## Micromechanical Accelerometers Classification

A variety of transduction mechanisms have been used in microaccelerometers. The main types of accelerometers are reviewed here.

1. Piezoresistive accelerometers: These accelerometers consist of a seismic mass carried by one or several suspension beams, with at least one of the suspension beams containing piezoresistors, in most cases four piezoresistors in a full Wheatstone bridge arrangement (Liu et al. 2014). As the support frame moves relative to the proof mass, the suspension beams will elongate or shorten, which changes their stress profile and hence the resistivity of their embedded piezoresistors. The main advantage of piezoresistive accelerometers is the simplicity of their structure and fabrication process, as well as their readout circuit. However, piezoresistive accelerometers have larger temperature sensitivity and smaller overall sensitivity compared to capacitive devices, and hence a larger proof mass is preferred for them.
2. Tunneling accelerometers: These accelerometers use a constant tunneling current between one tunneling tip (attached to a movable microstructure) and its counter electrode to sense displacement. Tunneling accelerometers can achieve very high sensitivity with small sizes since the tunneling current is highly sensitive to displacement, typically changed by a factor of 2 for each angstrom of displacement (Liu and Kenny 2001; Kubena et al. 1996; Dong et al. 2005). However, these devices usually have larger low-frequency noise levels and a high supply voltage (tens to hundreds of volts) is required for these devices, thus limiting their application.
3. Resonant accelerometers: Silicon resonant accelerometers are generally based on transferring the proof-mass's inertial force to axial force on the resonant beams and hence shifting their frequency (Pinto et al. 2009; Trusov et al. 2013). The resonant sensing technique is highly sensitive and has the potential for large dynamic range, good linearity, low noise and potentially low power. However, these devices typically have small bandwidths (less than a few hertz) and high temperature dependency.
4. Thermal accelerometers: These accelerometers are based on thermal transduction. One of the first thermal accelerometers used the principle that the temperature flux from a heater to a heat-sink plate is inversely proportional to their separation (Garraud et al. 2011; Rocha et al. 2011; Hiratsuka et al. 1991). Hence, by measuring the temperature using thermopiles, the change in separation between the plates (which is representative of acceleration) can be measured.
5. Capacitive accelerometers: Capacitive micromechanical accelerometers appeared in the late 1980s (Rudolf 1983). They rapidly established themselves as preferred devices as for the high sensitivity, good dc response and noise performance, low drift, low temperature sensitivity, low-power dissipation, and a simple structure. In addition, capacitors serve both as sensing and electrostatic actuation elements (Li et al. 2014; Xiao et al. 2016a). This dual nature of capacitive detection and electrostatic driving is one of the key advantages of this type of accelerometer,



which makes, for example, a self-test option easy to implement into the evaluation algorithm. In addition, force-feedback in a closed-loop evaluation circuit is possible, which keeps the seismic mass at a nearly fixed position by compensating external forces with corresponding forces from the electrostatic actuator. This reduces the effective deflection of the seismic mass by the loop gain factor, thus improving linearity, noise behavior, and dynamic response of the system in comparison to open-loop operation (Korvink et al. 2006). However, capacitive accelerometers can be susceptible to electromagnetic interference (EMI), as their sense node has high impedance. This issue can be addressed by proper packaging and by shielding the accelerometer and its interface circuit (Yazdi et al. 1998).

## Challenge for Micromachined Accelerometers

Today, MEMS accelerometers with low performances are in large volume production. However, the high-precision accelerometers are expensive and they are not evaluated completely as inertial grade yet. Therefore, the development of low-cost, inertial-grade accelerometers with sub-g noise levels, good long-term stability, and low temperature sensitivity still remains a challenge. The development of low-stress, low-drift packaging technologies for inertial-grade devices without affecting device performance and stability is also being pursued. A major challenge is the interface circuit, where low-drift readout/control circuitry with high sensitivity, low noise level, and large dynamic range is needed. Also, the highly stable (inertial-grade) multi-axis devices (Rodjegård et al. 2005a, b) and the integrating MEMS and ICs (Fischer et al. 2015) will be developed in the near future.

## Capacitive Torsional Accelerometers

To overcome the challenge presented above, various accelerometer structures are proposed. Among them, the capacitive torsional accelerometer is one of the most popular structures for its simple design and low cost. The typical torsional accelerometer consists of a torsional beam and two proof masses suspended on it (Ghisi et al. 2012; Tseng et al. 2014; Lee et al. 2005). When it is accelerated in the out-plane direction, the two masses lose their balance and the beam undergoes a torsional deformation, resulting in differential output capacitance.

While torsional accelerometers have been proved to be successful devices, significant challenges still remain in increasing their environment robustness and long-term performance. To get high-end accelerometers that can be used for precise measurements and in complex surroundings, environment robustness and stability of the accelerometers should be improved.

Some methods have been proposed to solve the problem. A fabrication process was presented to solve the structure curling problems (Tan et al. 2011) and environment-resistant packaging technology was developed to improve the temperature

robustness of MEMS sensors, but the methods are complicated and expensive (Lee and Najafi 2007).

Inspired from the high-performance butterfly gyroscope (Su et al. 2014; Andersson et al. 1999) and the differential vibrating beam accelerometer (Trusov et al. 2013), a double differential capacitive MEMS accelerometer is proposed (Xiao et al. 2016b).

## Overview

This chapter is organized in five sections. The first section reviews the fundamental operational principles of MEMS capacitive accelerometers. The second section reviews the design of the double differential accelerometer's mechanical sensing element. The third section reviews the fabrication process of the designed accelerometer. The fourth section reviews the readout circuit for the accelerometer, and the fifth section reviews the testing result of the accelerometer.

---

## Fundamentals of Micromechanical Capacitive Torsional Accelerometers

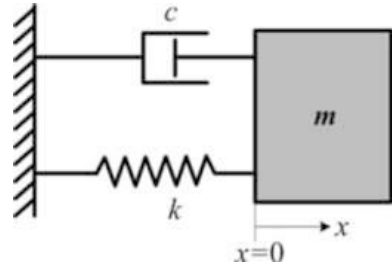
In this section, we review the fundamental operational principles of micro-mechanical capacitive torsional accelerometer. First, the dynamics of linear and torsional accelerometer sensing elements are developed. Then, the model of general torsional accelerometers is analyzed. Furthermore, the operational principle of the traditional and the novel double differential capacitive torsional accelerometers is analyzed. The models and characteristics analysis in this section are the basis of the mechanical design of capacitive torsional accelerometers.

### Linear Accelerometers and Torsional Accelerometers

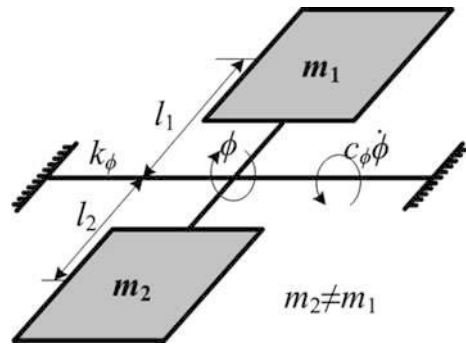
Mechanical accelerometers are mainly divided into two types based on the moving style of the proof mass: linear accelerometers and torsional accelerometers.

The linear or translational accelerometer generally consists of a proof mass suspended by compliant beams anchored to a fixed frame as shown in Fig. 1. The proof mass has a mass of  $m$ , the suspension beams have an effective spring constant of  $k$ , and there is a damping factor ( $c$ ) affecting the dynamic movement of the mass. The accelerometer can be modeled by a 1-DOF second-order mass-damper-spring system. External acceleration displaces the support frame relative to the proof mass, which in turn changes the internal stress in the suspension spring. Both this relative displacement and the suspension beam stress can be used as a measure of the external acceleration. By using Newton's

**Fig. 1** Mechanical model of the general linear or translational accelerometer



**Fig. 2** Mechanical model of the general torsional accelerometer



second law and the accelerometer model, the equation of motion of the system can be obtained

$$m \frac{d^2x(t)}{dt^2} + c \frac{dx(t)}{dt} + kx(t) = ma(t), \tag{1}$$

where  $a$  is the external acceleration, and  $x$  is the proof mass displacement.

In contrast, the torsional or rotation-based accelerometer generally consists of two unbalanced proof masses suspended on a torsional beam anchored to a fixed frame as shown in Fig. 2. The unbalanced structure has the rotational inertia of  $J = m_2 l_2^2 + m_1 l_1^2$ , the torsional beam has an effective torsional constant of  $k_\phi$ , and there is a damping factor ( $c_\phi$ ) affecting the torsional movement of the masses. Similar with the linear motion accelerometer, the torsional accelerometer can also be modeled by a 1-DOF second-order mass-damper-spring system. The equation of motion of the system can be expressed as

$$J \frac{d^2\phi(t)}{dt^2} + c_\phi \frac{d\phi(t)}{dt} + k_\phi\phi(t) = M_\phi a(t), \tag{2}$$

where  $a$  is the external acceleration along the sensitive direction,  $\phi$  is the twist angle of the beam, and  $M_\phi$  is the constant which can be calculated with  $M_\phi = m_2 l_2 - m_1 l_1$ .

## Model of Torsional Accelerometers

### Dynamic Equation and Transfer Function

The equation of motion of the torsional accelerometer system has been shown in Eq. 2 and it can be simplified as

$$\frac{d^2\phi(t)}{dt^2} + \frac{c_\phi}{J} \frac{d\phi(t)}{dt} + \frac{k_\phi}{J} \phi(t) = \frac{M_\phi}{J} a(t). \quad (3)$$

With the definition of the undamped natural frequency  $\omega_n = \sqrt{k_\phi/J}$  and the quality factor  $Q = \sqrt{k_\phi J}/c_\phi$ , the mechanical transfer function in Laplace domain can be obtained

$$\begin{aligned} H(s) &= \frac{\phi(s)}{a(s)} \\ &= \frac{\frac{M_\phi}{J}}{s^2 + \frac{c_\phi}{J}s + \frac{k_\phi}{J}} = \frac{\frac{M_\phi}{J}}{s^2 + \frac{\omega_n}{Q}s + \omega_n^2} \end{aligned} \quad (4)$$

### Mechanical Sensitivity

When the accelerometer is excited with a harmonic acceleration  $a = a_0 \sin \omega t$  at the frequency  $\omega$ , the steady-state component of the response is also harmonic, of the form (Thomson and Dahleh 1998)

$$\phi(t) = \phi_0 \sin(\omega t + \varphi), \quad (5)$$

$$\phi_0 = \frac{\frac{M_\phi}{J} \frac{1}{\omega_n^2} a_0}{\sqrt{\left[1 - \left(\frac{\omega}{\omega_n}\right)^2\right]^2 + \left(\frac{1}{Q} \frac{\omega}{\omega_n}\right)^2}}, \quad (6)$$

$$\varphi = -\tan^{-1} \frac{\frac{1}{Q} \frac{\omega}{\omega_n}}{1 - \left(\frac{\omega}{\omega_n}\right)^2}. \quad (7)$$

Herein,  $\phi_0$  is the dynamic twist angle of the beam, while in the presence of static acceleration ( $\omega = 0$ ), the static mechanical sensitivity of the accelerometer is defined as

$$S_{\text{static}} = \frac{\phi_{\text{static}}}{a} = \frac{M_\phi}{J} \frac{1}{\omega_n^2} = \frac{M_\phi}{k_\phi}. \quad (8)$$

As evident, the resonant frequency of the structure can be increased by increasing the torsional stiffness and decreasing the rotational inertia, while the quality factor of the device can be increased by reducing damping and increasing rotational inertia and torsional constant. Lastly, the static response of the device can be improved by reducing its resonant frequency.

**Bandwidth**

From Eq. 6, it can be seen that the dynamic sensitivity is the product of the static sensitivity and a gain whose magnitude affected by the frequency of input acceleration. Thus change of the gain magnitude will influence the detecting precision of the accelerometer. Bandwidth of the accelerometer is defined as the frequency range of the input acceleration that the accelerometer can precisely detect. Typically it is specified as the cut-off frequency coinciding to the -3dB point (Yazdi et al. 1998). However, as for the accelerometers with high quality factor, the gain will increase with the frequency of input acceleration increasing, which also makes the detecting result inaccurate. So the bandwidth of accelerometer will be cut-off frequency coinciding to the -3dB point or +3dB.

From Eq. 6, it is obvious that the dynamic sensitivity is relative with the resonant frequency and the quality factor. Define that

$$\lambda = \frac{\omega}{\omega_n}, \tag{9}$$

then the dynamic sensitivity can be simplified as

$$\phi_0 = \frac{\phi_{static}}{\sqrt{(1 - \lambda^2)^2 + \frac{1}{Q^2}\lambda^2}}, \tag{10}$$

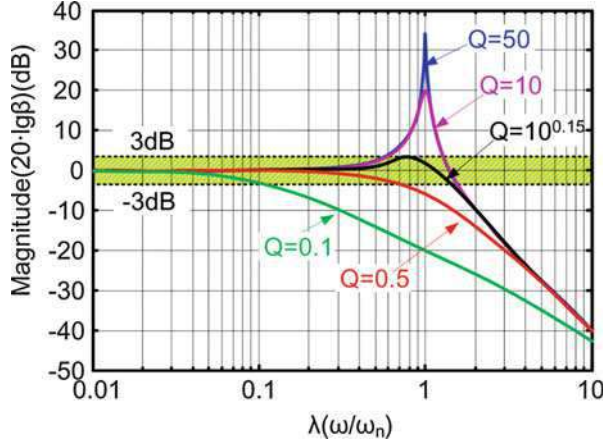
and the gain of dynamic and static sensitivity can be expressed as

$$\beta = \frac{\phi_0}{\phi_{static}} = \frac{1}{\sqrt{(1 - \lambda^2)^2 + \frac{1}{Q^2}\lambda^2}}. \tag{11}$$

Bode plot of Eq. 11 can then be obtained as Fig. 3.

It is obvious that three different cases can be distinguished according to the value of  $Q$ : under-damped system ( $Q > 0.5$ ), critically damped system ( $Q = 0.5$ ), and over-damped system ( $Q < 0.5$ ). In the critically and over-damped system, the gain has peak at the resonant frequency which means that the gain will always decrease with the frequency of input acceleration increasing. In contrast, in the under-damped system, the gain has a peak at the resonant frequency, which means that with the frequency of input acceleration increasing, the gain will firstly increase and then decrease. When the peak value is equal to +3dB at  $\lambda = 1$ ,  $Q$  can be calculated as

**Fig. 3** Bode plot of the gain of dynamic and static sensitivity at different quality factors



$10^{0.15}$ . So, when the quality factor  $Q < 10^{0.15}$ , the bandwidth should be specified as the cut-off frequency coinciding to the  $-3\text{dB}$  point. Then the bandwidth ( $\lambda\omega_n$ ) can be calculated with the equation

$$\beta_{(-3\text{dB})} = 10^{-1.5} = \frac{1}{\sqrt{(1 - \lambda^2)^2 + \frac{1}{Q^2}\lambda^2}}. \tag{12}$$

In contrast, when the quality factor  $Q \geq 10^{0.15}$ , the bandwidth should be specified as the cut-off frequency coinciding to the  $+3\text{dB}$  point. Then the bandwidth can be calculated with the equation

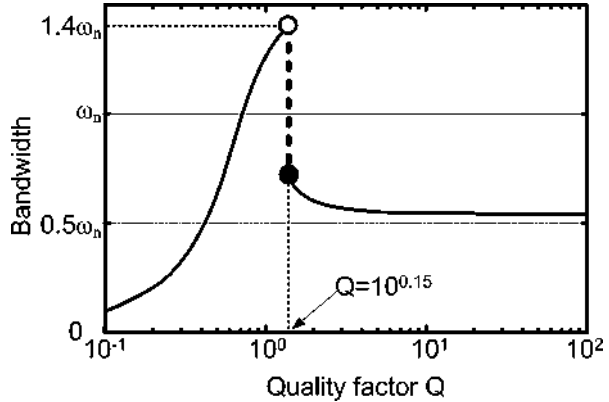
$$\beta_{(3\text{dB})} = 10^{1.5} = \frac{1}{\sqrt{(1 - \lambda^2)^2 + \frac{1}{Q^2}\lambda^2}}. \tag{13}$$

From the Eqs. 12 and 13,  $\lambda$  can be solved as

$$\lambda = \begin{cases} \sqrt{1 - \frac{1}{2Q^2} + \sqrt{\frac{1}{4Q^4} - \frac{1}{Q^2} + 10^{0.3}}} & Q < 10^{0.15} \\ \sqrt{1 - \frac{1}{2Q^2} - \sqrt{\frac{1}{4Q^4} - \frac{1}{Q^2} + 10^{-0.3}}} & Q \geq 10^{0.15} \end{cases} \tag{14}$$

And then the relationship between the bandwidth and the quality factor can be obtained as shown in Fig. 4. It is obvious that the bandwidth will be maximized when quality factor is slightly lower than  $10^{0.15}$ .

**Fig. 4** Relationship between the bandwidth and the quality factor



Exactly, this bandwidth is only for the mechanical sensing element of the accelerometer. In the whole system including the readout circuit, the bandwidth is also determined by the low pass filter, which will be introduced in section “[Readout Circuit of the Double Differential Capacitive Torsional Accelerometer.](#)”

Thus according to the requirements for bandwidth in different applications, the quality factor of accelerometer sensing element and the low pass filter in the readout circuit should be reasonably chosen.

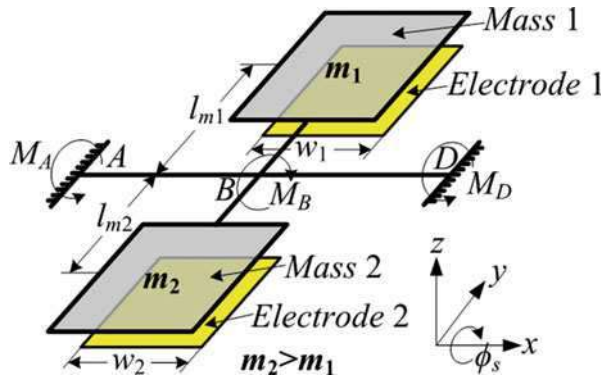
## Operational Principle of the Traditional Capacitive Torsional Accelerometer

In capacitive sensors, deflection of the seismic mass is usually transformed into corresponding capacitance changes in a differential capacitor arrangement. As the seismic mass approaches one detecting surface, thereby increasing the corresponding capacitance value, it simultaneously moves away from the other detecting surface, thereby decreasing the corresponding capacitance value.

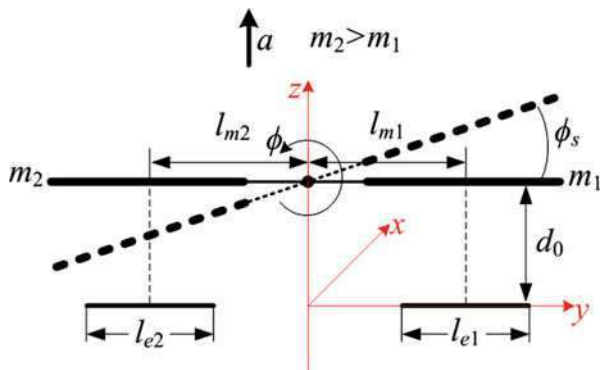
The torsional accelerometer, as shown in Fig. 5, is a typical differential capacitive micromechanical accelerometer. It is composed of the sensing structure and the corresponding detecting electrodes. The sensing structure consists of two masses whose mass are different and a torsional beam whose anchors are point A and D. The corresponding electrode and silicon mass form a capacitor and then the accelerometer’s sensing element can be regarded as two differential capacitors. Normally, the area of the electrode is smaller than that of the corresponding mass to make sure the capacitor’s relative area is constant even with some fabrication imperfections. As the mass of the masses are different, when static acceleration is applied out-of-plane, the unbalanced masses will have a rotating motion around the torsional beam, resulting in the difference of the two capacitors, as shown in Fig. 6.

Then the torque,  $T$ , of the unbalanced structure at the acceleration  $a$ , can be calculated

**Fig. 5** Schematic of the typical capacitive torsional accelerometer



**Fig. 6** Schematic of the torsional deflection of the capacitive torsional accelerometer when out-of-plane acceleration is applied



$$T = (m_2 l_{m2} - m_1 l_{m1}) a, \tag{15}$$

where,  $m_i$  is the mass of mass  $i$ , and  $l_m$  is the horizontal distance between the center of the mass and the beam. As shown in Fig. 5, the torsional beam is separated into two parts, AB and BD, whose length are  $l_{AB}$  and  $l_{BD}$ , respectively. According to the balanced equation of the beam, it can be obtained that

$$\begin{aligned} M_B &= T \\ -M_A - M_D + M_B &= 0, \end{aligned} \tag{16}$$

herein,  $M_A$  is the torque applied on the beam by the constraint at point A, also it is the torque applied on the part AB of the beam. Similarly,  $M_D$  is the torque applied on the beam by the constraint at point D, also it is the torque applied on the part BD of the beam. Meanwhile, according to the constraint conditions at the anchors, there is

$$\begin{aligned} \phi_{AD} &= \phi_{AB} + \phi_{BD} = 0 \\ \phi_A &= \phi_D = 0 \\ \phi_{AB} &= \phi_A - \phi_B = -\phi_B \\ \phi_{BD} &= \phi_B - \phi_D = \phi_B \end{aligned} \tag{17}$$



Where,  $\phi_{XY}$  is the relative twist angle between the point  $X$  and  $Y$ , while  $\phi_X$  is the absolute twist angle at point  $X$ . Based on the basic beam theory (Timoshenko and Goodier 1970), relationships between the twist angles and the torques are

$$\begin{aligned} \phi_{AB} &= -\frac{M_A l_{AB}}{GI_t}, \\ \phi_{BD} &= \frac{M_D l_{BD}}{GI_t} \end{aligned} \tag{18}$$

herein,  $G$  is the shear modulus, and  $I_t$  is torsional moment of inertia which is relative with the cross section. So it can be obtained that

$$\frac{M_D l_{BD}}{GI_t} - \frac{M_A l_{AB}}{GI_t} = 0. \tag{19}$$

By solving the Eqs. 16 and 19, the torques,  $M_A$  and  $M_B$ , can be obtained

$$M_D = \frac{Tl_{AB}}{l_{AB} + l_{BD}}, M_A = \frac{Tl_{BD}}{l_{AB} + l_{BD}}. \tag{20}$$

So the static twist angle of the beam,  $\phi_s$ , which is equal with  $\phi_B$ , can be calculated as

$$\phi_s = \phi_B = \frac{Tl_{AB}l_{BD}}{GI_t(l_{AB} + l_{BD})}. \tag{21}$$

Define that

$$\begin{aligned} l_{AB} &= kl_{BD} \\ L &= l_{AB} + l_{BD} = (k + 1)l_{BD}. \end{aligned} \tag{22}$$

Herein,  $L$  is the total length of the beam. Then the static twist angle can be expressed as

$$\phi_s = \frac{k}{(k + 1)^2} \frac{TL}{GI_t} = \frac{1}{\left(\sqrt{k} + \frac{1}{\sqrt{k}}\right)^2} \frac{TL}{GI_t}. \tag{23}$$

As known to us

$$\sqrt{k} + \frac{1}{\sqrt{k}} \geq 2, \tag{24}$$

and it has the minimum value at the condition

$$\sqrt{k} = \frac{1}{\sqrt{k}}, k = 1. \tag{25}$$

So the static twist angle,  $\phi_s$ , can be maximized when  $l_{AB} = l_{BD}$ , which means that  $B$  is the middle point of the beam. Meanwhile, the maximum static twist angle of the beam is

$$\phi_{s\max} = \frac{TL}{4GI_t}. \quad (26)$$

As shown in Fig. 6, the changed capacitance of capacitor 1,  $\Delta C_1$ , is

$$\begin{aligned} \Delta C_1 &= \int_{l_{m1} - \frac{1}{2}l_{e1}}^{l_{m1} + \frac{1}{2}l_{e1}} \frac{\varepsilon w_1}{d_0 + y \tan \phi_s} dy - C_{01} \\ &= \frac{\varepsilon w_1}{\tan \phi_s} \ln \frac{d_0 + \left(l_{m1} + \frac{1}{2}l_{e1}\right) \tan \phi_s}{d_0 + \left(l_{m1} - \frac{1}{2}l_{e1}\right) \tan \phi_s} - \frac{\varepsilon w_1 l_{e1}}{d_0}. \end{aligned} \quad (27)$$

Where,  $C_{01}$  is the original capacitance of capacitor 1,  $\varepsilon$  is the permittivity of the gas in the gap,  $w_1$  is the length of electrode 1 in  $x$ -axis,  $l_{e1}$  is the length of electrode 1 in  $y$ -axis, and  $d_0$  is the original gap of the capacitors. As the twist angle,  $\phi_s$ , is extremely small, Eq. 27 can be expanded with the Taylor Series at the point  $\phi_s = 0$ :

$$\Delta C(\phi_s) = \sum_{n=0}^{\infty} \frac{\Delta C^{(n)}(0)}{n!} (\phi_s - 0)^n, \quad (28)$$

and then we can obtain the approximate relationship between the capacitance and the twist angle:

$$\Delta C_1 \approx -\frac{w_1 \varepsilon l_{e1} l_{m1}}{d_0^2} \phi_s + \frac{(l_{e1}^3 + 12l_{e1} l_{m1}^2) w_1 \varepsilon}{12d_0^3} \phi_s^2. \quad (29)$$

Similarly, the capacitance of capacitor 2,  $\Delta C_2$ , can be calculated as:

$$\begin{aligned} \Delta C_2 &= \int_{-l_{m2} - \frac{1}{2}l_{e2}}^{-l_{m2} + \frac{1}{2}l_{e2}} \frac{\varepsilon w_2}{d_0 + y \tan \phi_s} dy - C_{02} \\ &= \frac{\varepsilon w_2}{\tan \phi_s} \ln \frac{d_0 + \left(-l_{m2} + \frac{1}{2}l_{e2}\right) \tan \phi_s}{d_0 + \left(-l_{m2} - \frac{1}{2}l_{e2}\right) \tan \phi_s} - \frac{\varepsilon w_2 l_{e2}}{d_0}, \end{aligned} \quad (30)$$

and the approximate relationship between the capacitance and the twist angle is

$$\Delta C_2 \approx \frac{w_2 \varepsilon l_{e2} l_{m2}}{d_0^2} \phi_s + \frac{(l_{e2}^3 + 12l_{e2} l_{m2}^2) w_2 \varepsilon}{12d_0^3} \phi_s^2. \quad (31)$$

Where,  $C_{02}$  is the original capacitance of capacitor 2,  $w_2$  is the length of electrode 2 in the  $x$ -axis, and  $l_{e2}$  is the length of electrode 2 in the  $y$ -axis. Then the differential capacitance of capacitor 1 and 2 can be calculated:

$$\Delta C = \Delta C_2 - \Delta C = (A_{e2}l_{m2} + A_{e1}l_{m1})\frac{\epsilon}{d_0^2}\phi_s + \sum_{i=1,2} (l_{e2}^2 + 12l_{m2}^2)A_{e2} \frac{\epsilon}{12d_0^3}\phi_s^2. \tag{32}$$

Where  $A_{ei}$  ( $A_{ei} = w_i l_{ei}$ ) is the relative area of capacitor  $i$ . To improve the linearity of the accelerometer, the coefficient of the second part in Eq. 32 should be zero, which is

$$(l_{e2}^2 + 12l_{m2}^2)A_{e2} = (l_{e1}^2 + 12l_{m1}^2)A_{e1}. \tag{33}$$

Although there are endless solutions to achieve the above equation, the simplest and most effective method is making

$$\begin{aligned} l_e &= l_{e1} = l_{e2} \\ l_m &= l_{m1} = l_{m2} \\ A_e &= A_{e1} = A_{e2} \end{aligned} \tag{34}$$

Then the capacitance sensitivity of the accelerometer can be obtained:

$$\begin{aligned} S_c &= \frac{\Delta C}{a} = \frac{2A_e l_m \epsilon}{ad_0^2} \phi_{s, \max} \\ &= \frac{(m_2 - m_1)A_e l_m^2 \epsilon L}{2GI_t d_0^2}. \end{aligned} \tag{35}$$

The corresponding torsional constant is

$$k_\phi = \frac{T}{\phi_{s, \max}} = \frac{4GI_t}{L}, \tag{36}$$

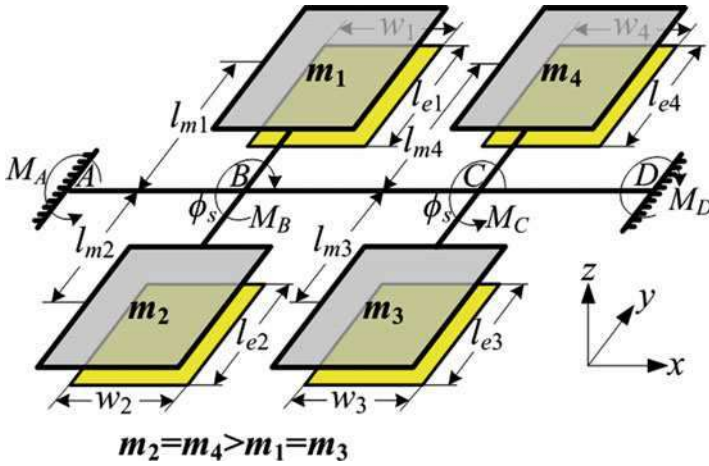
and natural frequency of the accelerometer is

$$\omega_n = \sqrt{\frac{k_\phi}{J}} = \sqrt{\frac{4GI_t}{(m_1 + m_2)l_m^2 L}}. \tag{37}$$

### Double Differential Capacitive Torsional Accelerometers

The torsional accelerometer with double differential operational principle is shown in Fig. 7.

The double differential torsional accelerometer consists of four masses and the four masses hang on a common torsional beam. The four masses can be separated to



**Fig. 7** Schematic of the double differential torsional accelerometer

two pairs of torsional structures: one pair includes mass 1 and 2, while the other one includes mass 3 and 4. Meanwhile, mass of the four masses are set as

$$m_2 = m_4 > m_1 = m_3. \tag{38}$$

And the distances between geometrical centers of the four masses and the torsional beam are set as

$$l_m = l_{m1} = l_{m2} = l_{m3} = l_{m4}, \tag{39}$$

to make the structure symmetrical. When an out-of-plane acceleration is applied on the accelerometer, the two pairs of torsional structures will rotate reversely, as the masses with the same mass locate at different sides of the beam. Then the total output capacitance can be calculated as by the double differential algorithm:

$$\begin{aligned} \Delta C &= (\Delta C_2 - \Delta C_1) - (\Delta C_3 - \Delta C_4) \\ &= (\Delta C_2 + \Delta C_4) - (\Delta C_1 + \Delta C_3). \end{aligned} \tag{40}$$

Similarly with the torques and twist angles calculation in traditional torsional accelerometers, the following relationship equations between the torques can be obtained:

$$\begin{aligned} M_B &= T = (m_2 - m_1)l_m a \\ M_C &= (m_4 - m_3)l_m a = M_B \\ -M_A + M_B - M_C + M_D &= 0, \end{aligned} \tag{41}$$

The torques' positive direction is shown in Fig. 7. Relationships between the twist angles are:

$$\begin{aligned}
 \phi_{AD} &= \phi_{AB} + \phi_{BC} + \phi_{BD} = 0 \\
 \phi_{AB} &= \phi_A - \phi_B = -\phi_B \\
 \phi_{BC} &= \phi_B - \phi_C \\
 \phi_{CD} &= \phi_C - \phi_D = \phi_C,
 \end{aligned}
 \tag{42}$$

and relationships between the twist angles and torques are:

$$\begin{aligned}
 \phi_{AB} &= -\frac{M_A l_{AB}}{GI_t} \\
 \phi_{BC} &= \frac{(M_B - M_A)l_{BC}}{GI_t} = -\frac{(M_D - M_C)l_{BC}}{GI_t}. \\
 \phi_{CD} &= -\frac{M_D l_{CD}}{GI_t}
 \end{aligned}
 \tag{43}$$

Solving the Eqs. 41 to 43, the torques and twist angles can be calculated as

$$\begin{aligned}
 M_A = M_D &= \frac{Tl_{BC}}{l_{AB} + l_{BC} + l_{CD}} \\
 \phi_B &= \frac{M_A l_{AB}}{GI_t} = \frac{Tl_{BC}l_{AB}}{GI_t(l_{AB} + l_{BC} + l_{CD})} \\
 \phi_C &= -\frac{M_D l_{CD}}{GI_t} = -\frac{Tl_{BC}l_{CD}}{GI_t(l_{AB} + l_{BC} + l_{CD})}.
 \end{aligned}
 \tag{44}$$

Herein,  $l_{AB}$ ,  $l_{BC}$ , and  $l_{CD}$  are the lengths of the beam’s parts AB, BC, and CD, respectively. It can be found that the two twist angles,  $\phi_B$  and  $\phi_C$  will be equal when the lengths  $l_{AB}$  and  $l_{CD}$  are equal with each other. With the definition that

$$\begin{aligned}
 l_e &= l_{AB} = l_{CD} \\
 \phi_s &= \phi_B = -\phi_C \\
 l_s &= l_{BC} = kl_e \\
 L &= l_s + 2l_e = (k + 2)l_e,
 \end{aligned}
 \tag{45}$$

herein  $L$  is the total length of the torsional beam, the static absolute twist angle of one torsional structure,  $\phi_s$ , can be expressed as

$$\begin{aligned}
 \phi_s &= \frac{Tl_s l_e}{GI_p(l_s + 2l_e)} \\
 &= \frac{k}{(k + 2)^2} \frac{TL}{GI_t} = \frac{1}{\left(\sqrt{k} + \frac{2}{\sqrt{k}}\right)^2} \frac{TL}{GI_p}.
 \end{aligned}
 \tag{46}$$

As

$$\sqrt{k} + \frac{2}{\sqrt{k}} \geq 2\sqrt{2},
 \tag{47}$$

and it has the minimum value when

$$\sqrt{k} = \frac{2}{\sqrt{k}}, k = 2, \quad (48)$$

so the static twist angle,  $\phi_s$ , can be maximized when  $k = 2$ . This means that to achieve the maximum mechanical sensitivity, the lengths of part BC and AB (CD) should keep the relationship

$$l_s = 2l_e. \quad (49)$$

And then the maximum twist angle

$$\begin{aligned} \phi_{s \max} &= \frac{(m_2 - m_1)l_m L}{8GI_t} a \\ &= \frac{(m_4 - m_3)l_m L}{8GI_t} a. \end{aligned} \quad (50)$$

Similar with the Eqs. 27 to 32, the double differential capacitance can be calculated as

$$\begin{aligned} \Delta C &= (\Delta C_2 + \Delta C_4) - (\Delta C_1 + \Delta C_3) \\ &= \left( \frac{\varepsilon l_m}{d_0^2} \sum_{i=1}^4 A_{ei} \right) \phi_s \\ &\quad + \left[ \frac{\varepsilon}{12d_0^3} \sum_{i=1}^4 (-1)^i (l_{ei}^2 + 12l_m^2) A_{ei} \right] \phi_s^2. \end{aligned} \quad (51)$$

Herein,  $A_{ei}$  is the relative area of the capacitor  $i$ ,  $l_{ei}$  is the length in  $y$ -axis of electrode  $i$ . To improve the linearity of the accelerometer, the coefficient of the second part in Eq. 51 should be zero, which is

$$\sum_{i=1}^4 (-1)^i (l_{ei}^2 + 12l_m^2) A_{ei} = 0. \quad (52)$$

Although there are endless results to achieve the above equation, the simplest and most effective method is making

$$\begin{aligned} l_e &= l_{e1} = l_{e2} = l_{e3} = l_{e4} \\ A_e &= A_{e1} = A_{e2} = A_{e3} = A_{e4}. \end{aligned} \quad (53)$$

Then the capacitance sensitivity of the double differential torsional accelerometer is obtained:

$$\begin{aligned}
 S_c &= \frac{\Delta C}{a} = \frac{4A_e l_m \varepsilon}{ad_0^2} \phi_{s \max} \\
 &= \frac{(m_2 - m_1)A_e l_m^2 \varepsilon L}{2GI_t d_0^2}.
 \end{aligned}
 \tag{54}$$

The corresponding torsional constant is

$$k_\phi = \frac{T}{\phi_{s \max}} = \frac{8GI_t}{L},
 \tag{55}$$

and natural frequency of the accelerometer is

$$\omega_n = \sqrt{\frac{k_\phi}{J}} = \sqrt{\frac{8GI_t}{(m_1 + m_2)l_m^2 L}}.
 \tag{56}$$

It can be seen that, comparing with the traditional single differential torsional accelerometers, the double differential torsional accelerometers have larger torsional stiffness and resonant frequencies, which benefits the improvement of the bandwidth of the accelerometer. Meanwhile, it has the same mechanical sensitivity with the traditional torsional accelerometer. Furthermore, the double differential accelerometer can effectively improve the temperature robustness, and demonstration of this advantage will be introduced in the following sections.

---

## Mechanical Design of the Double Differential Capacitive Torsional Accelerometer

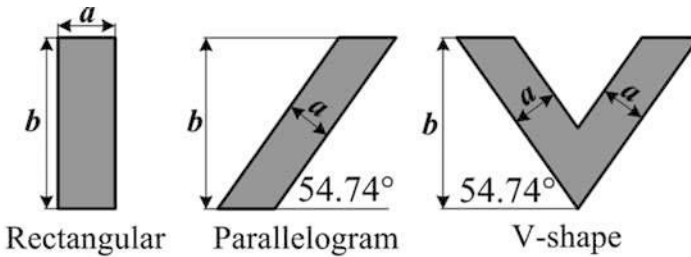
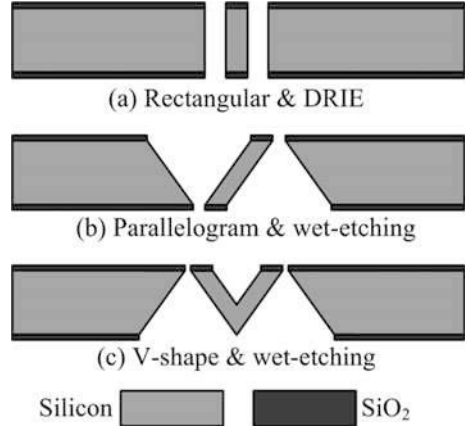
This section introduces the design process of the novel double differential capacitive torsional accelerometer proposed in this chapter, including design of the torsional beam, the unbalanced structure, the stress-released frame, and the detecting electrode. In the design, mathematical analysis and finite element method (FEM) are used for the analysis of mechanical sensitivity, mode, cross-axis error, temperature robustness, and shock resistance of the designed accelerometer.

### Design of the Torsional Beam

#### Three Types of Common Beams with Different Cross Sections

The torsional beam is the most sensitive element in the torsional accelerometer. We prefer to decrease the torsional stiffness of the beam to increase the accelerometer's sensitivity, while increase its bending stiffness to reduce the interference of the bending deflection to the preferred torsional deflection and the cross-axis errors. Thus, design of the torsional beam is the most important step of the accelerometer's design.

**Fig. 8** Schematic of the three types of beams with different cross sections and the corresponding fabrication techniques for them



**Fig. 9** Geometry parameters of the three types of beams with different cross sections

In MEMS sensors, the deep reactive ion etching (DRIE) technique and wet-etching techniques are the main methods for fabrication (Laermer et al. 1999). With the DRIE technique, the beam with a rectangular cross section is easy to achieve, as shown in Fig. 8a. While with the silicon anisotropic wet-etching technique (Elwenspoek and Wiegierink 2001), beams with parallelogram and V-shape cross sections are easy to achieve by covering different sections in the etching process, as shown in Fig. 8b, c. According to the character of anisotropic wet-etching, the angle formed by (100) and (111) crystal planes is 54.74°.

As shown in Fig. 9, the height of all the beams is defined as  $b$ , which is a certain value as for a certain silicon wafer, and width of the beams is defined as  $a$ . In this chapter, the beams with different cross sections are analyzed and compared using the mathematical method in terms of torsional stiffness and bending stiffness.

**Torsional Stiffness of Beams with Different Cross Sections**

It has been derived in Eq. 55 that the torsional constant of the beam in double differential torsional accelerometers can be calculated with

$$k_{\phi} = \frac{T}{\phi_{s\max}} = \frac{8GI_t}{L} = \frac{8}{L}T_{\phi}, \tag{57}$$



where,  $T_\phi = GI_t$  is defined as the torsional stiffness,  $I_t$  is the torsional moment of inertia and it is determined by the geometry and parameters of the cross section. As for the beams with rectangular cross section, the torsional moment of inertia can be calculated with (Timoshenko and Goodier 1970)

$$I_t = \beta h \delta^3. \tag{58}$$

Where,  $h$  is the length of the central line along the long side of the rectangular,  $\delta$  is the width and the coefficient  $\beta$  is determined by  $h$  and  $\delta$ :

$$\beta = \frac{1}{3} \left( 1 - \frac{192}{\pi^5} \frac{\delta}{h} \sum_{n=1,3,5}^{\infty} \frac{1}{n^5} \tanh \frac{n\pi h}{2\delta} \right). \tag{59}$$

As for the other thin-walled structure whose length of central line is much larger than its width, such as the parallelogram and V-shape cross section, their torsional moment of inertia is difficult to calculate due to the irregular geometries. Here, to simplify the analysis, the approximate algorithm is used, which approximately regards them as the combination of several rectangular cross sections. Then the length  $h$  is the longer central line of the thin-walled structure and  $\delta$  is the width of the thin-walled structure, then the torsional moment of inertia can be calculated with Eqs. 58 and 59. However, the calculation error will grow as the beam’s width increases.

By the approximate algorithm, torsional stiffness of beams with the above three types of cross sections can be obtained:

$$\begin{aligned} T_\phi^{rec} &= GI_t^{rec} = G\beta_{rec} ba^3 \\ T_\phi^V &= GI_t^V = G\beta_V \frac{2b \cos \theta - a}{\sin \theta \cos \theta} a^3. \\ T_\phi^{par} &= GI_t^{par} = G\beta_{par} \frac{b}{\sin \theta} a^3 \end{aligned} \tag{60}$$

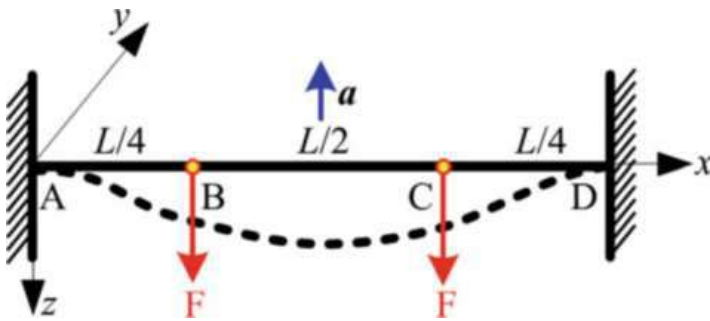
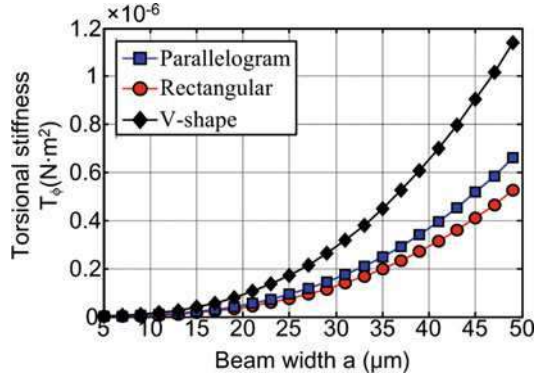
Given that  $b = 240 \mu\text{m}$  and  $G = 64.1 \text{ GPa}$ , then the torsional stiffness of three types of beams at different beam widths is calculated as shown in Fig. 10. It can be seen that the torsional stiffness of the three types of beams is near to each other when the beam width is small, which means that the accelerometers with the three types of beams have similar twist angles and static sensitivities.

**Out-of-Plane Bending Stiffness of Beams with Different Cross Sections**

When the acceleration is applied out-of-plane, the beam will have a torsional deflection as discussed above. Meanwhile, the beam will also have an out-of-plane bending deflection as shown in Fig. 11.

According the basic beam theory, displacement at point x can be calculated with

**Fig. 10** Torsional stiffness comparison of the beams with different cross sections



**Fig. 11** Schematic of the out-of-plane inertial force and bending deflection when out-of-plane acceleration is applied to the accelerometer

$$w_x = \begin{cases} \frac{FLx^2}{6EI_y} \left( \frac{9}{16} - \frac{x}{L} \right) & x \in \left[ 0, \frac{L}{4} \right] \\ \frac{FL^3}{96EI_y} \left( \frac{3x}{L} - \frac{3x^2}{L^2} - \frac{1}{4} \right) & x \in \left( \frac{L}{4}, \frac{3L}{4} \right) \\ \frac{FL(L-x)^2}{6EI_y} \left( \frac{x}{L} - \frac{7}{16} \right) & x \in \left[ \frac{3L}{4}, L \right] \end{cases} \quad (61)$$

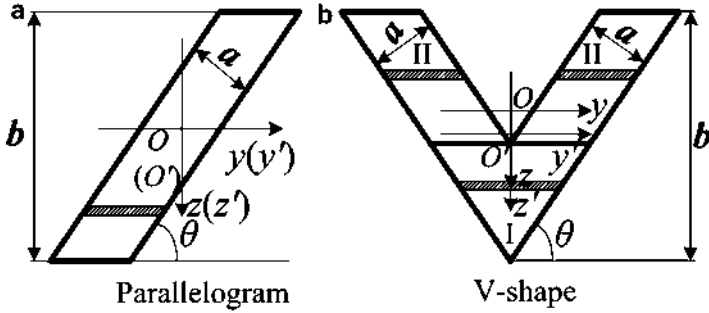
Herein, inertial force  $F$  is

$$F = (m_1 + m_2)a = (m_3 + m_4)a, \quad (62)$$

and the out-of-plane ( $z$ -axis) bending stiffness is defined as

$$B_z = EI_y, \quad (63)$$

where,  $E$  is the Young's modulus of silicon and  $I_y$  is area moment of inertia of the beam about the centroid axis ( $y$ -axis). As for the common rectangular cross section,



**Fig. 12** Coordinate system and micro-unit partition of parallelogram and V-shape cross section for out-of-plane bending stiffness calculation

$$I_y^{rec} = \frac{1}{12} ab^3. \tag{64}$$

As for the parallelogram cross section shown in Fig. 12a, the geometrical center,  $O'$ , coincides with the centroid of the parallelogram cross section,  $O$ . Then the area moment of inertia about the  $y$ -axis is

$$\begin{aligned} I_y^{par} &= I_{y'}^{par} \\ &= \int_{-\frac{b}{2}}^{\frac{b}{2}} \frac{a}{\sin \theta} z^2 dz. \\ &= \frac{1}{12} \frac{ab^3}{\sin \theta} \end{aligned} \tag{65}$$

As for the V-shape cross section shown in Fig. 12b, the geometrical center,  $O'$ , is not coincident with the centroid of the V-shape cross section,  $O$ . So the centroid's location needs to be calculated first. In coordinate system  $O'y'z'$ , the location of the centroid can be calculated as

$$\begin{aligned} z_0 = \frac{S_{y'}^I + S_{y'}^{II}}{S_V} &= \frac{\int_{-\frac{b}{2}}^{\frac{b}{2}} \left[ \frac{b}{b} + \frac{a}{\cos \theta} \frac{2z + b}{\tan \theta} \right] z dz + \int_{-\frac{b}{2}}^{\frac{b}{2}} \left[ \frac{b}{b} + \frac{a}{\cos \theta} \frac{2a}{\sin \theta} \right] z dz}{2 \left( b - \frac{a}{\cos \theta} \right) \frac{a}{\sin \theta} + \frac{a}{\cos \theta} \frac{a}{\sin \theta}} \\ &= \frac{-3ab \sec \theta + 2a^2 \sec^2 \theta}{-12b + 6a \sec \theta} \end{aligned} \tag{66}$$

Herein,  $S_{y'}$  is the static moment of the cross section about the  $y'$ -axis.  $S_V$  is the total area of the V-shape cross section. Then the area moment of inertia about the centroid axis can be calculated

$$\begin{aligned}
 I_y^V &= I_{y'}^I + I_{y'}^{II} - S_V z_O^2 = \int_{-\frac{b}{2}}^{-\frac{b}{2} + \frac{a}{\cos\theta}} \frac{2z + b}{\tan\theta} z^2 dz \\
 &+ \int_{-\frac{b}{2} + \frac{a}{\cos\theta}}^{\frac{b}{2}} \frac{2a}{\sin\theta} z^2 dz - S_V z_O^2 \\
 &= \frac{a \csc\theta}{36b - 18a \sec\theta} (6b^4 - 12ab^3 \sec\theta + 12a^2b^2 \sec^2\theta - 6a^3b \sec^3\theta + a^4 \sec^4\theta)
 \end{aligned}
 \tag{67}$$

Then out-of-plane bending stiffness of the above three types of beams can be obtained:

$$\begin{aligned}
 B_z^{rec} &= EI_y^{rec} \\
 B_z^V &= EI_y^V \\
 B_z^{par} &= EI_y^{par}.
 \end{aligned}
 \tag{68}$$

Given that  $b = 240 \mu\text{m}$  and  $E = 169 \text{ GPa}$ , then out-of-plane bending stiffness of the three types of beams at different beam widths are calculated as shown in Fig. 13. It can be seen that the out-of-plane bending stiffness of the V-shape beam is slightly larger than those of the other two types of beams, which can decrease the interference of the out-of-plane bending deflection on the preferred torsional motion of the beam.

### In-Plane Bending Stiffness of Beams with Different Cross Sections

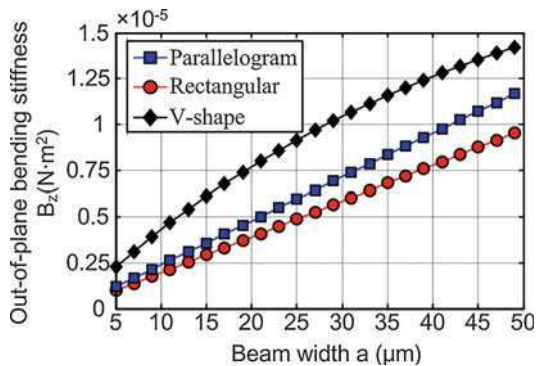
When the acceleration is applied in-plane, the beam will have an in-plane bending deflection. The displacement of the beam can also be calculated with Eq. 62 while the area moment of inertia should be  $I_z$  instead of  $I_y$ . Similarly define the in-plane bending stiffness as

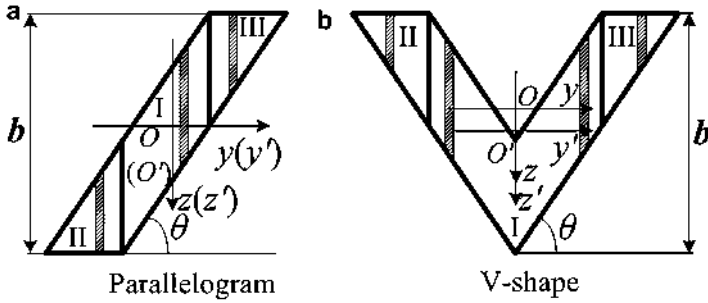
$$B_y = EI_z.
 \tag{69}$$

As for the common rectangular cross section,

$$I_z^{rec} = \frac{1}{12} ba^3.
 \tag{70}$$

**Fig. 13** Out-of-plane bending stiffness comparison of the beams with different cross sections





**Fig. 14** Coordinate system and micro-unit partition of parallelogram and V-shape cross section for in-plane bending stiffness calculation

As for the parallelogram cross section shown in Fig. 14a, the area moment of inertia about z-axis is

$$\begin{aligned}
 I_z^{par} &= I_{z'}^{par} \\
 &= I_z^I + I_z^{II} + I_z^{III} \\
 &= \frac{1}{12} (ab^3 \cot^2 \theta \csc \theta + a^3 b \csc^3 \theta).
 \end{aligned}
 \tag{71}$$

As for the V-shape cross section shown in Fig. 14b, the area moment of inertia about z-axis is

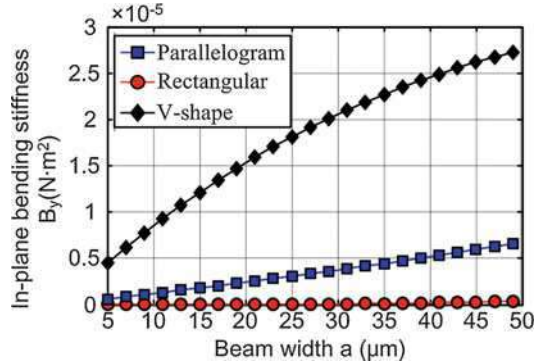
$$\begin{aligned}
 I_z^V &= I_{z'}^V \\
 &= I_z^I + I_z^{II} + I_z^{III} \\
 &= \frac{2ab^3 \cos^2 \theta}{3 \sin^3 \theta} - \frac{a^2 b^2 \cos \theta}{\sin^3 \theta} \\
 &\quad + \frac{2a^3 b}{3 \sin^3 \theta} - \frac{a^4 \sec \theta}{6 \sin^3 \theta}.
 \end{aligned}
 \tag{72}$$

Then the in-plane bending stiffness of the above three types of beams can be obtained:

$$\begin{aligned}
 B_y^{rec} &= EI_z^{rec} \\
 B_y^V &= EI_z^V \\
 B_y^{par} &= EI_z^{par}.
 \end{aligned}
 \tag{73}$$

Also given that  $b = 240 \mu\text{m}$  and  $E = 169 \text{ GPa}$ , then out-of-plane bending stiffness of three types of beams at different beam widths are calculated as shown in Fig. 15. It can be seen that the in-plane bending stiffness of the V-shape beam is much larger than those of the other two types of beams, which can decrease the interference of the in-plane bending deflection and the in-plane cross-axis errors.

**Fig. 15** In-plane bending stiffness comparison of the beams with different cross sections



From Figs. 10, 13, and 15, it can be seen that in the three types of beams with similar torsional stiffness, the V-shape beam has larger bending stiffness at either out-of-plane direction or in-plane direction, giving it better robustness for the interferential acceleration along the insensitive axes. So in this chapter, the V-shape beam is used for the sensitive torsional beam.

## Design of the Unbalanced Structure

There are two typical methods to achieve the unbalance of the masses on the two sides of the torsional beam.

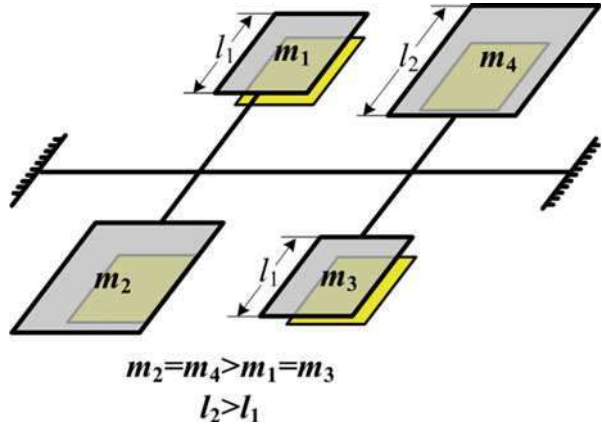
As shown in Fig. 18, the first method is making one mass smaller while the other one larger. This method is easy to achieve as no extra fabrication steps are needed to add, so it is used in many torsional accelerometer designs (Yang et al. 2015; Haris and Qu 2010). However, as derived in Eq. 53, the relative areas of the capacitors should be identical to improve the linearity of the accelerometer, so in this method the relative areas must be set according to the smaller mass to make them identical, and this will waste space under the larger mass. In addition, the asymmetric of the structure will trigger other problems (Fig. 16).

The second method is making holes in one mass as shown in Fig. 17 (Lee et al. 2005). This method is also easy to achieve, but the electrode under the holes has no effective relative area so it has the same problem as the first method.

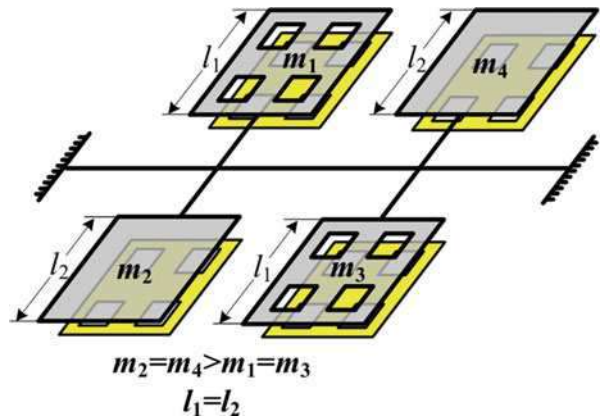
To overcome the disadvantages of the above typical methods, this chapter proposes a novel method to achieve the unbalanced structure, which is making a hollow on one mass, as shown in Fig. 18. As the bottoms of the masses are the same, they can be made to be the same size and the electrodes under them can also be set to be the same. So outside geometries of the masses and electrodes are symmetric about the torsional beam, which benefit the mechanical characteristics of the accelerometer.

As analyzed above, in the three methods of making the unbalanced masses, the mass with hollow is preferred, and as analyzed in section “[Design of the Torsional Beam](#),” of the four types of beams, the V-shape beam is preferred. The wet-etching method is more suitable and cheaper for the above structure, so it is chosen for the

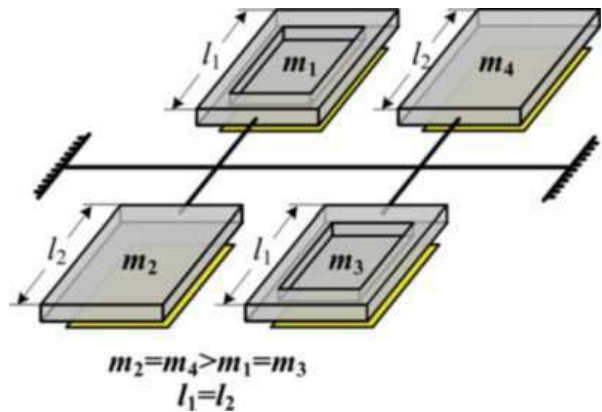
**Fig. 16** Schematic of the first method for making a pair of masses unbalanced, which is making one mass smaller while another one larger



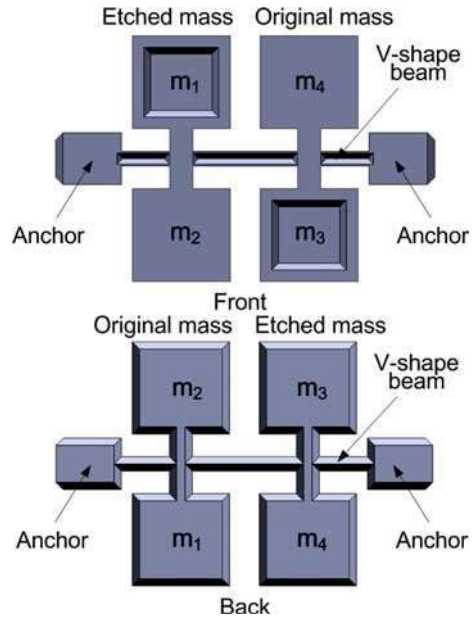
**Fig. 17** Schematic of the second method for making a pair of masses unbalanced, which is etching some holes on one mass



**Fig. 18** Schematic of the third method for making a pair of masses unbalanced, which is etching hollow on one mass while keeping another one unetched



**Fig. 19** The double differential torsional accelerometer structure without stress-released frame



fabrication method. The design model of the double differential torsional accelerometer without a stress-released frame is shown in Fig. 19. It can be seen that mass 1 and mass 3 are etched for a defined depth to form a hollow and achieve the unbalanced torsional structure. In addition, it is symmetric at the back view so the electrodes under them and the capacitors are also symmetric.

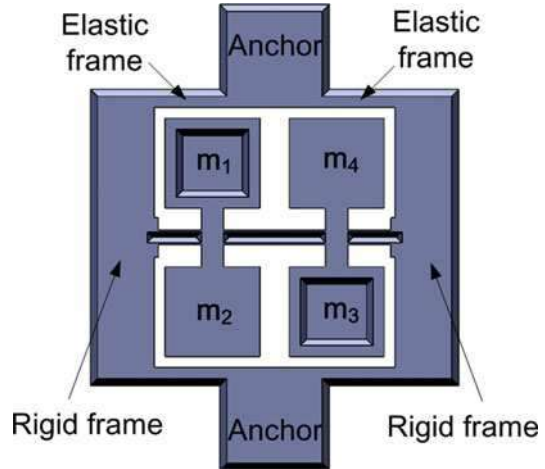
## Design of the Stress-Released Frame

The silicon structure of the accelerometer is attached to the glass substrate by anodic bonding process and then the die is packaged in the ceramic package. As material characteristics of the silicon structure, the glass substrate and the ceramic package are different; unavoidably there will be thermal stress on the attaching boundaries of them when the temperature changes. Thermal stress can be reduced by annealing and adopting compensation materials, but they are both limited to the fabrication process and the property of materials. Thus, a simple and effective stress-released structure is need for the accelerometer to decrease the temperature influence on it. Here, a stress-released frame is introduced, as shown in Fig. 20. The stress-released frame consists of four elastic frames and two relatively rigid frames. The torsional structure is suspended on the rigid frames, while the elastic frames are connected with the anchors.

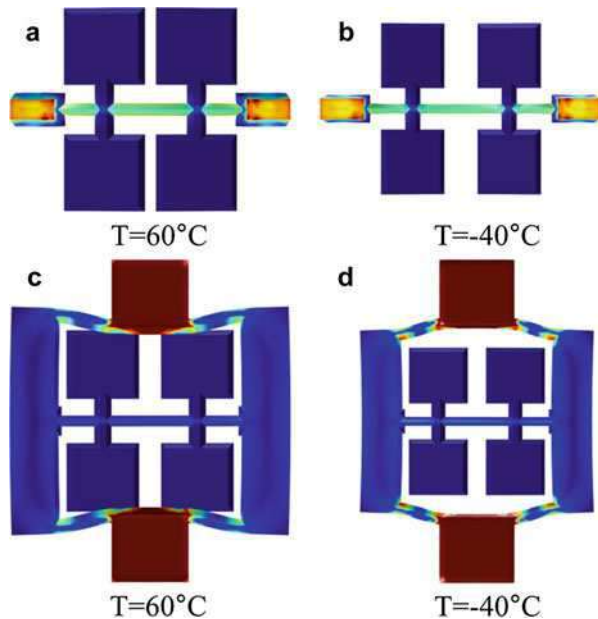
When the temperature changes, thermal stress occurs on the boundaries between the glass substrate and the silicon structure at the anchor locations. This stress then transfers to the torsional beam, which further results in the change of the beam's stiffness. As shown in Fig. 21, in the structure without a stress-released frame, thermal



**Fig. 20** The double differential torsional accelerometer structure with stress-released frame

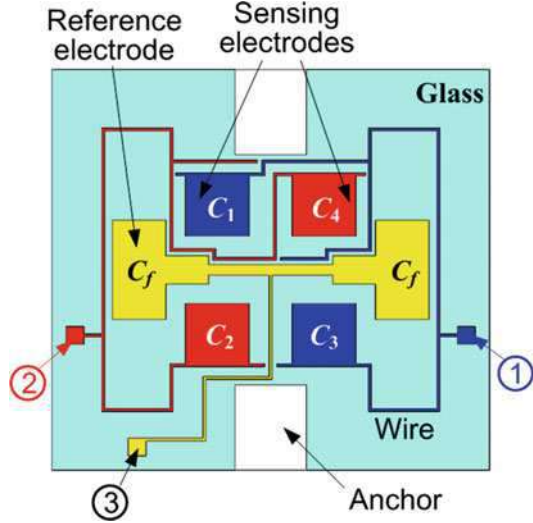


**Fig. 21** Thermal stress distribution of the accelerometer without and with the stress-released frame at the temperature 60 °C and -40 °C respectively



stress will transfer to the torsional beam easily and has large influence on the performance of the accelerometer. In contrast, in the structure with the stress-released frame, thermal stress will transfer to the elastic frame, but less to the torsional beam because of the large difference of the stiffness between the elastic frames and the relative rigid frames. It demonstrates that by using the stress-released frame, the influence of the thermal stress on core torsional beam can be greatly decreased and the temperature robustness of the accelerometer can be greatly improved.

**Fig. 22** Electrode design of the designed accelerometer



## Electrode Design

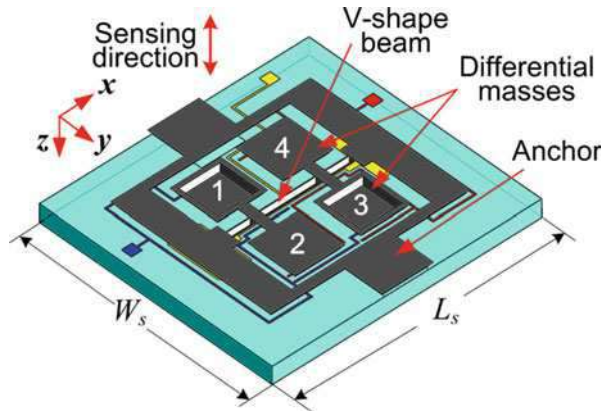
The silicon structure is bonded on the glass substrate on which the aluminum electrodes are made. As shown in Fig. 22, the electrode is separated into three parts: two sensing electrodes and one reference electrode. Sensing electrodes are located directly under the masses in the silicon structure to form the sensing capacitors  $C_1 \sim C_4$ . To ensure the relative areas are stable when the silicon masses have deformation, the areas of electrodes are smaller than those of the silicon masses.

As shown in Eq. 40,  $C_2$  and  $C_4$ ,  $C_1$  and  $C_3$  are added respectively and then the difference is calculated. In this accelerometer, the capacitances' sum is achieved by the direct connection of the electrodes with the connecting wires on the glass substrate, and the added signals are led to the pads. It can be seen that the connecting wires have the same length and they are symmetrically located under the silicon structure to ensure identical parasitic capacitances. The reference capacitor is set under the frame to further improve the temperature robustness of the readout circuit (the details will be introduced in section “[Readout Circuit of the Double Differential Capacitive Torsional Accelerometer](#)”).

In the sections above, the design and analysis of all the parts in the accelerometer have been discussed. The total sensing element of the accelerometer, which is composed of the silicon structure and the glass substrate, is shown in Fig. 23. Sensitive direction is out-of-plane and coincident with  $z$ -axis.

According to the actual requirements and application, partial geometrical parameters of the accelerometer are designed first, as shown in Table 1. The (100) silicon wafer with 240  $\mu\text{m}$  thickness is used for the silicon structure fabrication. The masses with hollows are etched for 220  $\mu\text{m}$  with 20  $\mu\text{m}$  left. The length of the torsional beam is set as 6 mm. The size of the sensing element is designed to be  $11 \times 11$  mm and the capacitors' relative areas are set to be  $2 \times 2$  mm. The capacitors' original gap of the

**Fig. 23** Design model of the sensing element of the double differential torsional accelerometer



**Table 1** Partial geometrical parameters of the new accelerometer

Description	Symbol	Value
Wafer thickness	$t$	240 $\mu\text{m}$
Hollow depth	$h$	220 $\mu\text{m}$
Length of the torsional beam	$L$	6 mm
Length of the sensing element	$L_s$	11 mm
Width of the sensing element	$W_s$	11 mm
Distance between the beam and the mass's center	$l_m$	0.87 mm
Relative area of each sensing capacitor	$A_e$	4.0 mm <sup>2</sup>
Original gap of the capacitor	$d_0$	8.0 $\mu\text{m}$
Theoretical original capacitance of each sensing capacitor	$C_0$	4.43 pF

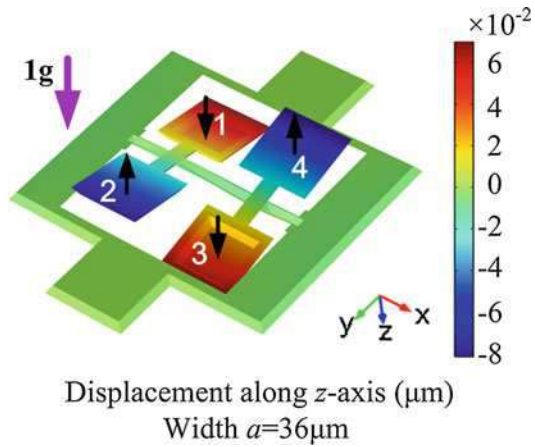
accelerometer is designed as 8  $\mu\text{m}$  and then the theoretical original capacitance of each sensing capacitor can be obtained to be 4.43 pF.

After confirming the above geometrical parameters, analysis including mathematical calculation and finite element method (FEM) simulation are made for the design of the left parameters such as the sensitivity, natural frequency, and so on. The results are shown in the following sections.

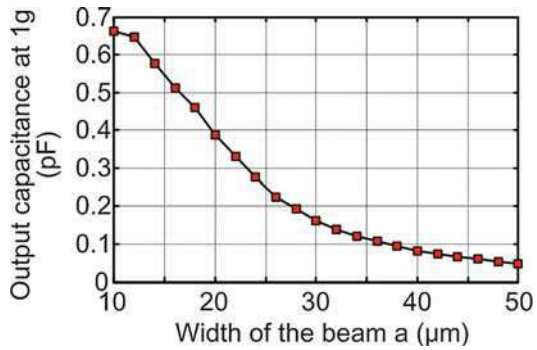
### Analysis of the Static Mechanical Sensitivity

It has been derived that the mechanical sensitivity of the accelerometer can be calculated with Eq. 54. But to simplify the analysis process, this chapter uses the FEM simulation for the sensitivity analysis. As the length of the torsional beam has been confirmed, the static mechanical sensitivity is only relative to the width of the beam, whose symbol is “ $a$ ” as shown in Fig. 9. In the simulation, the anchors are fixed and 1 g acceleration is applied along the out-of-plane direction. Then the deformations of the masses’ vertexes are measured, the twist angles are calculated

**Fig. 24** Deformation along z-axis of the accelerometer with 0.036 mm beam width at 1 g acceleration



**Fig. 25** The static mechanical sensitivities of the accelerometers with different beam widths



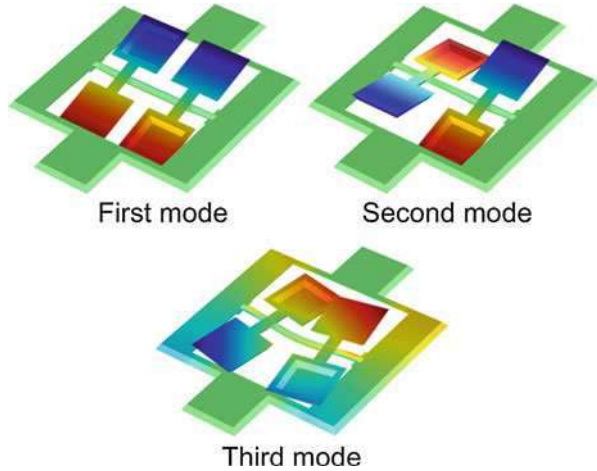
and further the mechanical sensitivities are obtained. Deformation along the z-axis of the accelerometer with 36  $\mu\text{m}$  beam width is shown in Fig. 24. The static mechanical sensitivities of the accelerometers with different beam width are shown in Fig. 25.

### Mode Analysis

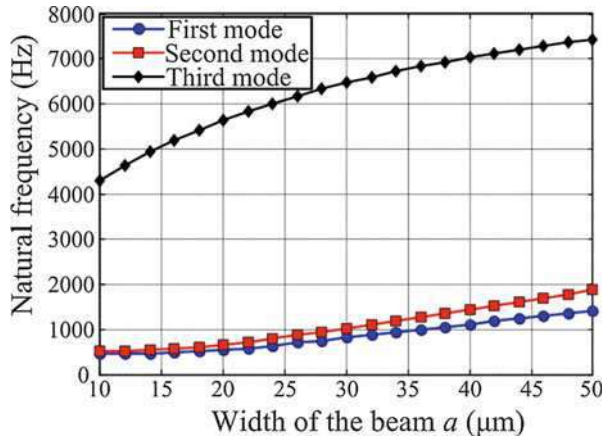
The natural frequency of the sensing mode is an important factor that influences the dynamic sensitivity, noise and bandwidth of the accelerometer. For this accelerometer, the first three modes are shown in Fig. 26 and the second mode is the working mode.

Same with the static mechanical sensitivity, the natural frequencies of the three modes are also mainly determined by the width of the torsional beam,  $a$ . By FEM simulation, the natural frequencies of the accelerometers with different beam widths

**Fig. 26** The first three modes of the double differential torsional accelerometer with V-shape beam. Deformation shown in the figure is the deformation along z-axis

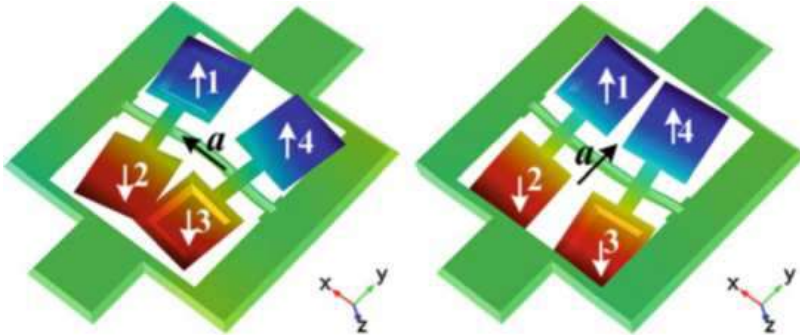


**Fig. 27** Natural frequencies of the first three modes of the first three modes of the double differential torsional accelerometer with V-shape beam at different beam widths



are obtained as shown in Fig. 27. It can be seen that the first and second modes are near each other when the beam width is small, and when the width increases, the frequency split also increases. In accelerometers, we prefer the frequency split between the working mode and the other modes to be as large as possible to decrease the interference to each other.

In addition, an increase in natural frequency results in a decrease in sensitivity and an increase in noise, thus a lower frequency is preferred. However, the bandwidth will also be lower with a decreased frequency, so a suitable resonant frequency must be chosen according to the special application and requirements for sensitivity, noise, and bandwidth. The accelerometer designed here is used in navigation which has no strict requirement for bandwidth, so the resonant frequency is designed to be about 1500 Hz and the beam width about 40  $\mu\text{m}$ .



**Fig. 28** FEM simulation when the acceleration is applied along the insensitive direction, the positive  $x$  and  $y$  direction, at room temperature. Deformation shown in the figure is the deformation along  $z$ -axis

### Analysis of the Cross-Axis Error

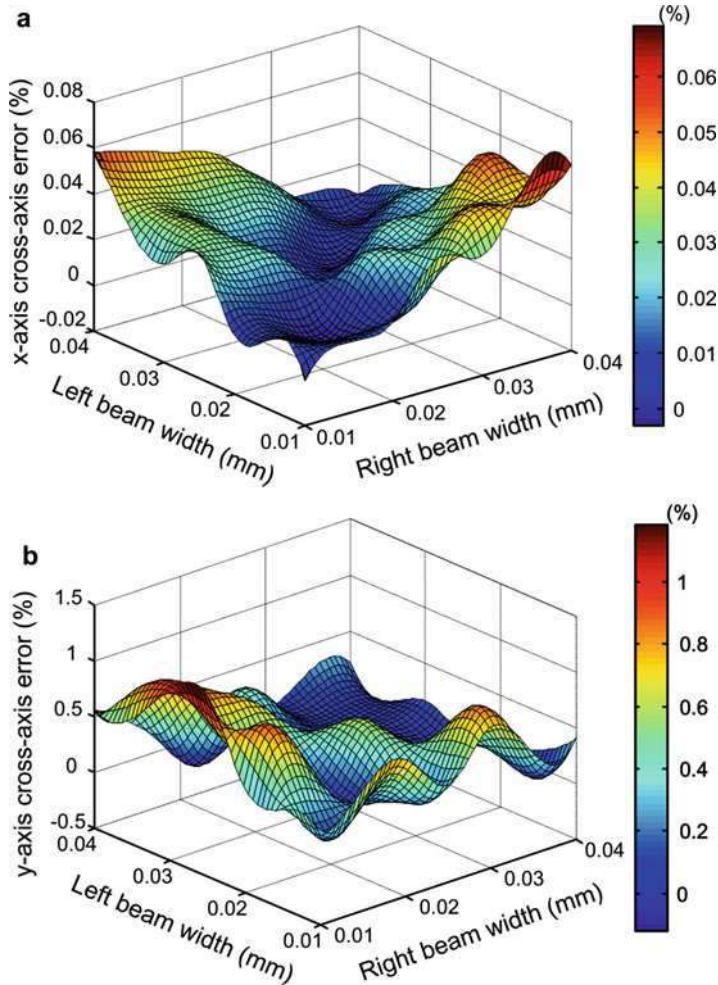
As analyzed in section “[Design of the Torsional Beam](#),” V-shape beam has a larger ratio of bending stiffness and torsional stiffness, improving mechanical sensitivity while decreasing cross-axis error. Ideally, the cross-axis error should be zero, but it cannot be eliminated completely due to the beams’ slight bending and unavoidable fabrication imperfections. Deformations of the accelerometer under  $x$ -axis and  $y$ -axis accelerations are shown in Fig. 28.

In the designed accelerometer with a V-shape beam, the cross-axis error is mainly related to the side walls of the beam. So it is necessary to deliberately set some imperfections to analyze their influences on the cross-axis error. Here, we use the FEM analysis for the cross-axis error by making the V-shape beam have different widths on the two sides. The  $x$ -axis cross-axis error simulation results are shown in Fig. 29a, and demonstrate that the  $x$ -axis cross-axis error has the minimum value when the left and right beam widths are the same, while the error will increase with the difference expansion of them. The  $y$ -axis cross-axis error simulation results are shown in Fig. 29b, and show that the  $y$ -axis cross-axis error has no obvious relationship with the difference of the two sides’ beam widths, but it has the minimum value at the points where the left and right beam width are both about  $35\ \mu\text{m}$ .

Comprehensively considering the static mechanical sensitivity, natural frequency, and cross-axis error, the width of the V-shape beam is chosen as  $38\ \mu\text{m}$  and the mechanical sensitivity, natural frequency,  $x$ -axis cross-axis error, and  $y$ -axis cross-axis error are  $0.0936\ \text{pF/g}$ ,  $1358\ \text{Hz}$ ,  $0.0055\%$ , and  $0.1499\%$ , respectively.

### Analysis of the Temperature Robustness

Material properties of the silicon beam is easily to be influenced by the temperature, meanwhile, the coefficients of thermal expansion of silicon and glass are different,



**Fig. 29** (a) x-axis and (b) y-axis cross-axis error simulation when the V-shape beam has different width on the two sides

resulting the temperature dependency of the device. In the previous traditional torsional accelerometers, consisting of only one pair of differential masses, the output errors caused by temperature drift cannot be declined. In contrast, the accelerometer proposed in this chapter can achieve the temperature self-calibration by the design of double differential configuration, greatly improving the temperature robustness.

As shown in Fig. 21, when the temperature change, the frame will have out-of-plane linear deformations along  $z$  direction and torsional deformation around  $x$  and  $y$  axis, which will induce the output drift in the single differential configuration accelerometers. However, in the new designed double differential



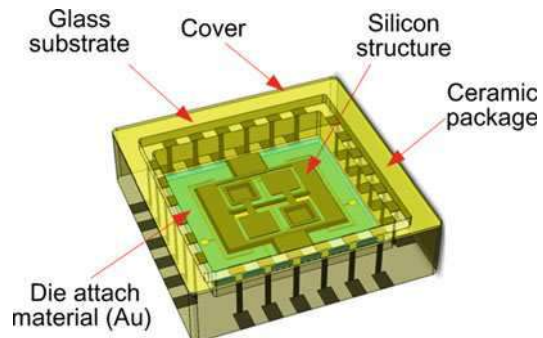
configuration accelerometer, all these errors can be reduced. For example, when the frame has an out-of-plane linear deformation along  $z$  direction, all the capacitances will be smaller or bigger simultaneously with the same degree but the double differential output calculated by Eq. 40 will keep stable. Also, when the frame has a torsional deformation around  $x$  axis, the capacitances 1 and 2 will have the same change while the capacitances 3 and 4 will have the inverse change, which also make the double differential output be stable. At last, when the frame has a torsional deformation around  $y$  axis, the capacitances 2 and 3 will have the same change while the capacitances 1 and 4 will have the inverse change, which also make the double differential output be stable. So, theoretically, the double differential structure can greatly reduce the error induced by environment changing.

Here, in this chapter, the temperature performance simulation is taken for the whole device to demonstrate the temperature robustness analyzed above. The whole device including package is shown in Fig. 30. The bonded accelerometer die, composed of a silicon structure and a glass substrate, is attached on the ceramic package bottom by an Au film and the cover is used to package the device at vacuum environment. As all the materials and connections will produce thermal stress with the temperature changing, so the temperature performance simulation should be taken for the whole system.

Here, COMSOL is used for the FEM simulation of the device at a temperature range from  $-40$  to  $140$  °C while the reference temperature is  $20$  °C. The constraint condition is shown in Fig. 31a. Point 1 is fixed completely while the other points are constrained in special directions, making the whole system free to expand. Then the deformations in this state at high and low temperature are shown in Fig. 31b. The temperature drift of the device, which is determined by the changed capacitance of each capacitor, is calculated using the differential deformation of the central points on the corresponding masses and electrodes.

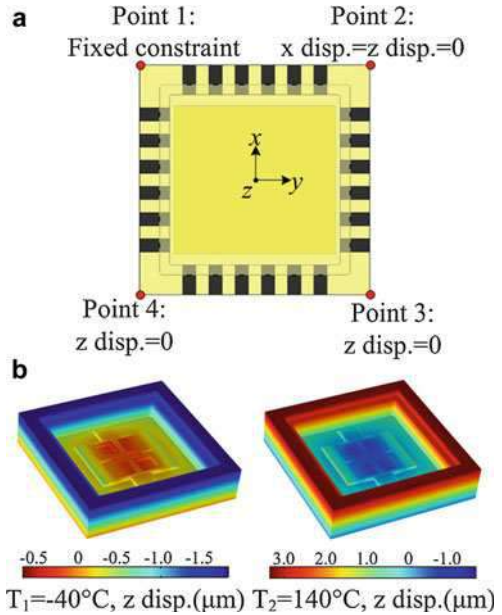
The simulation results are shown in Fig. 32. It shows that, in the full temperature range, each capacitor has about  $3500$  mg output drift and it is a large drift for a micromechanical accelerometer. However, in our design, the output is the double differential capacitance which is calculated with Eq. 40. And by the double differential method, the drift in full temperature range is reduced to  $17$  mg, which

**Fig. 30** Schematic of the packaged device

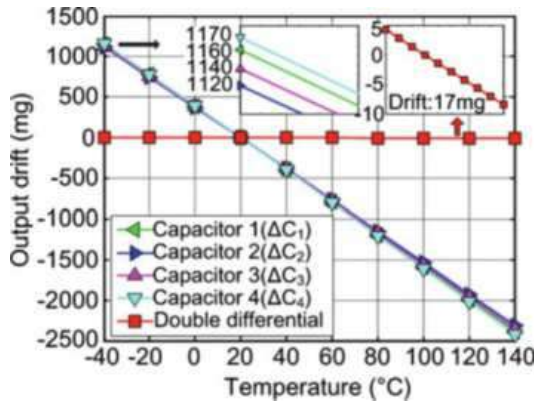




**Fig. 31** Schematic of (a) the constraint condition and (b) the deformation along z-axis in the temperature performance simulation



**Fig. 32** Temperature performance simulation results for the device at the full temperature range from  $-40$  to  $140^\circ\text{C}$

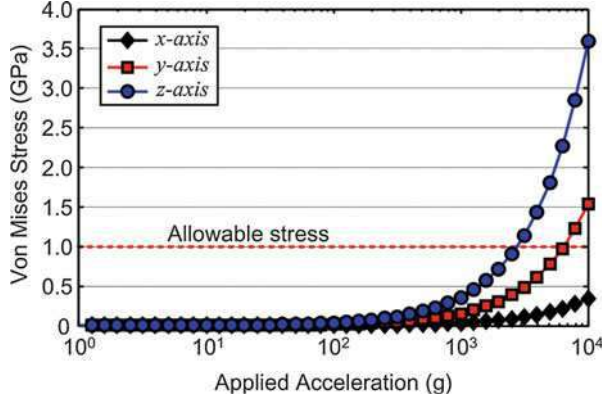


effectively demonstrates the advantage of double differential design in this chapter's accelerometer.

### Analysis of the Shock Resistance

In micromachined accelerometers, the most common failure mechanism due to shock is fracture. Crystalline silicon is a purely brittle material, which means that it deforms elastically until the maximum stress level reaches the yield strength. Even though the yield strength of Silicon is 7 GPa, to minimize the probability of fracture

**Fig. 33** Stress simulation of the designed accelerometer at the acceleration from 0 to 10,000 g



at a given shock level, a practical design criteria is to keep the maximum von-Mises stress lower than 1 GPa (Acar et al. 2008). Almost all FEA software packages have the capability to model applying acceleration on the structure and calculating the von-Mises stress distribution.

In torsional accelerometers, the highest stress concentration points are usually torsional beam connections. As for the new accelerometer design, we have made the stress simulation from 1 to 10,000 g along all three axes. The results are shown in Fig. 33, as the acceleration range is so large, a logarithmic expression is used instead of a linear one for the horizontal axis. The results show that when acceleration is applied along the z-axis, the maximum stress in this accelerometer is larger than that when the same acceleration is applied along x-axis or y axis. This means that the accelerometer can endure higher acceleration along x-axis and y-axis. Meanwhile, the results show that the accelerometer can endure about 3000 g acceleration along z-axis.

## Fabrication Technique

This section introduces the fabrication process of the novel double differential capacitive torsional accelerometer, including the silicon structure fabrication, the glass substrate fabrication, and the trimming and packaging technique.

### Fabrication of the Silicon Structure

For the designed accelerometer, the supporting V-shape beam with dimensions of microns is of great importance, since its quality plays a crucial role in the accelerometer's performance improvement. The silicon anisotropic wet chemical etching and its self-stop phenomenon have been known for a long time. In this chapter, the self-stop phenomenon is used for the fabrication of the V-shape beam.

In addition, the masses with different depths are difficult to achieve with traditional fabrication techniques. A method based on the combination of wet etching and the after thermal oxidation (ATO) technique is proposed to fabricate structures with normative polygon cross section (PCS) shapes in a single crystal silicon substrate. This method can also be used for the fabrication of the masses with different depths (Xiao et al. 2013). However, this method requires thermal oxidization in the fabrication process, making the fabrication more complex. Here, a novel method based on the pre-buried mask technique is used to fabricate masses with different depths.

The fabrication process, as shown in Fig. 34, mainly includes: (a) the starting material, a 240  $\mu\text{m}$  thick (100) silicon wafer with 400 nm silicon dioxide on the both sides; (b) photoresist spincoating on both sides and photomask on the top side with the first mask; (c) etch the silicon dioxide layers for half of the thickness, about 200 nm to form the pre-buried layer pattern; (d) remove all the photoresist; (e) photoresist spincoating on both sides for the second time and photomask on the bottom side with the second mask; (f) etch the silicon dioxide for the second time for 400 nm to reveal out the silicon plate and move; (g) etch the mechanical structure in 25% TMAH and 3% oxidizer at 80 °C for 20  $\mu\text{m}$ ; (h) etch the silicon dioxide for the third time for 200 nm to reveal out the per-buried pattern; (i) etch the silicon from both sides at the same time for another 20  $\mu\text{m}$  to etch through the silicon, now, the hexagon cantilever is formed; (j) scour off the silicon dioxide to complete the fabrication of the silicon structure.

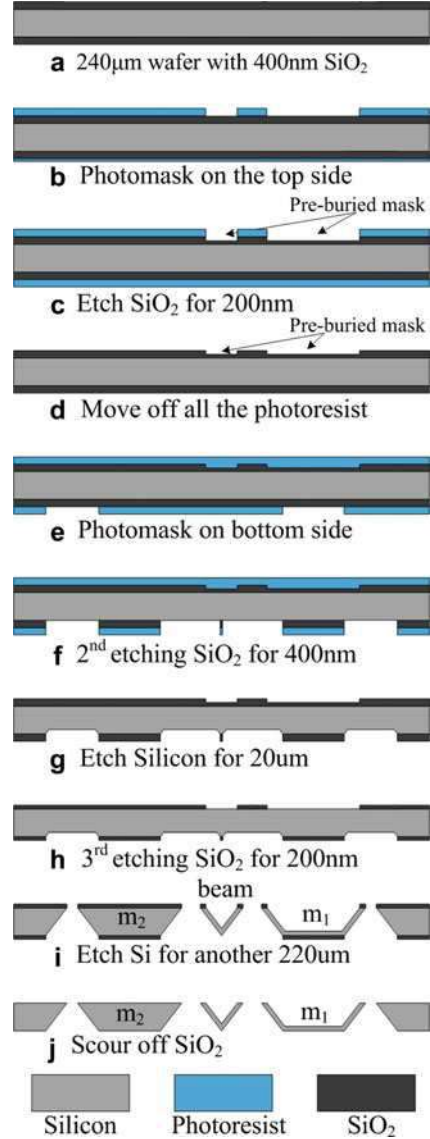
## Fabrication of the Glass Substrate and Anodic Bonding

Pyrex7740 is chosen as the glass substrate for its similar characteristics with single crystal silicon. The fabrication technology for pyrex glass have been widely researched (Tay et al. 2006; Iliescu et al. 2008). So it is briefly introduced here. Process of the glass substrate is shown in Fig. 35, it mainly includes: (a) the starting material, a 500  $\mu\text{m}$  thick Pyrex7740 glass wafer; (b) Cr-Au film coating and lithographing to pattern the mask; (c) glass wafer etching for 8  $\mu\text{m}$  and Cr-Au film is removed, the cavity depth determines the electrode gap; (d) sputtering of 500 nm aluminum and lithography to pattern the electrodes, wires, and bonding pads; (e) the silicon wafer and glass substrate are connected together with anodic bonding method.

## Trimming and Packaging

After anodic bonding, laser trimming is used to divide the wafer into separate dies. There are 16 chips on the 4-in. silicon wafer. Optical and SEM images of the wafer and microstructures are shown in Fig. 36. Then the dies are then packaged in the vacuum ceramic package and attached on the printed circuit board (PCB). Ceramic packages provide excellent mechanical and electrical characteristics. The most important advantage of ceramic packages is that the die is placed in a cavity which

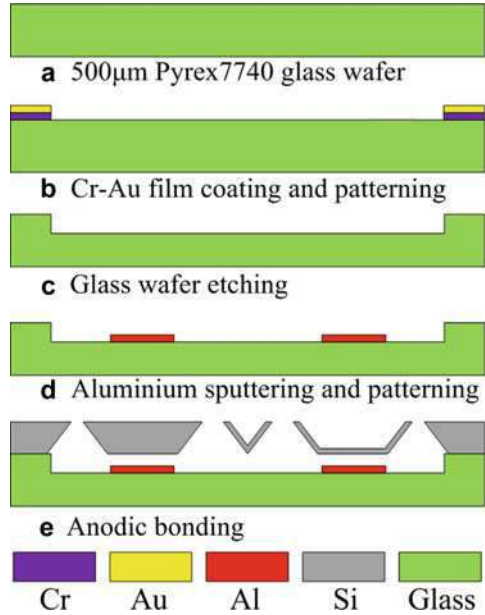
**Fig. 34** Fabrication process of the silicon structure of the accelerometer



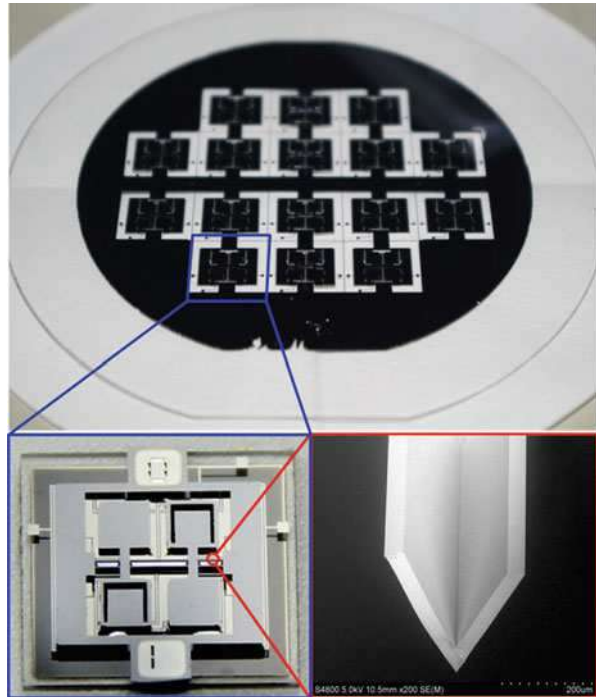
is hermetically sealed, eliminating packaging stresses and hermeticity problems. Ceramic packages are typically more expensive but more robust.

After fabrication, the static capacitances of the 16 accelerometer chips on one silicon wafer are measured to evaluate the identity of the fabrication. The results are shown in Fig. 37. It can be obtained that average capacitance and standard deviation of the sensing capacitors  $C_1 + C_3$  are 8.32 pF and 0.51 pF, respectively. Average capacitance and standard deviation of the sensing capacitors  $C_2 + C_4$  are 8.67 pF and

**Fig. 35** Fabrication process of the glass substrate and anodic bonding of the accelerometer



**Fig. 36** The optical image of the silicon wafer and microstructures and SEM image of the V-shape beam's cross section



0.36 pF, respectively. Average capacitance and standard deviation of the reference capacitor are 9.98 pF and 0.64 pF, respectively. It demonstrates that the fabrication method proposed in this chapter has a good identity.

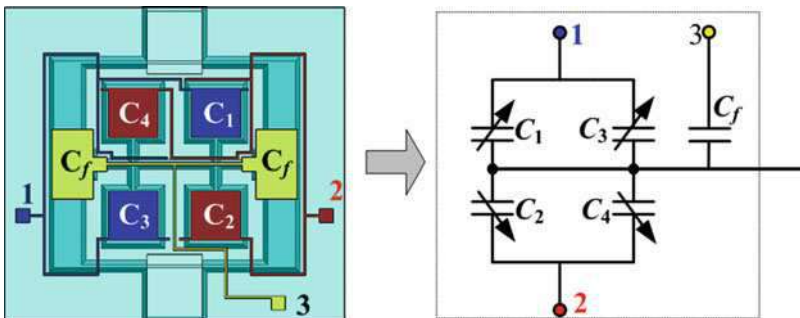
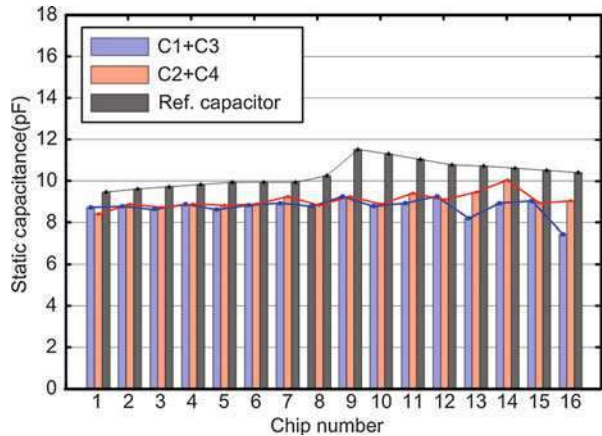
### Readout Circuit of the Double Differential Capacitive Torsional Accelerometer

This section firstly introduces the readout circuit of the traditional capacitive accelerometer, and then proposes the novel readout circuit with integrated reference capacitor to improve the temperature robust of the accelerometer.

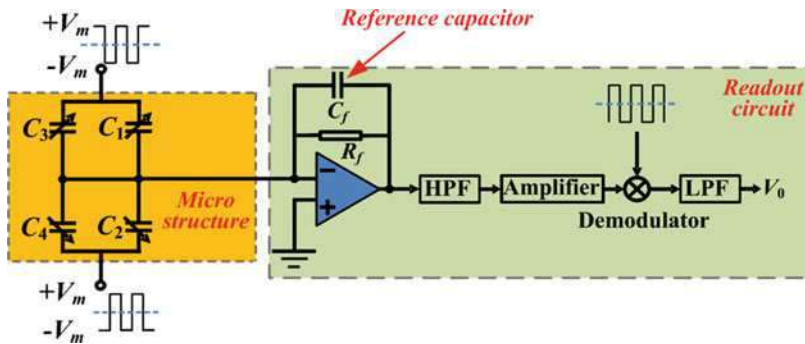
### Electrical Model of the Accelerometer

The design model of the double differential accelerometer and its electrical model are shown in Fig. 38. The electrode of  $C_1$  and  $C_3$  are wired together to achieve the

**Fig. 37** Static capacitance measurement for the 16 chips on one wafer



**Fig. 38** Design model of the accelerometer and schematic of its electrical model



**Fig. 39** Schematic of the traditional readout circuit, in which the reference capacitor is located on the PCB or ASIC

parallel connection and the signal is lead out at pad 1. Similarly, the electrodes of  $C_2$  and  $C_4$  are wired together to achieve the parallel connection and the signal is lead out at pad 2. The reference capacitor signal,  $C_f$ , is lead out at pad 3. The silicon structure can be seen as the communal pad of the four capacitors. It is directly wired out at the anodic bonding area.

### Traditional Capacitance Readout Circuit

A schematic of the traditional readout circuit is shown in Fig. 39. It mainly consists of the sensing element, charge amplifier, high pass filter, second amplifier, demodulator, and low pass filter. The silicon structure, as the communal pad of the four capacitors, is directly connected to the charge amplifier. A square wave is applied to the connecting point of  $C_1$  and  $C_3$  while an anti-phase square wave is applied to the connecting point of  $C_2$  and  $C_4$ , achieving the double differential detection as described in Eq. 40.

The charge amplifier is the most important part of the readout circuit as it converts the capacitance into voltage, it can be expressed as (Bao 2005)

$$V_0 \approx \frac{-(C_u - C_d)V_R}{C_f} \approx \frac{-\Delta C}{C_f} V_m \sin \omega t. \tag{74}$$

Herein,  $C_f$  is the reference capacitance. It can be seen that the output voltage is linearly relative with the differential capacitance and the amplification factor can be increased by decreasing the value of the reference capacitance  $C_f$ .

The following high pass filter is used for removing the interfering signals whose frequencies are lower than the modulating signal. Afterwards, a demodulator and another amplifier are applied to obtain the DC voltage that is proportional to the capacitance. The low pass filter is used for removing the interfering signals whose frequencies are higher than the bandwidth of the filter. The overall sensitivity of the readout circuit is given by

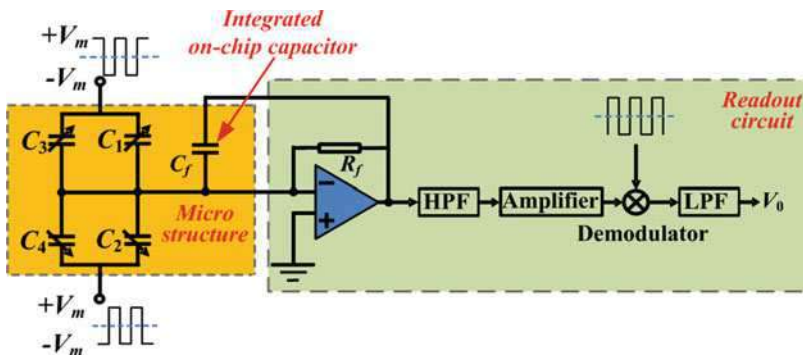
$$\begin{aligned}
 V_o &= kV_m \frac{(C_1 + C_3) - (C_2 + C_4)}{C_f} \\
 &= kV_m \frac{2 \left( \frac{\varepsilon A_0}{d_0 - \Delta x} - \frac{\varepsilon A_0}{d_0 + \Delta x} \right)}{C_f} \\
 &\approx 4kV_m \varepsilon A_0 \frac{\Delta x}{d_0^2 C_f}
 \end{aligned} \tag{75}$$

Herein,  $\Delta x$  is the deformation of each mass block,  $d_0$  is the original gap,  $A_0$  is the plate area of each capacitor,  $k$  is the gain of the signal amplifier, and  $\varepsilon$  is the permittivity of the gas in the gap.

### Readout Circuit with Integrated Reference Capacitor

Temperature dependency of the accelerometer not only results from the sensing chip but also results from the readout circuit. In conversional sensors, the reference capacitor of the amplifier,  $C_f$ , is the main factor that influences the capacitive detecting sensitivity, and is usually fabricated on ASIC or PCB. As the temperature coefficient of the outer reference capacitance and the inner sensing capacitances are different, temperature drift will greatly influence the sensitivity of the readout circuit. In this chapter, the temperature-dependent reference capacitor is shifted into the microstructure as shown in Fig. 38, and the schematic of the novel proposed detecting circuit is shown in Fig. 40.

The on-chip reference capacitor locates under the suspending frame, insuring gaps of the reference capacitors and that of the sensing capacitors are identical. Thus, the output voltage can be calculated with the equation:



**Fig. 40** Schematic of the novel proposed detecting circuit, in which the reference capacitor is shifted into the microstructure



$$\begin{aligned}
 V_o &= kV_m \frac{(C_1 + C_3) - (C_2 + C_4)}{C_f} \\
 &\approx 4kV_m \frac{\frac{\epsilon A_0 \Delta x}{d_0^2}}{\frac{\epsilon A_f}{d_0}} = 4kV_m \frac{A_0 \Delta x}{A_f d_0}.
 \end{aligned}
 \tag{76}$$

Where,  $A_f$  is the plate area of the on-chip integrated reference capacitor. It can be seen that the reference capacitor  $C_f$  in Eq. 75, which is easy to be influenced by temperature, is replaced with the plate area of reference capacitor,  $A_f$ , in Eq. 76. The factor  $A_f$  will be also influenced by the temperature because a nonzero temperature coefficient cannot be avoided. However, the change of  $A_f$  has a same coefficient with that of  $A_0$ , resulting that the ratio of  $A_0$  and  $A_f$  keeps constant. Thus, the temperature robustness of this micro accelerometer is expected to be improved by using integrated reference configuration.

Disadvantage of the design is that the modulation signal will have some influence on the mechanical structure, resulting errors of the reference capacitance. However, it can be learnt from FEM simulation that the maximum deformation induced by the electrostatic force ( $7.26 \times 10^{-6}$  N) is  $3.07 \times 10^{-4}$   $\mu\text{m}$ , which is only 0.0038% of the capacity gap (8  $\mu\text{m}$ ). In addition, resonant frequencies of all the modals relative with the frame are between 7 and 15 kHz, which is far lower than frequency of the electrostatic force (500 kHz). Therefore, the influence of the modulation signal is very small and can be ignored. The completed device with readout circuit is shown in Fig. 41.



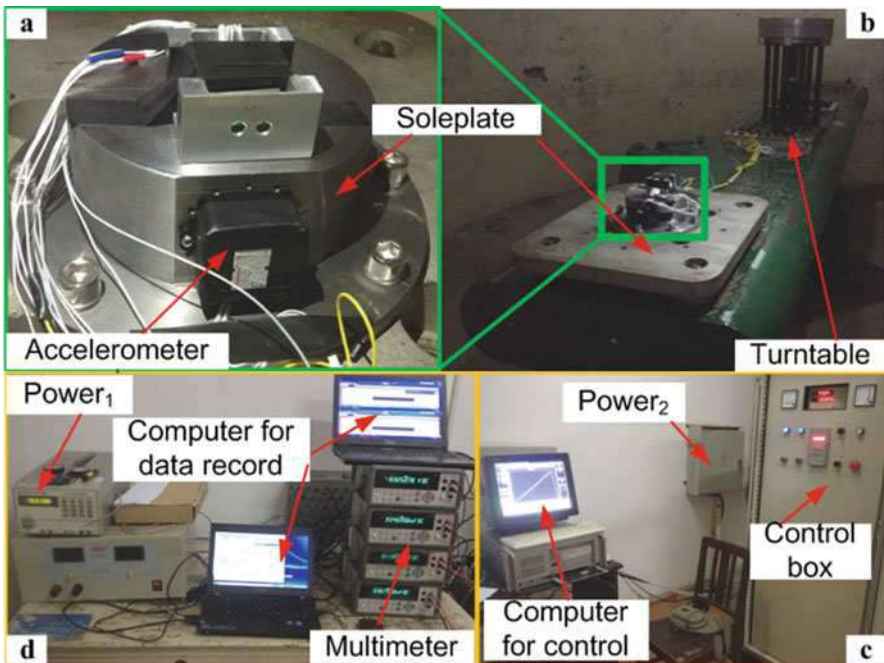
**Fig. 41** Wire bonding of the sensing element in the ceramic package and the completed device with readout circuit

## Experiment and Results Discussion

In this section, performances of the fabricated accelerometer prototypes are tested. The performances are concluded as following: scale factor is about 140 mV/g, nonlinearity is about 0.22%,  $x$ -axis cross-axis error is 0.04%,  $y$ -axis cross-axis error is 0.69%, quality factor is about 28.02, resonant frequency is about 1,500 Hz, bandwidth is 57 Hz, Allan deviation of the 1 g output in 1 h is 8.7  $\mu$ g, temperature sensitivity of the scale factor is 49.9 ppm/ $^{\circ}$ C, offset temperature coefficient of this accelerometer is under 0.22 mg/ $^{\circ}$ C, and Allan deviation of the output in the full temperature range is 121.8  $\mu$ g.

### Test of the Scale Factor

Scale factor is the ratio of a change in output to a change in the input intended to be measured, and evaluated as the slope of the least squares straight line fit to input-output data. The test instruments are composed of the data collection devices, the data-recording computers, power supply, control box for the centrifugal turntable, as shown in Fig. 42. The fabricated accelerometer with PCB is fixed on the centrifuge as the sensitive axis coincident with the direction of centrifugal acceleration.



**Fig. 42** The test instruments of scale factor and nonlinearity in the full range from  $-15$  to  $+15$  g

When the centrifuge rotates, the centrifugal acceleration is applied on the accelerometer, and the acceleration improves with the increase of the rotating speed. The acceleration applied on the device is increased by steps of 1 g from +1 to +15 g and then the accelerometer is fixed in reverse to get the opposite acceleration from -1 to -15 g. At each acceleration step, 60 output voltage sample points are taken and averaged. Then the 31 points are fitted by MATLAB and the slope is the scale factor of accelerometer, 140 mV/g, as shown in Fig. 43.

### Test of the Nonlinearity

Nonlinearity is the systematic deviation from the straight line that defines the nominal input-output relationship. From the testing results shown in Fig. 43, the nonlinearity of the accelerometer can also be calculated as

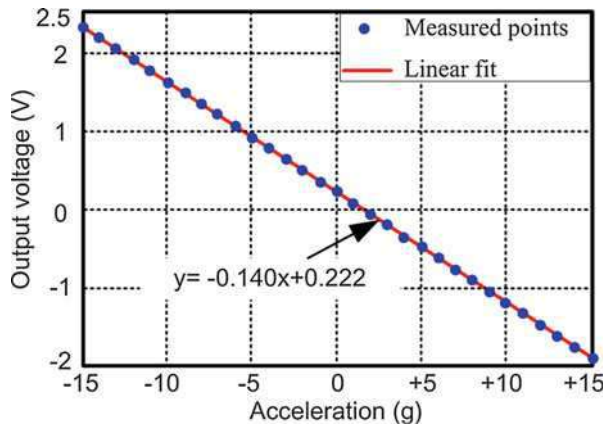
$$K = \frac{U_j^* - U_j}{|U_{\max} - U_{\min}|_{\max}} \tag{77}$$

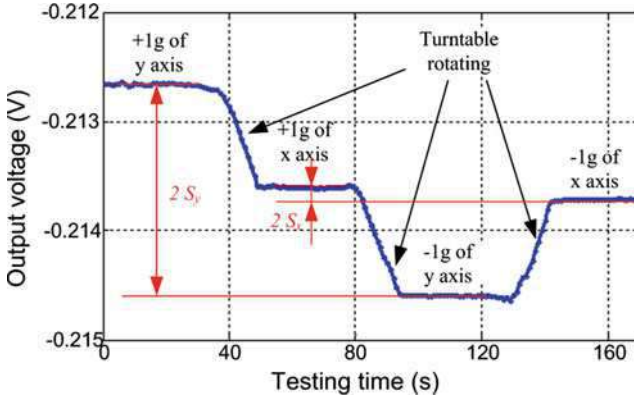
Herein,  $U_j$  is the average value of the sample points at each acceleration,  $U_j^*$  is the corresponding points of each direction on the fitting lines,  $U_{\max}$  and  $U_{\min}$  are the maximum and minimum voltage of the output, respectively. Nonlinearity of this accelerometer is calculated as 0.22%.

### Test of Cross-Axis Error

Cross-axis error is the ratio of the accelerometer outputs when the accelerations are applied along insensitive and sensitive directions. Cross-axis error of  $x$ -axis and  $y$ -axis are measured on the turntable. The turntable rotates in steps of  $90^\circ$  with the input

**Fig. 43** Output of the accelerometer in the full range from -15 to +15 g





**Fig. 44** The testing results of cross-axis error of the designed accelerometer

acceleration on  $x$ -axis and  $y$ -axis varying from  $+1$  to  $-1$  g, then the sensitivity of  $x$ -axis and  $y$ -axis is calculated. The cross-axis error is calculated by the equation:

$$S_c = S_i/S_s. \quad (78)$$

Herein,  $S_s$  is the sensitivity of sensing axis and  $S_i$  is the sensitivity of the off-sensing axis. The data recorded is shown in Fig. 44. The  $x$ -axis and  $y$ -axis cross-axis error are 0.04% and 0.69%, respectively.

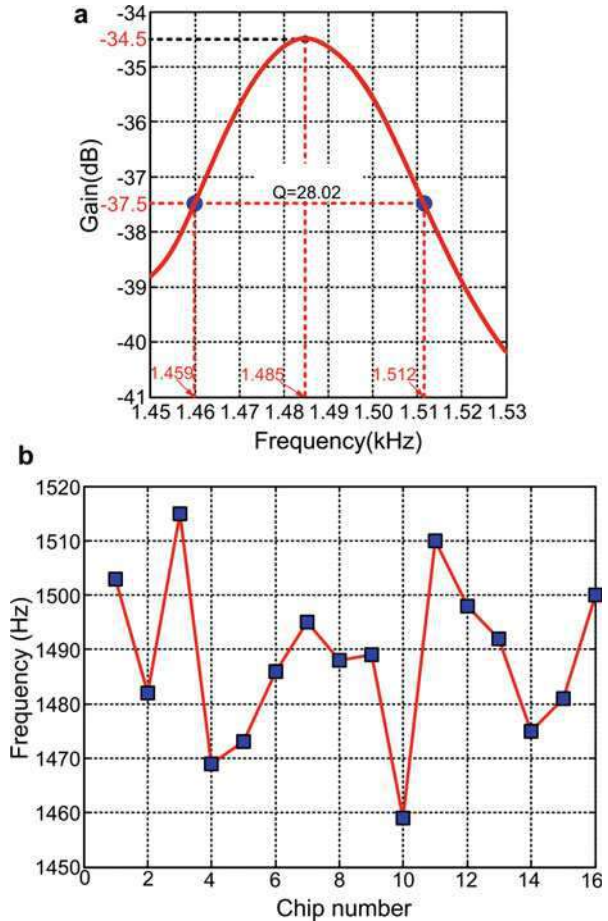
## Test of Natural Frequency and Q

Resonant frequency and quality factor of the accelerometers in vacuum environment are tested on FRA5087. One sample is tested as shown in Fig. 45a, and it shows that the resonant frequency and quality factor are 1.485 kHz and 28.02, respectively. To verify the performance of the fabrication progress, resonant frequencies of 16 chips are measured. As shown in Fig. 45b, the resonant frequencies of all chips are close to 1500 Hz: the maximum is 1515 Hz, the minimum is 1473 Hz, and the difference between these two values is 42 Hz, showing that the fabricated beams have good consistency, which verifies the excellence of the fabrication process.

## Test of the Bandwidth

This accelerometer is mainly used in navigation, which does not require high bandwidth. So, in the design of the readout circuit as shown in Fig. 40, the cut-off frequency of the low pass filter is set to 60 Hz to get lower noise and higher detecting accuracy. In addition, the bandwidth of this accelerometer is limited in 60 Hz, although it has a resonant frequency larger than 1000 Hz. The bandwidth of the accelerometer is tested on the vibrating table. The sine acceleration amplitude is set

**Fig. 45** (a) The resonant frequency and quality factor measurement and (b) resonant frequencies of 16 chips

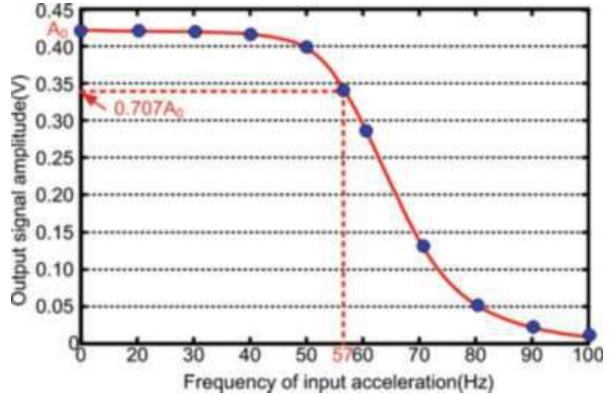


to 3 g and the frequencies are swept from 0 to 100 Hz with a step of 10 Hz. By measuring the output signal amplitude at each frequency point, the frequency response can be obtained as shown in Fig. 46. It can be seen that the bandwidth of this differential torsional micro-accelerometer is about 57 Hz. For further applications of wider bandwidth, cut-off frequency of the low pass filter in the readout circuit should be adjusted.

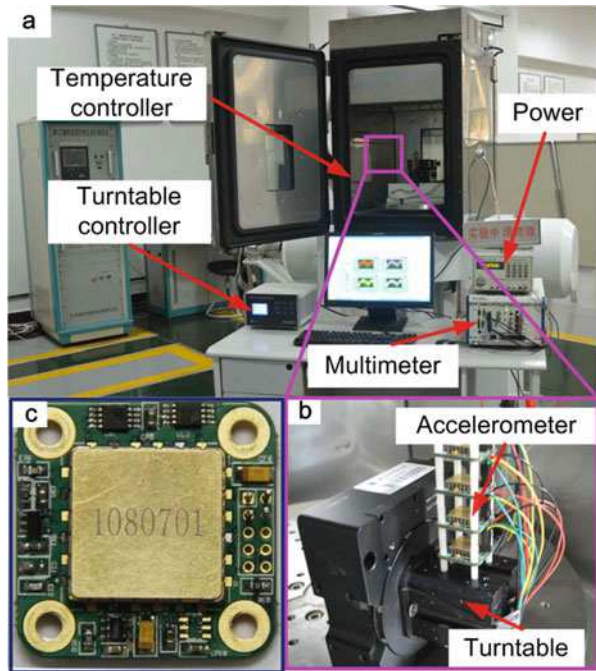
### Test of Bias Stability

Bias stability testing instruments are shown in Fig. 47. It includes a temperature control box, a highly precise turntable, a voltage supply, a data acquisition system including the acquisition card and a computer. The temperature control box is used for temperature performance testing, while the turntable, which is

**Fig. 46** Bandwidth test of the designed accelerometer



**Fig. 47** (a) Testing instruments for stability, (b) turntable with accelerometers, and (c) prototype of the accelerometer device



placed vertically, can be used to adjust the gravity acceleration on the device in a range of  $\pm 1$  g. The zero bias stability performance is measured at room temperature for 60 min after supplying power for 20 min, as shown in Fig. 48a. The accelerometer is applied with 0 g acceleration. Sampling frequency is set at 1 Hz, getting 4800 sample points, and the bias stability which is calculated with the deviation of the last 3600 sample points is 0.11 mg. The Allan deviation for the last 3600 s is 8.7  $\mu$ g, as shown in Fig. 48b.

### Noise of the Accelerometer

As for the double differential accelerometer shown in Fig. 19, the limiting mechanical noise occurs at the condition when the Brownian force applied on mass 1 and 3 are along the positive z-axis meanwhile the Brownian force applied on mass 2 and 4 are along the negative z-axis. The spectral density of the Brownian force is

$$F_{Brownian} = \sqrt{\frac{4k_B T m \omega}{Q}} \left( N / \sqrt{Hz} \right) \tag{80}$$

Herein,  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature,  $m$  is the proof mass,  $\omega$  is the resonant angular frequency, and  $Q$  is the quality factor. Under the limiting mechanical noise condition, the torque caused by Brownian force is

$$M_{Brownian} = F_{Brownian\_m1} \cdot l_m + F_{Brownian\_m2} \cdot l_m \tag{81}$$

Where,  $l_m$  is the distance between the beam and the mass’s center. Then the noise-equivalent acceleration is expressed as

$$\begin{aligned} a_{Brownian} &= \frac{M_{Brownian}}{9.8(m_2 - m_1)} \\ &= \frac{\sqrt{m_1} + \sqrt{m_2}}{9.8(m_2 - m_1)} \sqrt{\frac{4k_B T \omega_0}{Q}} \left( g / \sqrt{Hz} \right) \end{aligned} \tag{82}$$

And the mechanical noise of the accelerometer in this chapter can be calculated as 0.099  $\mu g / \sqrt{Hz}$ . The electronic noise can be calculated according to the noise floor of the electronic devices and scale factor of the circuit. In this accelerometer, electronic noise is calculated as 3.3  $\mu g / \sqrt{Hz}$ . So the overall noise-equivalent acceleration is calculated as

$$\begin{aligned} a_{Noise} &= \sqrt{a_{Brownian}^2 + a_{Circuitry}^2} \\ &= 3.3015 \left( \mu g / \sqrt{Hz} \right) \end{aligned} \tag{83}$$

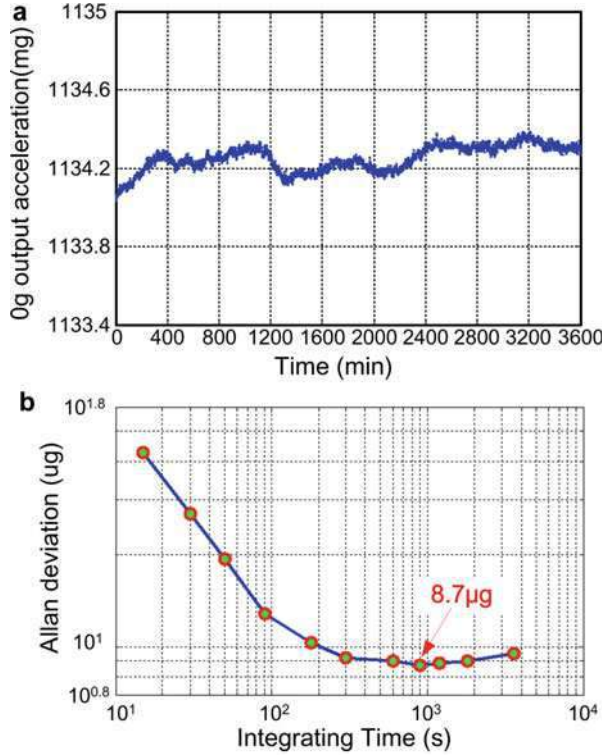
In the testing results shown in Fig. 48a, the noise of the accelerometer is measured to be 11.7  $\mu g / \sqrt{Hz}$ . It can be learnt that the actual noise is larger than the theoretical noise, which is mainly resulted by the environmental noise. By improving the package and testing condition, the tested noise floor of the accelerometer can be further reduced.

### Test of the Temperature Performance

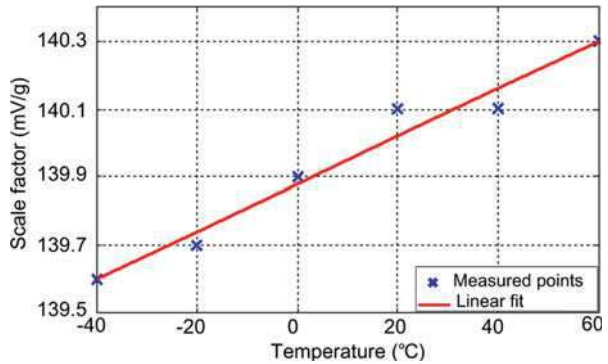
The test of temperature sensitivity of the scale factor is taken in the temperature control box. After keeping the temperature at the specific points (−40 °C,



**Fig. 48** (a) Bias stability testing data and (b) Allan deviation of the bias stability for the last 3600 s



**Fig. 49** The temperature sensitivity of the scale factor for the device

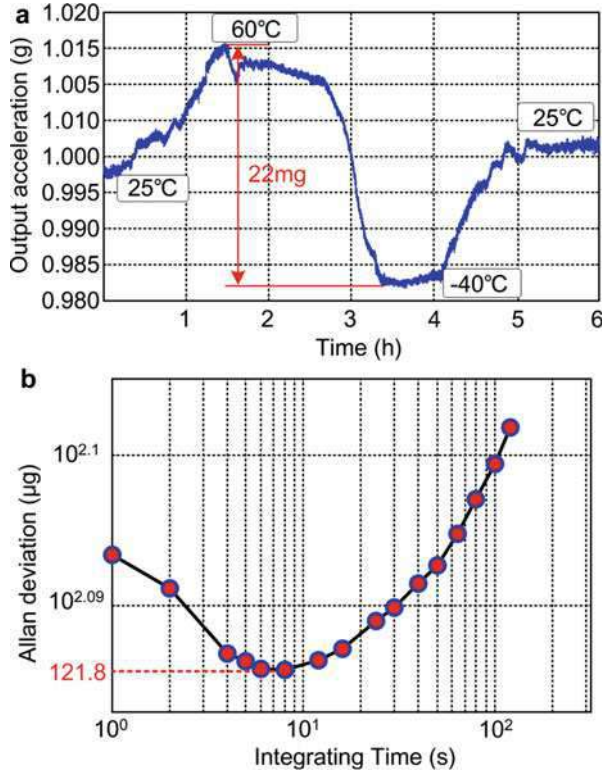


$-20\text{ }^{\circ}\text{C}$ ,  $0\text{ }^{\circ}\text{C}$ ,  $20\text{ }^{\circ}\text{C}$ ,  $40\text{ }^{\circ}\text{C}$ ,  $60\text{ }^{\circ}\text{C}$ ) for an hour, the turntable is rotated for  $180^{\circ}$  to get the scale factors at each temperature point, as shown in Fig. 49. Then, temperature sensitivity of the scale factor can be calculated with the equation:

$$S_T = SF_T/SF_0. \tag{84}$$



**Fig. 50** (a) Output drift and (b) Allan deviation of the bias stability in the full temperature range for the prototype



Herein,  $SF_T$  is coefficient of the scale factor drifting in the full temperature range, equal with the slope of the linear fit line.  $SF_0$  is the scale factor at room temperature. Temperature sensitivity of the scale factor of this accelerometer is 49.9 ppm/°C.

Output temperature drift of the accelerometer with temperature compensation is measured in the temperature control box. Firstly, the output of the accelerometer under different temperature from  $-40$  to  $60$  °C is tested and the least square fitting is used to obtain the simple linear relationship of the output and temperature. And then by real time dynamically testing the environmental temperature using the temperature sensor in the Single-chip Microcomputer (SCM) on the readout circuit, the output is dynamically compensated. In the temperature performance testing, the temperature is turned from the room temperature to the highest temperature,  $+60$  °C, then turned to the lowest temperature,  $-40$  °C, and then returning to room temperature again. The testing results are shown in Fig. 50. It can be seen that the offset temperature coefficient of this accelerometer is under  $0.22$  mg/°C, and Allan deviation of the output in the full temperature range is  $121.8$  µg. The offset drift with temperature mainly comes from the manufacture deviation, package stress, and the drift of readout circuit.

## Conclusion

The chapter presents the fundamental theory, design, fabrication, and performances test of a double differential capacitive torsional accelerometer. The topics include the brief introduction of the accelerometer, the fundamentals of micromechanical capacitive torsional accelerometer, mechanical design of the proposed accelerometer and its fabrication technique and readout circuit. At last, the completed accelerometers are tested to demonstrate the effectivity of the design. Testing results show that the accelerometer designed in this chapter has satisfactory performances. This chapter supplies a new direction for the accelerometer performance improvement, more works on the structure optimization and further research on the temperature robustness improvement will be developed in the future.

---

## References

- Acar C, Shkel A, SpringerLink EE (2008) MEMS vibratory gyroscopes: structural approaches to improve robustness. Springer London, Guildford/Boulder
- Andersson G, Hedenstierna N, Svensson P (1999) A novel silicon bulk gyroscope. *TRANSDUCERS '99, Sendai*, pp 902–905
- Bao M (2005) Analysis and design principles of MEMS devices. Elsevier Science [Imprint], San Diego
- Dong H, Jia Y, Hao Y et al (2005) A novel out-of-plane MEMS tunneling accelerometer. *Sensors Actuators A Phys* 120(2):360–364
- Elwenspoek M, Wiegink R (2001) Silicon micromachining. In: Suleman A (ed) *Smart structures: applications and related technologies*. Springer Vienna, Vienna, pp 241–264
- Fischer AC, Forsberg F, Lapisa M et al (2015) Integrating MEMS and ICs. *Microsyst Nanoeng* 1:15005
- Garraud A, Giani A, Combette P et al (2011) A dual axis CMOS micromachined convective thermal accelerometer. *Sensors Actuators A Phys* 170(1):44–50
- Ghisi A, Mariani S, Corigliano A et al (2012) Physically-based reduced order modelling of a uniaxial polysilicon MEMS accelerometer. *Sensors (Basel)* 12(10):13985–14003
- Haris M, Qu HW (2010) Fully differential CMOS-MEMS z-axis accelerometer with torsional structures and planar comb fingers. *J Micro-Nanolithography Memos* 9(1):13031
- Hiratsuka R, van Duyn DC, Otaredian T et al (1991) A novel accelerometer based on a silicon thermopile. pp 420–423
- Iliescu C, Chen B, Miao J (2008) On the wet etching of Pyrex glass. *Sens Actuators A Phys* 143(1):154–161
- Korvink JG, Paul O, Ebrary E et al (2006) MEMS: a practical guide to design, analysis, and applications. W. Andrew, Norwich/Heidelberg
- Kubena RL, Atkinson GM, Robinson WP et al (1996) A new miniaturized surface micromachined tunneling accelerometer. *IEEE Electron Device Lett* 17(6):306–308
- Laermer F, Schilp A, Funk K et al (1999) Bosch deep silicon etching: improving uniformity and etch rate for advanced MEMS applications
- Lee S, Najafi K (2007) A generic environment-resistant packaging technology for MEMS. In: *The 14th international conference on solid-state sensors, actuators and microsystems. TRANSDUCERS 07, Lyon, June 2007*, pp 335–338
- Lee I, Yoon GH, Park JY et al (2005) Development and analysis of the vertical capacitive accelerometer. *Sensors Actuators A Phys* 119(1):8–18

- Li W, Song ZH, Li XL et al (2014) A novel sandwich capacitive accelerometer with a double-sided 16-beam-mass structure. *Microelectron Eng* 115:32–38
- Liu C, Kenny TW (2001) A high-precision, wide-bandwidth micromachined tunneling accelerometer. *J Microelectromech Syst* 10(3):425–433
- Liu Y, Zhao Y, Tian B et al (2014) Analysis and design for piezoresistive accelerometer geometry considering sensitivity, resonant frequency and cross-axis sensitivity. *Microsyst Technol* 20(3):463–470
- Pinto D, Mercier D, Kharrat C (2009) A small and high sensitivity resonant accelerometer. *Procedia Chem* 1:4
- Rocha LA, Silva CS, Cerqueira MF et al (2011) A microinjected 3-axis thermal accelerometer. *Procedia Eng* 25:607–610
- Rodjegård H, Anderson GI, Rusu C et al (2005a) Capacitive slanted-beam three-axis accelerometer II. Characterization. *J Micromech Microeng* 15:1997–2002
- Rodjegård H, Andersson GI, Rusu C et al (2005b) Capacitive slanted-beam three-axis accelerometer: I. Modelling and design. *J Micromech Microeng* 15:1989–1996
- Rudolf F (1983) A micromechanical capacitive accelerometer with a two-point inertial-mass suspension. *Sensors Actuators* 4:191–198
- Su JB, Xiao DB, Wang X et al (2014) Vibration sensitivity analysis of the ‘Butterfly-gyro’ structure. *Microsyst Technol Micro Nanosyst -Inf Storage Process Syst* 20(7):1281–1290
- Tan SS, Liu CY, Yeh LK et al (2011) A new process for CMOS MEMS capacitive sensors with high sensitivity and thermal stability. *J Micromech Microeng* 21(3):35005
- Tay FEH, Iliescu C, Jing J et al (2006) Defect-free wet etching through pyrex glass using Cr/Au mask. *Microsyst Technol* 12(10):935–939
- Thomson WT, Dahleh MD (1998) *Theory of vibration with applications*. Prentice Hall, Upper Saddle River
- Timoshenko S, Goodier JN (1970) *Theory of elasticity*. McGraw-Hill, New York
- Trusov AA, Zotov SA, Simon BR et al (2013) Silicon accelerometer with differential frequency modulation and continuous self-calibration. *Micro Electro Mechanical Systems (MEMS)*, Jan 2013, pp 29–32
- Tseng SH, Wu PC, Tsai HH et al (2014) Monolithic z-axis CMOS MEMS accelerometer. *Microelectron Eng* 119:178–182
- Xiao DB, Wang XH, Zhou ZL et al (2013) A novel fabrication method based on an after thermal oxidation process for the realization of silicon-beams with normative polygon cross sections shapes. *Microsyst Technol Micro Nanosyst -Inf Storage Process Syst* 19(7):1081–1086
- Xiao DB, Li QS, Hou ZQ et al (2016a) A novel sandwich differential capacitive accelerometer with symmetrical double-sided serpentine beam-mass structure. *J Micromech Microeng* 26(2):25005
- Xiao D, Xia D, Li Q et al (2016b) A double differential torsional accelerometer with improved temperature robustness. *Sensors Actuators A Phys* 243:43–51
- Yang B, Wang X, Dai B et al (2015) A new Z-axis resonant micro-accelerometer based on electrostatic stiffness. *Sensors* 15(1):687–702
- Yazdi N, Ayazi F, Najafi K (1998) Micromachined inertial sensors. *Proc IEEE* 86(8):1640–1659



# Micromachined Gas Inertial Sensors

Rong Zhu

## Contents

Overview .....	518
Theory and Operation Principle of Thermal Gas Inertial Sensors .....	519
Design and Fabrication of Thermal Gas Inertial Sensor .....	527
Gas MIMU System .....	531
Configuration of Thermal Gas MIMU .....	531
Error Source Analysis of Thermal Gas MIMU .....	532
Error Compensation Method Based on Neural Network .....	534
Concluding Remarks .....	536
References .....	536

## Abstract

Development of micromachined inertial sensors has been widely addressed for many years. Most micromachined inertial sensors generally use a mechanical structure including a solid proof mass suspended on springs, which raises the complexity of structure and fabrication and particularly restricts the high shock resistance of the sensor. In this chapter, we introduce a kind of micromachined thermal gas inertial sensor by using thermally driven gaseous flow instead of solid proof mass. The sensor generally consists of one or several heaters and multiple thermistors, which detects the deflections of temperature profile induced by inertial quantities. The thermal inertial sensors, including thermal convective accelerometer and thermal gas gyroscope, have exhibited unique advantages of simple structure, low cost, and high shock resistance.

---

R. Zhu (✉)

State Key Laboratory of Precision Measurement Technology and Instrument, Department of Precision Instrument, Tsinghua University, Beijing, China  
e-mail: [zr\\_gloria@mail.tsinghua.edu.cn](mailto:zr_gloria@mail.tsinghua.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,  
[https://doi.org/10.1007/978-981-10-5945-2\\_11](https://doi.org/10.1007/978-981-10-5945-2_11)

517

---

**Keywords**

Thermal gas inertial sensor · Micromachined sensor · Thermal convection · Thermal expansion

---

**Overview**

Micromachined inertial sensors have been developed for many years. Conventionally micromachined inertial sensor uses a mechanical structure including a solid proof mass suspended on springs (Yazdi et al. 1998). For example, micromachined vibrating gyroscopes use vibrating shells (Ayazi and Najafi 2001), tuning forks (Bernstein et al. 1993), and vibrating beams (Maenaka and Shiozawa 1994). These micro inertial sensors are based on the movement of a seismic proof mass. The mechanical proof mass raises the complexity of structure and fabrication, and particularly restricts the high shock resistance of the sensor. Another concept of fluid inertial sensor using fluid medium instead of mechanical proof mass as key moving and sensing element has been reported (Shiozawa et al. 2004; Takemura et al. 2009; Leung et al. 1997, 1998). The sensing principle of the fluid inertial sensor is that the inertial-induced deflection of the gaseous or liquid flow is detected by using distributed thermistors. There are different driving modes for generating fluid flows. Accordingly, the fluid inertial sensors are categorized into jet gyroscope using jet flow (Shiozawa et al. 2004; Takemura et al. 2009), thermal convective accelerometer (Leung et al. 1997, 1998), and thermal gas gyroscope based on heat transfer (Zhu et al. 2005, 2006, 2014; Cai et al. 2013). The former sensor uses a jet flow that is driven by piezoelectric diaphragm pump (Shiozawa et al. 2004) or electro-conjugate fluid (Takemura et al. 2009). To generate an appropriate jet flow, elaborate structures are needed, such as diaphragm, nozzle orifice, or needle electrode, which unavoidably debase the feasibility for miniaturization and simplification of the devices. Comparatively, thermal gas inertial sensor possesses a relative simple structure with one or several heaters and multiple thermistors, which detects the deflections of temperature profile induced by inertial quantities (Zhu et al. 2005). The thermal inertial sensors, including thermal convective accelerometer and thermal gas gyroscope, have exhibited unique advantages of simple structure, low cost, and high shock resistance (Zhu et al. 2006).

Micromachined thermal accelerometer is based on free-convection heat transfer of a tiny hot air bubble in an enclosed chamber, which was firstly reported in the 1990s by Albert M. Leung et al. (1997, 1998). Since then, many researches have been developed on design, fabrication, and testing of the sensors. Up to now, the thermal accelerometers have approached maturity and become commercialized. The thermal gas gyroscope uses a reciprocating gaseous flow driven by thermal convection or thermal expansion, which was reported in the 2000s by Rong Zhu et al. (2005, 2006, 2014; Cai et al. 2013). The convective gyroscope is based on natural convection due to temperature differences which affect the density, and thus relative buoyancy, of the fluid. However, the convective gyroscope has a coupling effect with external acceleration because the natural convection originates from buoyancy that is

proportional to the external acceleration (Zhu et al. 2005; Zhu and Zhou 2006). Therefore another gas flow driven by thermal expansion/contraction is utilized to develop an alternative gas gyroscope (Cai et al. 2013; Zhu et al. 2014). The expansion flow is independent from external linear acceleration and thus overcomes the coupling problem between the acceleration and the angular rate measurements of the sensor. Due to use a similar sensing principle with thermal accelerometer to detect the Coriolis acceleration, the thermal gas gyroscope can simultaneously detect acceleration and thus can be also served as an accelerometer.

---

## Theory and Operation Principle of Thermal Gas Inertial Sensors

Natural convection is a mechanism, or type of heat transport, in which fluid motion is not generated by any external mechanical source (like a pump, fan, suction device, etc.) but only by density differences in fluid due to temperature gradients. In natural convection, the fluid surrounding a heat source receives heat, becomes less dense, and rises. The surrounding, cooler fluid moves to replace it. The cooler fluid is then heated and the process continues, forming a convective current. The driving force for natural convection is buoyancy, a result of differences in fluid density. Because of this, the presence of acceleration such as gravity, or an equivalent force (linear acceleration, centrifugal force, or Coriolis acceleration), is essential for natural convection. Due to the above reason, the convection flow can be used to sense rotation, but it is greatly influenced by the body acceleration due to the gravity and centrifugal effect. Considering another kind of flow in the fluid environment, for example, jet flow or thermal expansion flow which is driven and dominated by an external source rather than the body acceleration, the coupling between the flow and the body acceleration is therefore diminished. Thermal expansion is a tendency of matter to change in volume in response to a change in temperature (Tipler and Mosca 2008). Compared with the jet flow, the thermal expansion flow is easily generated by alternately heating and cooling two thermal elements. As one of element pair heats up and the other cools down, the fluid particles between two elements move about and are conveyed from the hot element to the cold element. As the cold element heats up and the previous heating element cools down, the fluid particles run reversely and generate the backflow between the two elements.

Thermal expansion-induced gaseous flow generally exists in nature, but it is rarely served as an actuating source of a fluid movement because it is weak in macroscale compared with other driving forces, such as forced convection flow generated by external mechanical drive or natural convection flow generated by density differences in fluid due to temperature gradients. However, when a thermal expansion or contraction is driven in a micro-scale chamber by alternately heating and cooling, the gas flow can be greatly enhanced due to short travel distance so as to be possible serving as an effective driving source. Therefore, a gyroscope can be implemented by using the thermal expansion of gaseous medium in a micro chamber to generate a seismic flow along a specific path and detecting the deflection of the flow induced by the Coriolis acceleration that is proportional to the angular rate.

To understand the mechanism of thermally driven gaseous flow, a computational fluid dynamic (CFD) model is developed as follows to describe a three-dimensional flow (ignoring jet flow) in a chamber, which can be solved to obtain a transient flow velocity, temperature, and pressure in the chamber:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{U}) = 0 \quad (1)$$

$$\frac{\partial \rho \vec{U}}{\partial t} + \rho \vec{U} \cdot \nabla \vec{U} = -\nabla P + \rho \vec{f} + \mu \nabla \cdot \tau \quad (2)$$

$$\frac{\partial T}{\partial t} + \vec{U} \cdot \nabla T = \alpha \nabla^2 T \quad (3)$$

$$P = \rho R_g T \quad (4)$$

where  $\tau = \nabla \vec{U} + (\nabla \vec{U})^T - \frac{2}{3} \sigma \nabla \cdot \vec{U}$ ,  $\sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ;  $\vec{r} = (x, y, z)$  refers to position vector;  $\vec{U} = (U_x, U_y, U_z)^T$  refers to velocity vector;  $\vec{f} = (g_x, g_y, g_z)^T$  refers to acceleration vector;  $\rho$ ,  $P$ ,  $\mu$ , and  $T$  are gas density, pressure, dynamic viscosity, and temperature;  $\alpha = \frac{\lambda}{\rho C_p}$ ;  $\lambda$  is thermal conductivity;  $C_p$  is specific heat; and  $R_g$  is gas constant. The vector operator  $\nabla$  is defined as gradient  $\nabla \equiv i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z}$ . The operator  $\nabla \cdot$  is defined as divergence  $\nabla \cdot \vec{U} \equiv \frac{\partial U_x}{\partial x} + \frac{\partial U_y}{\partial y} + \frac{\partial U_z}{\partial z}$ .

The above momentum (Eq. (2)) characterizes the kinetic process of three-dimensional flow. Considering about the rotation vector  $\vec{\Omega} = (\Omega_x, \Omega_y, \Omega_z)$  that refers to angular rate vector, and expanding the two items in the left side of Eq. (2), the momentum equation can be equivalently converted to the following formula:

$$\begin{aligned} \frac{\rho \partial \vec{U}}{\partial t} + \rho \vec{U} \cdot \nabla \vec{U} = & -\nabla P + \rho \vec{f} - \vec{U} \frac{\partial \rho}{\partial t} - 2\rho \vec{\Omega} \times \\ & \vec{U} - \rho \vec{\Omega} \times (\vec{\Omega} \times \vec{r}) + \mu \nabla \cdot \tau \end{aligned} \quad (5)$$

Defining the coefficient of thermal expansion as  $\alpha_v$  and assuming admitting a small variation of density with temperature (Lienhard et al. 2012), then

$$\alpha_v = \frac{1}{V} \frac{\partial V}{\partial T} \Big|_p = -\frac{1}{\rho} \frac{\partial \rho}{\partial T} \Big|_p \approx -\frac{1}{\rho} \frac{\rho - \rho_0}{T - T_0} \quad (6)$$

$$\nabla P = \rho_0 \vec{f} \quad (7)$$

where  $T_0$  refers to the ambient temperature and  $\rho_0$  refers to the gas density at the ambient temperature. Substituting Eqs. (6) and (7) into the momentum Eq. (5), then there is

$$\begin{aligned} \frac{\partial \vec{U}}{\partial t} + \vec{U} \cdot \nabla \vec{U} &= \alpha_v (T - T_0) \vec{f} + \alpha_v \vec{U} \frac{\partial T}{\partial t} \\ -2\vec{\Omega} \times \vec{U} - \vec{\Omega} \times (\vec{\Omega} \times \vec{r}) &+ \frac{\mu}{\rho} \nabla \cdot \tau \end{aligned} \quad (8)$$

Define  $\vec{\bar{U}} = \frac{\vec{U}}{(v/H)} = \left( \frac{U_x}{v/H}, \frac{U_y}{v/H}, \frac{U_z}{v/H} \right)^T$ ,  $\vec{\bar{r}} = \frac{\vec{r}}{H} = \left( \frac{x}{H}, \frac{y}{H}, \frac{z}{H} \right)^T$ ,  $\Delta\theta = \frac{T-T_0}{T_h-T_0}$ ,  $\vec{\bar{\Omega}} = \frac{\vec{\Omega}}{(v/H^2)} = \left( \frac{\Omega_x}{v/H^2}, \frac{\Omega_y}{v/H^2}, \frac{\Omega_z}{v/H^2} \right)^T$ ,  $\delta = tw/H^2$ , where  $T_h$  is the heater temperature,  $\nu$  ( $\nu = \frac{\mu}{\rho}$ ) is the momentum viscosity, and  $H$  is the feature dimension. Put a non-dimensionalization to the momentum Eq. (8) and get a nondimensional model:

$$\frac{\partial \vec{\bar{U}}}{\partial \delta} + \vec{\bar{U}} \cdot \vec{\nabla} \vec{\bar{U}} = \vec{Gr} \Delta\theta + \alpha_v \frac{\partial T}{\partial \delta} \vec{\bar{U}} - 2\vec{\bar{\Omega}} \times \vec{\bar{U}} - \vec{\bar{\Omega}} \times (\vec{\bar{\Omega}} \times \vec{\bar{r}}) + \vec{\nabla} \cdot \vec{\bar{\tau}} \quad (9)$$

where the Grashof number  $\vec{Gr} = (Gr_x, Gr_y, Gr_z) = \left( \frac{\alpha_v (T_h - T_0) g_x H^3}{\nu^2}, \frac{\alpha_v (T_h - T_0) g_y H^3}{\nu^2}, \frac{\alpha_v (T_h - T_0) g_z H^3}{\nu^2} \right)$ ,  $\vec{\nabla} = \left( \frac{\partial}{\partial \bar{x}}, \frac{\partial}{\partial \bar{y}}, \frac{\partial}{\partial \bar{z}} \right)^T$ ,  $(\bar{x}, \bar{y}, \bar{z}) = \left( \frac{x}{H}, \frac{y}{H}, \frac{z}{H} \right)$ , and  $\vec{\bar{\tau}} = \vec{\nabla} \cdot \left( \vec{\bar{\nabla}} \vec{\bar{U}} \right)^T - \frac{2}{3} \sigma \vec{\nabla} \cdot \vec{\bar{U}}$ . The first item  $\vec{Gr} \Delta\theta$  at the right of Eq. (9) is a driving force to generate a heat convective flow that is correlative with the Grashof number  $\vec{Gr}$ ; the second item  $\alpha_v \frac{\partial T}{\partial \delta} \vec{\bar{U}}$  at the right of Eq. (9) is another driving force to generate an expansion flow that is correlative with the transient variation rate of the temperature, which is induced by alternately heating and cooling thermal elements.

From Eq. (9), we can analyze the inertial sensing mechanisms for thermal accelerometer and thermal gyroscope. For sensing acceleration, the above model can be simplified as the following formula by setting the rotation as zero and ignoring the transient variation of the temperature:

$$\frac{\partial \vec{\bar{U}}}{\partial \delta} + \vec{\bar{U}} \cdot \vec{\nabla} \vec{\bar{U}} = \vec{Gr} \Delta\theta + \vec{\nabla} \cdot \vec{\bar{\tau}} \quad (10)$$

The fluid flow  $\vec{\bar{U}}$  is generated by the buoyancy, a result of the temperature difference  $\Delta T = T_h - T_0$  and the acceleration  $\vec{g} = (g_x, g_y, g_z)$ . This fluid flow induces a deflection of the temperature profile based on the energy Eq. (3) that can be detected by using several distributed thermistors. Consequently, it is seen that the output of the thermal accelerometer is dominated by the Grashof number  $\vec{Gr}$ .

To sense an angular rate of rotation, consider flow vectors generated by two driving sources: buoyancy and volumetric thermal expansion. Two driving sources result in two flow vectors, defined as  $\vec{\bar{U}}_c$  for the convective flow and  $\vec{\bar{U}}_e$  for the



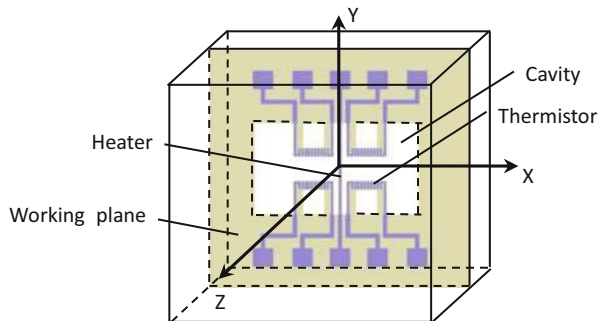
expansion flow. When the rotation  $\vec{\Omega}$  occurs, the nondimensional Coriolis acceleration  $-2\vec{\Omega} \times (\vec{U}_c + \vec{U}_e)$  is engendered that induces a deflection of the temperature field and generates the output of the gyroscope. Obviously, the convective flow  $\vec{U}_c$  interconnects with the acceleration, and the expansion flow  $\vec{U}_e$  is absolutely independent from the acceleration. Accordingly, the thermal convection gyroscope exists a high coupling with the external acceleration; specifically the scale factor (i.e. sensitivity) of the gyroscope is dependent on the acceleration (Zhu and Zhou 2006; Zhu et al. 2005). In contrast, the thermal expansion gyroscope is independent from the acceleration. The sensitivity of the thermal expansion gyro is correlated with the expansion number  $\alpha_v \frac{\partial T}{\partial \delta}$  that is proportional to the transient variation rate of the temperature.

A schematic view of a thermal gas inertial sensor is shown in Fig. 1. X- and Y-axes are central axes located on the upper surface of the cavity, and Z-axis constructs a Cartesian coordinate with X- and Y-axes. A suspended central heater placed in the X-Y Cartesian plane (specified as working plane, abbreviated to WP) heats up and lowers the density of the surrounding gas. Four suspended thermistor wires placed symmetrically on both sides of the heater detect the local temperatures and gas flows. The cavity, heater, and thermistors are assembled and packaged in a sealed chamber to prevent disturbances from external environment.

As mentioned above, inertial forces and temperature differences drive the gas sealed in the chamber moving, regular flow engenders. For example, when gravity is applied on the reverse direction of the Z-axis and the central heater heats up, the convection-induced gas flow in the region of the hermetic chamber is depicted in Fig. 2a, b. The two gas flows on two sides of the heater in the WP move along X-axis and are antisymmetric about Y-axis.

Considering the working plane, apply an angular rate rotation  $\vec{\omega}_z$  around the Z-axis. It induces Coriolis acceleration  $\vec{a}_c$  normal to the flow direction  $\vec{v}_x$ .  $\vec{v}_x$  represents the component of the flow velocity along X in the WP and is antisymmetric about Y-axis. Combining non-Coriolis acceleration  $\vec{a}_y^0$  with  $\vec{a}_c$ , the total acceleration  $\vec{a}_y$  along the Y-axis is expressed by

**Fig. 1** Schematic view of a gas inertial sensor (Zhu et al. 2005)



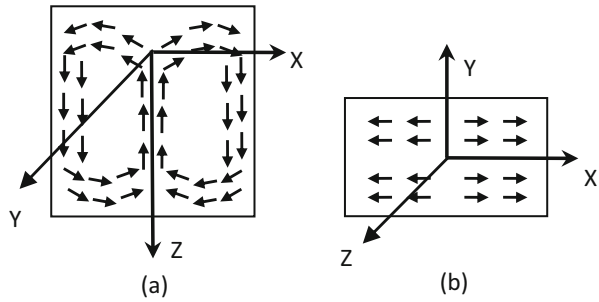
$$\vec{a}_y = \vec{a}_c + \vec{a}_y^0 = 2\vec{\omega}_z \times \vec{v}_x + \vec{a}_y^0 \tag{11}$$

$$a_y = 2\omega_z \cdot v_x + a_y^0$$

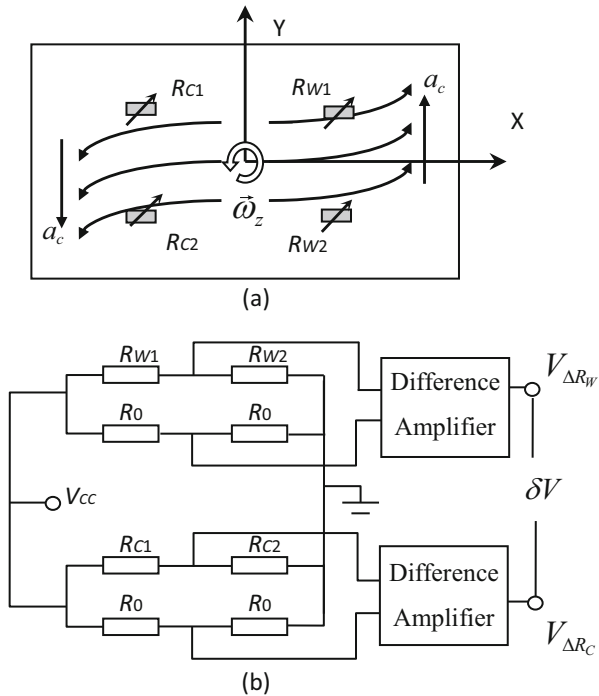
i.e.,  $Opp.a_y' = -2\omega_z \cdot v_x + a_y^0$  (12)

where  $a_y$  and  $a_y'$  represent the accelerations along Y-axis on the two sides of the heater, respectively. It is obvious that the Coriolis accelerations  $\vec{a}_c$  on the two sides of the heater point to the opposite directions so as to drive the flows  $\vec{v}_x$  on the two sides of the heater deflecting in opposite directions of Y as shown in Fig. 3a. The

**Fig. 2** Schematic view of thermal convection flow in region of a hermetic chamber (Zhu et al. 2005)



**Fig. 3** The inverse symmetric flow in the WP is deflected in opposite directions of Y due to Coriolis acceleration, and Wheatstone bridge circuit outputs the angular rate (Zhu et al. 2005)



flow deflection causes opposite cooling effects between the two thermistors in pair along the Y-axis (specifically  $R_{W1}$  and  $R_{W2}$  form one pair,  $R_{C1}$  and  $R_{C2}$  form another pair), and as a result, the resistances of  $R_{W1}$  and  $R_{W2}$  ( $R_{C1}$  and  $R_{C2}$ ) change oppositely. Wheatstone bridge circuit as shown in Fig. 3b converts the resistance changes of the two thermistor pairs  $R_{W1} - R_{W2}$  and  $R_{C1} - R_{C2}$  to the voltages  $V_{\Delta R_W}$  and  $V_{\Delta R_C}$  and further makes an output voltage difference  $\delta V$  between them. Under a constant voltage  $V_{cc}$ , the voltage  $V_{\Delta R_W}$  is proportional to  $R_{W1} - R_{W2}$ . Therefore,  $V_{\Delta R_W}$  is proportional to the acceleration  $a_y$ . Similarly, the voltage  $V_{\Delta R_C}$  is proportional to the acceleration  $a_y'$ . The output voltage difference  $\delta V$  can be given by

$$\delta V = V_{\Delta R_W} - V_{\Delta R_C} \propto a_y - a_y' = 4\omega_z \cdot v_x \tag{13}$$

Equation (13) indicates that the voltage difference  $\delta V$  is proportional to the angular rate  $\omega_z$ . The differential processing can eliminate non-Coriolis accelerations  $a_y^0$ , which are equal on two sides of the heater and are subtracted from the output  $\delta V$  via the difference of  $V_{\Delta R_W} - V_{\Delta R_C}$ . On the other hand, the sum of  $\delta V' = V_{\Delta R_W} + V_{\Delta R_C} \propto a_y + a_y' = 2a_y^0$ , which is proportional to the external acceleration  $a_y^0$ . Therefore, by using differential and summing amplifiers, both angular rate and acceleration can be extracted simultaneously.

The entire working process of the thermal gas inertial sensor consists of multi-physics interactions: electrical-thermal conversion of heaters, heat conduction, thermal convective/expansion flow, and thermal-electrical conversion of thermistors. A block diagram of system model, including heating source, heat transfer, gas flow, and temperature sensing, is established as shown in Fig. 4 (Zhu et al. 2010). The established transfer function model for the system can help to characterize the sensing mechanism and optimize the sensor design.

Firstly, consider the heating source. The electric power supplied to the heating resistor is dissipated by heat transfer toward the ambient fluidic medium and also toward heating resistor, which leads to a temperature difference between the heater and ambience. According to the energy principle (Lienhard 1987), the dynamic process of the heating can be modeled by

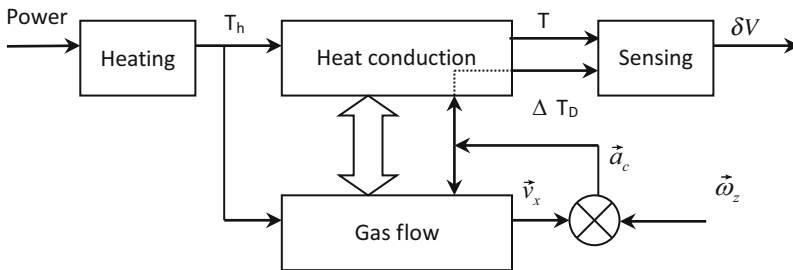


Fig. 4 Block diagram of the gas inertial sensor model (Zhu et al. 2010)

$$\begin{aligned} C \frac{\partial T_h}{\partial t} &= P_h - hg_0 \Delta T_h \\ \Delta T_h &= (T_h - T_a) \end{aligned} \quad (14)$$

where  $T_h$  and  $T_a$  refer to the temperatures of the heater and ambience, respectively,  $C$  is the thermal capacity of the heater,  $P_h$  is the electrical power,  $h$  is the heat transfer coefficient, and  $g_0$  is a constant coefficient depending on the geometrical parameters of the heater. According to linear perturbation theory, the heat transfer coefficient  $h$  can be considered to be constant. Performing Laplace transform to Eq. (14), the transfer function of the heating source can be formulated by a first-order model, where  $s$  represents differential operator:

$$G_1(s) = \frac{T_h(s)}{P_h(s)} = \frac{k_1}{1 + \tau_1 s} \quad (15)$$

where  $\tau_1 = C/hg_0$  and  $k_1 = 1/hg_0$ .

Secondly, analyze the process of the heat conduction in gas medium. The temperature difference between the heater and ambience leads to a conductive heat transfer in the gaseous medium within the chamber. According to the heat transfer principle (Lienhard 1987), the temperature  $T$  at any point in the chamber can be ruled by

$$\frac{\partial T}{\partial t} = \alpha \cdot \nabla^2 T \quad (16)$$

Only consider the heat conduction in the working plane and define  $x$  as characteristic dimension for the device. Therefore Eq. (16) can be reduced to  $\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2}$ . Solving the partial differential equation using separation variable (Tannehill et al. 1997) together with the boundary conditions  $T_h$  at the heater, the following first-order transfer relationship can be obtained:

$$G_2(s) = \frac{T(s)}{T_h(s)} = \frac{1}{1 + \tau_2 s} \quad (17)$$

where  $\tau_2 = - \iint T_0(x) dx^2 / (\alpha T_0(x))$ , and  $T_0(x)$  is a normalized shape function of temperature profile.

In the process of engendering gas flow, the flow is generated by the source of buoyancy or thermal expansion/contraction. According to Navier-Stokes equation (Eq. 8) and omitting the quadratic terms for simplification, the engendered flow of the gas can be ruled by

$$\frac{\partial \vec{U}}{\partial t} = \alpha_v (T - T_0) \vec{f} + \alpha_v \vec{U} \frac{\partial T}{\partial t} + \frac{\mu}{\rho} \nabla \cdot \nabla \vec{U} \quad (18)$$

Solving Eq. (18) using separation variable approach, the transfer function between the temperature  $T$  and the flow velocity  $\vec{U}$  of the gas can be given by

$$G_3(s) = \frac{\vec{U}(s)}{T(s)} = \frac{k_2 + k_3s}{1 + \tau_3s} \quad (19)$$

where  $\tau_3 = -\vec{U}_0 / \left( \frac{\mu}{\rho} \nabla \cdot \nabla \vec{U}_0 \right)$ ,  $k_2 = -\alpha_v \vec{f} / \left( \frac{\mu}{\rho} \nabla \cdot \nabla \vec{U}_0 \right)$ ,  $k_3 = -\alpha_v \vec{U}_0 / \left( \frac{\mu}{\rho} \nabla \cdot \nabla \vec{U}_0 \right)$ , and  $\vec{U}_0$  is a normalized shape function of  $\vec{U}$ . Consider the angular rate  $\Omega_z$  around the Z-axis that induces Coriolis acceleration  $\omega_z \cdot v_x$  along the Y-axis, which leads to a deflection of the temperature profile that is detected by the thermistors. Therefore, define the temperature difference  $\Delta T_D$  across the thermistors that is proportional to Coriolis acceleration given by

$$\Delta T_D \propto \omega_z \cdot v_x \text{ or } \Delta T_D = k_4 \cdot \omega_z \cdot v_x \quad (20)$$

where  $k_4$  is a constant coefficient depending on thermal and fluidic properties of gas.

The thermistors convert the temperature signals into the resistances. Due to thermal inertia of the thermistors, another first-order transfer function is considered since thermistors have to be in equilibrium with the local temperature of gas to convert temperature variations into electrical resistance variations. The first-order transfer function represents the signal transfer from the local temperature difference  $\Delta T_D$  to the temperature difference on the thermistors  $\Delta T_d$ :

$$G_4(s) = \frac{\Delta T_d}{\Delta T_D} = \frac{1}{1 + \tau_4s} \quad (21)$$

where  $\tau_4$  represents the time constant of the thermal inertia of the thermistors.

Using a Wheatstone bridge circuit, the temperature difference among the thermistors is proportionally converted into a voltage difference  $\delta V$  (Zhu et al. 2005). This process can be formulated by

$$\delta V = k' \Delta T_d \quad (22)$$

where  $k'$  is a constant coefficient depending on the parameters of the circuit.

Combining the transfer functions of all subsystems and ignoring the high-order terms to yield a simplified second-order spring-damping model, the entire transfer function from the heating power  $P_h$  to the output voltage can be given by

$$H(s) = \frac{Y(s)}{X(s)} = \frac{\lambda \cdot \omega_z \cdot (k_2 + k_3s)}{s^2 + cs + k} \quad (23)$$

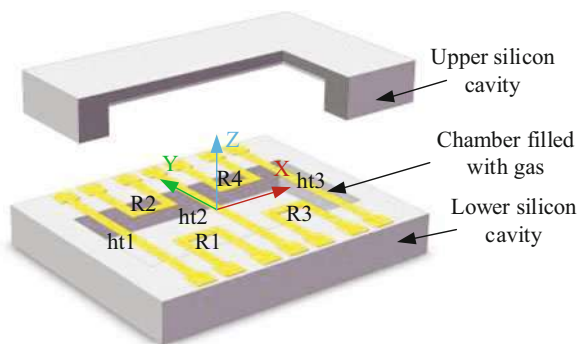
where  $X(s)$  refers to the Laplace vector of the applied electrical power  $P_h$  on the heater,  $Y(s)$  refers to the Laplace vector of the output voltage  $\delta V$ ,  $k = 1/\tau''$  and  $c = \tau'/\tau''$  denote equivalent stiffness and damping coefficient,  $\lambda = k'k_1k_4/\tau''$  is a gain representing the sensor sensitivity,  $\tau' = \sum_{i=1}^4 \tau_i$ , and  $\tau'' = \tau_1\tau_2 + \tau_1\tau_3 + \tau_1\tau_4 + \tau_2\tau_3 + \tau_2\tau_4 + \tau_3\tau_4$ . In practical use, the coefficients  $k$ ,  $c$ , and  $\lambda$  can be determined through experimental calibration.

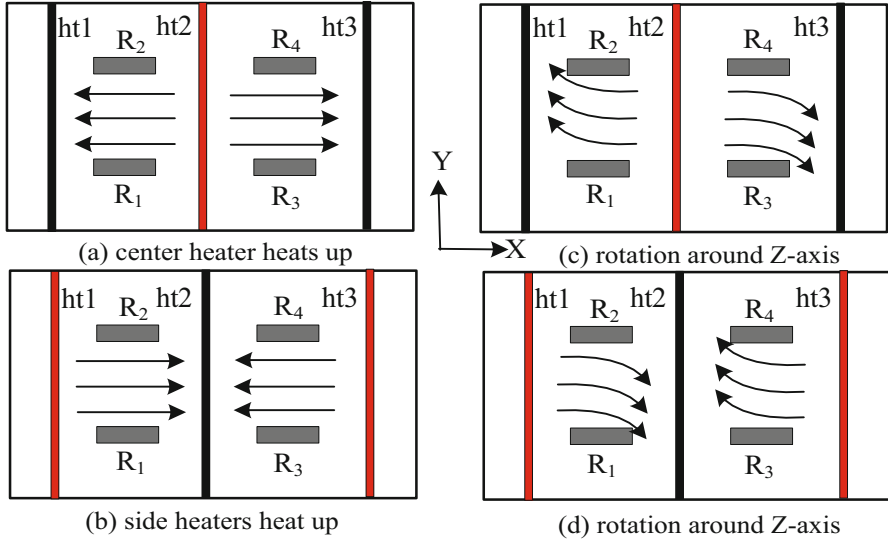
The above modeling processes rely on the fundamentals of fluid mechanics and heat transfer, in association with empirical techniques. Further analysis based on the above modeling concluded that the nonlinearity of the sensor output was mainly attributable to the variable stiffness of the sensor system and the structural asymmetry due to nonideal fabrication (Zhu et al. 2010).

## Design and Fabrication of Thermal Gas Inertial Sensor

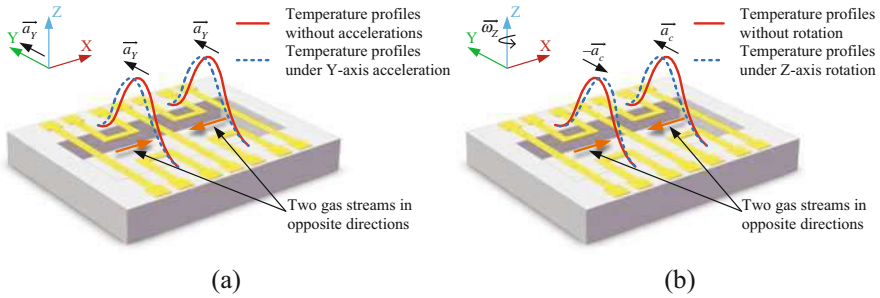
As mentioned above, there are two types of gas inertial sensors, thermal convective inertial sensor and thermal expansion inertial sensor. The thermal convective sensor has great cross-coupling effects between acceleration and rotation measurements due to buoyancy dependence, while the thermal expansion-based sensor reduces the cross-coupling by employing expansion flow instead of convective flow. A thermal gas inertial sensor is composed of at least one heater and multiple thermistors distributed around the heater. The heater is used for engendering gas flow in a micro chamber, while the thermistors are used for detecting the temperature variation in the chamber induced by inertial quantity. One configuration of thermal convective inertial sensor is shown in Fig. 1, where one center heater and four thermistors are structured. Another configuration of inertial sensor based on thermal expansion is shown in Fig. 5, comprised of three heater wires (ht1, ht2, ht3), four thermistor wires (R1, R2, R3, R4), and a sealed micro chamber, which is filled with gas medium (Cai et al. 2013; Zhu et al. 2014). Three heaters are suspended over a cavity etched on silicon substrate, while four thermistors are suspended on the cavity and symmetrically distributed between two adjacent heaters. The heaters and thermistors form a working plane, where X- and Y-axes of Cartesian coordinate are located. Three heaters are adopted, where the center heater ht2 and the two side heaters ht1 and ht3 are electrically heated in alternating way so as to generate two flows in opposite direction as show in Fig. 6. When a rotation around the Z-axis occurs, the flow along the X-axis will be deflected to the Y-axis due to the Coriolis acceleration, resulting in the deflection of temperature profile along the Y-axis that is detected by the thermistors symmetrically distributed around the heaters.

**Fig. 5** Schematic diagram of a thermal expansion inertial sensor (Liu and Zhu 2016)





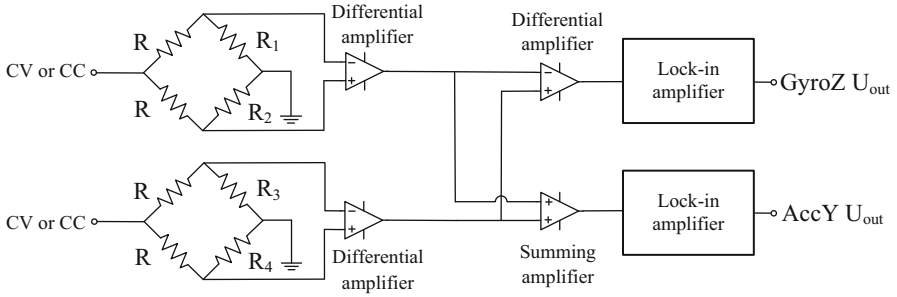
**Fig. 6** Flow field in a thermal expansion sensor (Zhu et al. 2014)



**Fig. 7** Temperature distribution deflected by (a) a linear acceleration and (b) Coriolis acceleration (Liu and Zhu 2016)

The temperature profiles in the Y-axis direction are shown in Fig. 7, where Fig. 7a represents the deflection of the temperature profile induced by a linear acceleration  $\vec{a}_y$ ; Fig. 7b represents the deflection of the temperature profile induced by the Coriolis acceleration due to the rotation  $\vec{\omega}_z$  (Liu and Zhu 2016). The Coriolis accelerations on two sides of the Y-axis are opposite and thus deflect the temperature profiles to the opposite directions along the Y-axis. Therefore the Z-axis rotation can be read out by the following formula:

$$\text{Output}_{\text{GyroZ}} \propto [T(R_1) - T(R_2)] - [T(R_3) - T(R_4)] \propto \omega_z \tag{24}$$



**Fig. 8** Schematic diagram of the read-out circuit of the four thermistors (Zhu et al. 2014)

where  $T(R_1)$ ,  $T(R_2)$ ,  $T(R_3)$ , and  $T(R_4)$  are the temperatures of the four thermistors. On the other hand, the linear acceleration deflects the temperature profiles on the two sides of the Y-axis in the same direction, and thus the Y-axis acceleration is read out by the following formula:

$$\text{Output}_{\text{AccY}} \propto [T(R_1) - T(R_2)] + [T(R_3) - T(R_4)] \propto a_y \tag{25}$$

Thermistors can be operated either in constant voltage (CV) or constant current (CC) modes as shown in Fig. 8, where R is the reference resistance used for balancing the Wheatstone bridge. According to Eq. (24), the Z-axis rotation outputs under CV and CC, respectively, can be deduced by.

$$\text{Output}_{\text{GyroZ}}^{\text{CV}} \propto \frac{U_c R(R_1 - R_2)}{(R + R_1)(R + R_2)} - \frac{U_c R(R_3 - R_4)}{(R + R_3)(R + R_4)} \approx \frac{\alpha R_0 U_c \Delta T_G}{4R} \tag{26}$$

$$\text{Output}_{\text{GyroZ}}^{\text{CC}} \propto \frac{I_c R(R_1 - R_2)}{2R + R_1 + R_2} - \frac{I_c R(R_3 - R_4)}{2R + R_3 + R_4} \approx I_c \alpha R_0 \Delta T_G / 4 \tag{27}$$

where  $U_c$  and  $I_c$  are the constant voltage and constant current applied on the bridge in the CV and CC modes,  $R_0$  is the resistance of the thermistor at 0°C, and  $\alpha$  is the temperature coefficient of resistance (TCR) of the thermistor (assume four thermistors have same  $R_0$  and TCR).  $\Delta T_G = [T(R_1) - T(R_2)] - [T(R_3) - T(R_4)]$ . According to Eq. (25), the Y-axis acceleration outputs under CV and CC, respectively, can be deduced by

$$\text{Output}_{\text{AccY}}^{\text{CV}} \propto \frac{U_c R(R_1 - R_2)}{(R + R_1)(R + R_2)} + \frac{U_c R(R_3 - R_4)}{(R + R_3)(R + R_4)} \approx \frac{\alpha R_0 U_c \Delta T_A}{4R} \tag{28}$$

$$\text{Output}_{\text{AccY}}^{\text{CC}} \propto \frac{I_c R(R_1 - R_2)}{2R + R_1 + R_2} + \frac{I_c R(R_3 - R_4)}{2R + R_3 + R_4} \approx I_c \alpha R_0 \Delta T_A / 4 \tag{29}$$

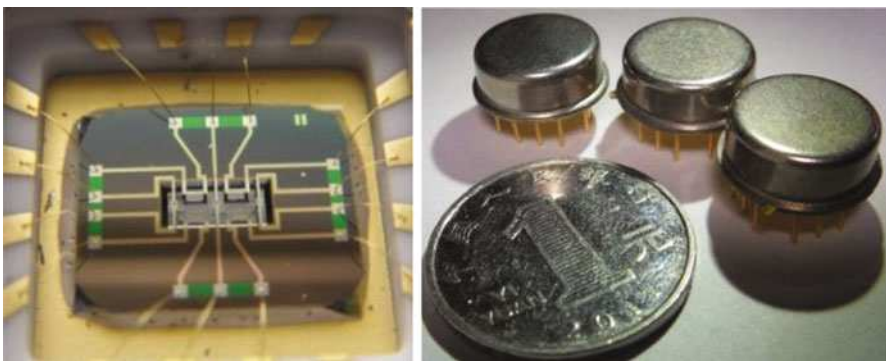
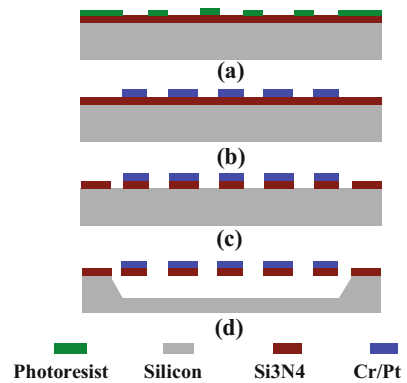
where  $\Delta T_A = [T(R_1) - T(R_2)] + [T(R_3) - T(R_4)]$ .



From Eqs. (26), (27), (28), and (29), it is seen that the sensitivities of the thermal gas inertial sensor including rotation sensitivity and acceleration sensitivity are proportional to the applied power (applied voltage in CV or applied current in CC) and also proportional to the TCR of the thermistors. The TCRs of the thermistors are dominated by the material, geometry, and fabrication of the elements. Thermal sensitive materials with high TCRs, such as platinum (Pt), nickel (Ni), and doped silicon, can be used. Fabrication should be designed in consideration of its compatibility, stability, and simplification.

A fabrication of the sensor is illustrated in Fig. 9, where Pt is used as the material of the heaters and thermistors considering its excellent thermal sensitivity and stability. The fabrication process starts with depositing low-stress silicon nitride by LPCVD on silicon substrate. Then a thermosensitive metal layer, including platinum and Cr adhesion layer, is deposited and patterned (Fig. 9a, b). RIE (reactive-ion etching) is utilized to open the etching window of the cavity (Fig. 9c). Afterward, as shown in Fig. 9d, the cavity is formed using wet anisotropic etching. Finally, the chip is hermetically packaged and filled with a given gas medium. The prototype of the fabricated sensor is shown in Fig. 10.

**Fig. 9** Fabrication process



**Fig. 10** Prototype of fabricated thermal gas inertial sensor

The above thermal gas inertial sensor capably simultaneously measures one-axis angular rate ( $\pm 3000$  °/s, the nonlinearity  $< 0.57\%$ ) and one-axis acceleration ( $\pm 10$  g) in one chip by using the differential operational circuit shown in Fig. 8 and is provided with low coupling effect via using thermal expansion flow instead of buoyancy-based convection flow (Cai et al. 2013; Zhu et al. 2014). The sensor demonstrated an angular rate sensitivity of about  $95 \mu\text{V}/^\circ/\text{s}$  with the amplification of 36,000 and an acceleration sensitivity of 300 mV/g with the amplification of 10,000, as well as extremely high shock resistance ( $> 20,000$  g) (Cai et al. 2013; Zhu et al. 2014).

---

## Gas MIMU System

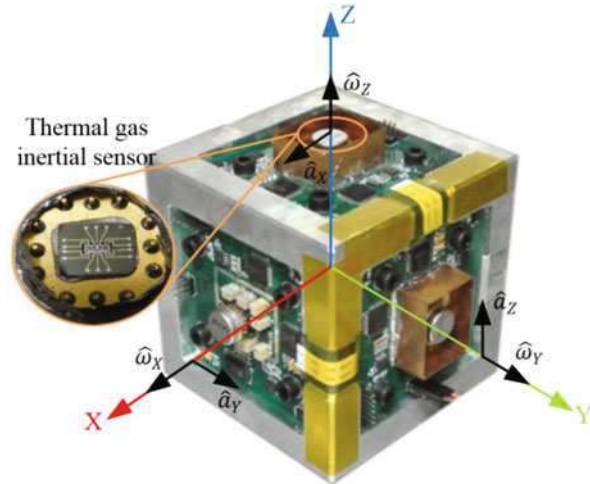
### Configuration of Thermal Gas MIMU

Micromachined inertial measurement unit (MIMU) has played an important role in civil and military applications (Barbour and Schmidt 2001; Höflinger et al. 2013; Brown and Lu 2004; Hanse 2004). In general, a conventional MIMU with six degrees of freedom is a three-dimensional integrated system consisting of three uniaxial micromachined accelerometers (or a three-axis accelerometer) and three uniaxial micromachined gyroscopes (or a three-axis gyroscope) mounted orthogonally in a Cartesian coordinate system. The accelerometers and gyroscopes measure the accelerations and angular rates around corresponding three axes, respectively. Theoretically, one output of an inertial sensor is proportional to an angular rate or acceleration around the corresponding axis of the system, and thus three orthogonal angular rates and three orthogonal accelerations are measured from six sensor outputs in MIMU.

In this section, a thermal gas MIMU system using only three thermal gas inertial sensors to implement measurements of three-axis angular rates and three-axis accelerations is introduced (Liu and Zhu 2016). As mentioned above, each thermal gas inertial sensor enables to simultaneously measure one-axis acceleration and one-axis rotation by using the conditioning circuit shown in Fig. 8. Three gas inertial sensors are adequate to measure three-axis accelerations and three-axis rotations, which allow to simplify the MIMU structure as well as the conditioning circuit, therefore miniaturize the system size, and reduce the power consumption and cut down the fabrication cost of the MIMU.

Three thermal gas inertial sensors are integrated as shown in Fig. 11. Each sensor is mounted on a PCB, where the conditioning circuit of the sensor is located. Three PCBs are assembled orthogonally to form a three-dimensional frame. A rectangular coordinate frame X-Y-Z is defined as the MIMU's body frame, where X-, Y-, and Z-axes are served as sensing axes of the MIMU. On each axis of X-Y-Z, there locates a thermal gas inertial sensor, the outputs of which correspond to one-axis acceleration and one-axis angular rate. Specifically, the outputs of the sensor on the X-axis ( $\hat{a}_Y, \hat{\omega}_X$ ) correspond to Y-axis acceleration  $a_Y$  and X-axis angular rate  $\omega_X$ , that of the sensor on the Y-axis ( $\hat{a}_Z, \hat{\omega}_Y$ ) correspond to Z-axis acceleration  $a_Z$  and Y-axis angular

**Fig. 11** Configuration of the thermal gas MIMU (Liu and Zhu 2016)



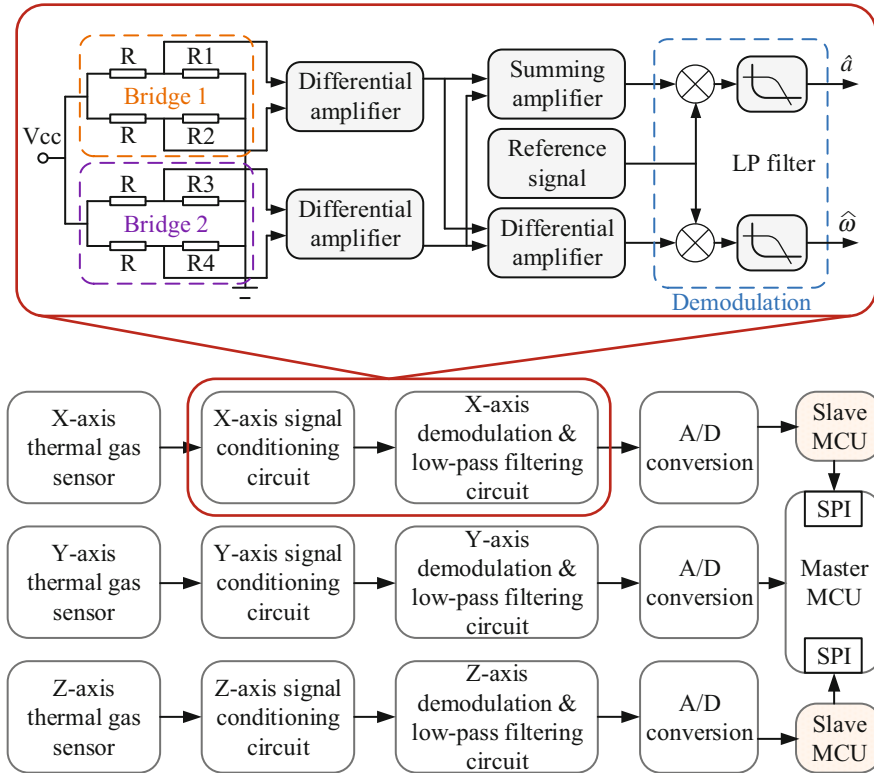
rate  $\omega_y$  and that of the sensor on the Z-axis ( $\hat{a}_x$ ,  $\hat{\omega}_z$ ) correspond to X-axis acceleration  $a_x$  and Z-axis angular rate  $\omega_z$ .

A diagram of the measurement and control circuit of the MIMU is illustrated in Fig. 12. Each sensor circuit consists of two Wheatstone bridges, two-level differential amplifying circuit, demodulation circuit, and low-pass filtering circuit, which are used to implement the signal conversions from the deflections of the temperature profile to the quantities of acceleration and rotation. Each sensor has one A/D converter and one micro control unit (MCU) to extract one-axis acceleration and one-axis angular rate that makes each circuit board independent and thus avoids cross talk. A high-performance MCU acts as a master for operating compensation algorithm.

## Error Source Analysis of Thermal Gas MIMU

Theoretically, the output of the inertial sensor in a MIMU is proportional to an angular rate or acceleration around the corresponding axis of the system, and intuitively three orthogonal angular rates and three orthogonal accelerations are measured directly from the sensor outputs in MIMU. As a matter of fact, however, the outputs of the inertial sensors may be coupled with each other, containing deterministic errors and stochastic errors, all of which possibly corrupt the system accuracy (Zhu et al. 2009).

Error sources of a MIMU mainly consist of cross-coupling, misalignment, eccentricity, noise, etc. Among them, the cross-coupling error is significant for a thermal gas MIMU due to the complex fluidic motion in the sensor chamber (Zhu et al. 2009). The cross-coupling effects include the coupling errors among the measurements of three-axis accelerations and three-axis angular rates. Although impressive decrease of the coupling has been achieved via using the thermal expansion instead of thermal



**Fig. 12** Schematic diagram of the circuit system for a thermal gas MIMU (Liu and Zhu 2016)

convection (Zhu et al. 2014), the cross-coupling error cannot be ignored yet for the thermal gas MIMU in fact.

From Eq. 9, it is theoretically demonstrated that the thermal convection flow is generally existent in a fluid dynamic system, which unavoidably results in a cross-coupling effect between acceleration and rotation measurements, for example, the influence of the gravitational acceleration on the gyro output at a vertical position of the sensor. In real application, the problem is more complicated due to misalignment and asymmetry of the sensor structure induced by unsatisfied fabrication and thermal distortion in the operation of the sensor, all of which will induce the cross-coupling effects. Consequently, the real couplings exhibit a complex interactive relationship among not only rotation and acceleration measurements but also rotation measurements among axes and acceleration measurements among axes. In addition, it is seen that other inertial sensors based on thermal principle all suffer cross-coupling effects, such as thermal accelerometers. Besides, the misalignment is another inevitable error source for the MIMU. Misalignment errors generally exist when assembling multiple sensors into a three-dimensional system, which bring about negative effects on the system accuracy. In addition, assembly position of individual sensor deviated

from the centroid of the MIMU system will also result in assembling errors (termed as eccentric errors) in the MIMU system.

## Error Compensation Method Based on Neural Network

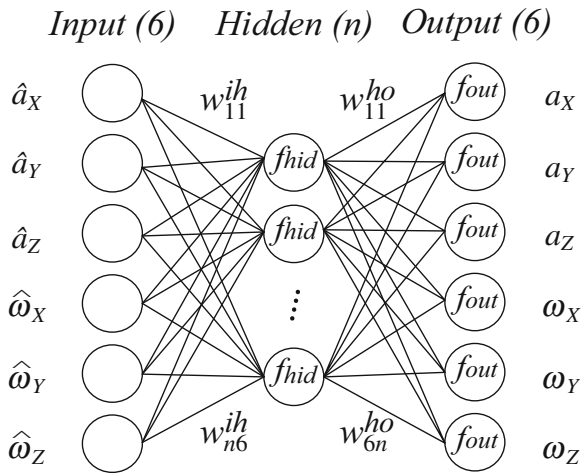
It is seen that the errors of a MIMU contain stochastic errors and deterministic errors. Stochastic errors are usually represented as bias instability and random walk which can be restrained with filtering technologies, such as prediction error minimization (PEM) method, autoregressive (AR) processes, wavelet multiresolution techniques, and so on (Jafari et al. 2014; Naseri and Homaeinezhad 2014). Deterministic errors in a MIMU system are mainly induced by cross-coupling, misalignment, and eccentric problems when assembling multiple sensors into a three-dimensional system. It has experimentally demonstrated that these errors result in navigation errors in quadratic or cubic increasing with the growth of time (El-Sheimy 2006). Compensation, often termed MIMU calibration, is generally used to eliminate the deterministic errors by using error model-based calibration. Up to date, many methods of MIMU calibration have been reported, such as multi-position static calibration (Chatfield 1997; Nieminen et al. 2010; Zhang et al. 2009), fast field calibration (Ma et al. 2014), and optical alignment calibration (Zhu and Zhou 2006). These methods usually need to establish a physical error model prior to calibration process through an analytical method. For example, the model of misalignment error can be formulated by a direction cosine matrix (Zhu and Zhou 2006). However, some errors, such as cross-coupling, are generally difficult to be physically modeled due to their complex, multivariate, and nonlinear coupling. Therefore, an effective system identification method needs to be investigated to solve this problem and consequently eliminate the system errors to most extent.

The basic idea of the proposed compensation method is to establish a relationship model between the outputs of the sensors and the physical quantities to be measured. For a MIMU, the quantities to be measured are the three-axis accelerations and three-axis angular rates along X-Y-Z axes denoted as  $\mathbf{a} = (a_x, a_y, a_z)$  and  $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ , respectively. The outputs of the sensors in the MIMU are the voltage outputs of the sensors representing three-axis acceleration outputs and three-axis angular rate outputs denoted as  $\hat{\mathbf{a}} = (\hat{a}_x, \hat{a}_y, \hat{a}_z)$  and  $\hat{\boldsymbol{\omega}} = (\hat{\omega}_x, \hat{\omega}_y, \hat{\omega}_z)$ , respectively. Obviously, the relationship model of the MIMU is a multi-input and multi-output (MIMO) coupling system, which can be formulated by

$$[\mathbf{a}, \boldsymbol{\omega}] = f[\hat{\mathbf{a}}, \hat{\boldsymbol{\omega}}] \quad (30)$$

In general, the main work for error compensation focuses on mapping relationship  $f$  which is a multivariate and nonlinear function. Analytical method is one of the commonly used methods for approximating  $f$ , but it is usually too complicated and inefficient, especially for cross-coupling effects as mentioned above. Alternative method is experimental identification, which is more reliable and effective. In this section, a system identification method based on experiment to model  $f$  is introduced

**Fig. 13** Structure of three-layer back-propagation (BP) neural network to model the MIMU system errors (Liu and Zhu 2016)



using a three-layer back-propagation (BP) neural network as the model structure (Lau 1991). BP neural network is simpler and more mature than other neural networks. Moreover, it has been theoretically proved that three layers of neural network could solve arbitrarily complicated nonlinear mapping problems (Lau 1991).

Illustrated in Fig. 13, the structure of the BP neural network with six inputs  $[\hat{a}, \hat{\omega}]$  and six outputs  $[a, \omega]$  is constructed and used to represent the model of  $f$ . The network includes several hidden neurons, the number of which is determined by trial. In Fig. 13, the transfer function of the hidden layer  $f_{hid}$ , the function of output layer  $f_{out}$ , and transfer matrix  $w^{ih}$  and  $w^{ho}$  are the key parameters of the network model structure. These parameters, including the transfer functions  $f_{hid}$  and  $f_{out}$ , the number of hidden neurons, and the transfer weights  $w^{ih}$  and  $w^{ho}$ , need to be determined through training using sufficient experimental data, in which the outputs of the thermal gas inertial sensors are used as the inputs of the network while the corresponding data of the actual accelerations and angular rates are used as the outputs of the network. The training of the neural network is to search for appropriate network parameters to minimize the sum-squared error between the network outputs and actual values by using an iteration algorithm. In the standard back-propagation learning algorithm (Lau 1991), two key parameters including the learning rate and momentum parameter need to be configured properly. The learning rate determines the increment of the transfer weights in each iteration step. A large learning rate leads to a fast learning but may results in non-convergence. The additional momentum makes it possible to avoid dropping into local minimum which improves the efficiency of convergence in training. The transfer functions  $f_{hid}$  and  $f_{out}$  play important roles in the model structure and accuracy. The hidden layer  $f_{hid}$  usually uses tan-sigmoid or log-sigmoid function, and the output layer  $f_{out}$  often uses log-sigmoid or pure-line function. The number of hidden neurons is another key parameter dominating the system accuracy and simplicity. A small number of hidden

neurons will cause a large fitting error because it is not enough to model the MIMU system accurately, while too many hidden neurons will be redundant which need more calculation but do not help to improve the model accuracy. Calibration is an experiment method that uses a group of sample data for training the neural network model. The sample data should be sufficient, covering all working conditions of the MIMU.

The demonstration of the above compensation method can be found in literature (Liu and Zhu 2016). One experimental example indicates that by using a neural network-based compensation method, the measurement errors of three-axis angular rates and three-axis accelerations of the gas MIMU can be reduced to less than 1% and 3% of the uncompensated errors.

A neural network is used to comprehensively eliminate the system errors of the thermal gas MIMU, including cross-coupling errors, misalignment errors, eccentric errors, and any other deterministic errors, which are complex multivariate and nonlinear coupling. The neural network-based error compensation method is easily operated, physical-model-free, and universal, which is applicable to not only compensate the system errors of the thermal gas MIMU but also capably overcomes various deterministic errors in general integrated sensor systems.

---

## Concluding Remarks

Micromachined thermal gas inertial sensors have promising prospects due to their unique merits of simple structure, easy and low-cost fabrication, and high shock resistance. Using fluid in substitution for solid proof mass allows inertial measurement to be realized in a chamber filled with gas medium. Dual functions with rotation and acceleration measurability can be readily integrated in one chip of a thermal gas inertial sensor and thus make MIMU system much simple. Although with the above impressive features, hydrothermal fluid in the sensor chamber makes the gas motion and thermodynamic behavior complex and multidimensional coupling. Therefore, cross-coupling effect, thermal compensation, and data fusion should be concerned in practical sensor applications.

---

## References

- Aggarwal P, Syed Z, Niu X, El-Sheimy N (2008) A standard testing and calibration procedure for low cost MEMS inertial sensors and units. *J Navig* 61(02):323–336
- Ayazi F, Najafi K (2001) A HARPSS polysilicon vibrating ring gyroscope. *J Microelectromech Syst* 10(2):169–179
- Barbour N, Schmidt G (2001) Inertial sensor technology trends. *IEEE Sensors J* 1(4):332–339
- Bernstein J, Cho S, King A, Kourepenis A, Maciel P, Weinberg MA (1993) Micromachined comb-drive tuning fork rate gyroscope. In: *Proceedings of the IEEE micro electro mechanical systems (MEMS)*. IEEE, pp 143–148
- Brown A, Lu Y (2004) Performance test results of an integrated GPS/MEMS inertial navigation package. In: *Proceedings of ION GNSS*, S 825–832

- Cai SL, Zhu R, Ding HG, Yang YJ, Su Y (2013) A Micromachined integrated gyroscope and accelerometer based on gas thermal expansion. In: 2013 Transducers & Eurosensors XXVII: the 17th international conference on solid-state sensors, actuators and microsystems (TRANSDUCERS & EUROSENSORS XXVII). IEEE, pp 50–53
- Chatfield AB (1997) Fundamentals of high accuracy inertial navigation, Bd 174. American Institute of Aeronautics and Astronautics, Reston
- El-Sheimy N (2006) Inertial techniques and INS/DGPS integration. In: Engo 623-Course Notes, Department of Geomatics Engineering, University of Calgary, Calgary
- Hanse JG (2004) Honeywell MEMS inertial technology & product status. In: Position location and navigation symposium, PLANS 2004. IEEE, pp 43–48
- Höflinger F, Müller J, Zhang R, Reindl LM, Burgard W (2013) A wireless micro inertial measurement unit (IMU). IEEE Trans Instrum Meas 62(9):2583–2595
- Jafari M, Najafabadi TA, Moshiri B, Tabatabaei SS, Sahebjamayan M (2014) PEM stochastic modeling for MEMS inertial sensors in conventional and redundant IMUs. IEEE Sensors J 14(6):2019–2027
- Lau C (1991) Neural networks: theoretical foundations and analysis. IEEE Press, New York
- Leung AM, Jones J, Czyzewska E, Chen J, Pascal M (1997) Micromachined accelerometer with no proof mass. In: International electron devices meeting, IEDM'97. Technical Digest. IEEE, pp 899–902
- Leung A, Jones J, Czyzewska E, Chen J, Woods B (1998) Micromachined accelerometer based on convection heat transfer. In: Proceedings of the IEEE micro electro mechanical systems (MEMS), pp 627–630
- Lienhard JH (1987) A heat transfer textbook, 3rd edn. Prentice-Hall, Englewood Cliffs
- Lienhard JH, JH IV, L V (2012) A heat transfer textbook. Phlogiston Press, Cambridge, MA
- Liu SQ, Zhu R (2016) System error compensation methodology based on a neural network for a micromachined inertial measurement unit. Sensors 16(2):175
- Ma L, Chen W, Li B, You Z, Chen Z (2014) Fast field calibration of MIMU based on the Powell algorithm. Sensors 14(9):16062–16081
- Maenaka K, Shiozawa T (1994) A study of silicon angular rate sensors using anisotropic etching technology. Sensors Actuators A Phys 43(1–3):72–77
- Naseri H, Homaeinezhad M (2014) Improving measurement quality of a MEMS-based gyro-free inertial navigation system. Sensors Actuators A Phys 207:10–19
- Nieminen T, Kangas J, Suuriniemi S, Kettunen L (2010) An enhanced multi-position calibration method for consumer-grade inertial measurement units applied and tested. Meas Sci Technol 21(10):105204
- Shiozawa T, Dau V, Dao DV, Kumagai H, Sugiyama S (2004) A Dual axis thermal convective silicon gyroscope. In: Proceedings of the 2004 international symposium on micro-nanomechanics and human science. IEEE, pp 277–282
- Takemura K, Yokota S, Suzuki M, Edamura K, Kumagai H, Imamura T (2009) A liquid rate gyroscope using electro-conjugate fluid. Sensors Actuators A Phys 149(2):173–179
- Tannehill J, Anderson D, Pletcher R (1997) Computational fluid mechanics and heat transfer, 2nd edn. Washington, DC, Taylor & Francis Ltd, pp 15–22
- Tipler PA, Mosca GP (2008) Physics for scientists and engineers, vol 1, 6th edn. Worth Publishers, New York
- Yazdi N, Ayazi F, Najafi K (1998) Micromachined inertial sensors. In: Proceedings of the IEEE. vol 8, pp 1640–1659
- Zhang H, Wu Y, Wu W, Wu M, Hu X (2009) Improved multi-position calibration for inertial measurement units. Meas Sci Technol 21(1):015107
- Zhu R, Zhou Z (2006) Calibration of three-dimensional integrated sensors for improved system accuracy. Sensors Actuators A Phys 127(2):340–344
- Zhu R, Su Y, Ding H (2005) A MEMS hybrid inertial sensor based on convection heat transfer. In: The 13th international conference on solid-state sensors, actuators and microsystems. Digest of technical papers. TRANSDUCERS'05. IEEE, pp 113–116
- Zhu R, Ding H, Su Y, Zhou Z (2006) Micromachined gas inertial sensor based on convection heat transfer. Sensors Actuators A Phys 130:68–74



- 
- Zhu R, Ding H, Yang Y, Su Y (2009) Sensor fusion methodology to overcome cross-axis problem for micromachined thermal gas inertial sensor. *IEEE Sensors J* 9(6):707–712
- Zhu R, Ding H, Su Y, Yang Y (2010) Modeling and experimental study on characterization of micromachined thermal gas inertial sensors. *Sensors* 10(9):8304–8315
- Zhu R, Cai S, Ding H, Yang YJ, Su Y (2014) A micromachined gas inertial sensor based on thermal expansion. *Sensors Actuators A Phys* 212:173–180



# A Micromachined Thermal Wind Sensor

Zhenxiang Yi, Ming Qin, and Qing-An Huang

## Contents

Introduction .....	540
Principle .....	541
Hot-Wire Principle .....	541
Calorimetric Principle .....	542
Time-of-Flight (TOF) Principle .....	543
Operation Mode .....	543
Constant Power Mode .....	544
Constant Temperature Difference Mode .....	544
Temperature Balance Mode .....	544
Constant Voltage Mode .....	544
Constant Current Mode .....	544
MEMS Thermal Wind Sensors .....	545
1D Wind Sensors .....	545
2D Wind Sensors .....	546
Sensitivity Improvement and Power Consumption Reduction .....	551
Substrate Etching Technology .....	551
Substrate with Low Thermal Conductivity .....	553
Full-Bridge Technology .....	557
Other Methods .....	559
Package .....	565
Direct Chip Attach Package .....	566
Flip Chip Package .....	566
Au-Au Bonding Package .....	569

---

Z. Yi (✉) · M. Qin

Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, China

e-mail: [xp@seu.edu.cn](mailto:xp@seu.edu.cn); [mqin@seu.edu.cn](mailto:mqin@seu.edu.cn)

Q.-A. Huang

Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing,

Jiangsu, China

e-mail: [hqa@seu.edu.cn](mailto:hqa@seu.edu.cn)

Environmental Effect .....	570
Temperature Effect .....	572
Humidity Effect .....	572
Conclusion .....	573
References .....	575

### Abstract

This chapter introduces the micromachined thermal wind sensor by MEMS technology. First, three working principles and operation modes are addressed. Then, several typical wind sensors including 1D and 2D devices are presented and the experimental results are given. In order to improve the sensitivity and reduce the power consumption, several methods, such as thinned substrate and low-thermal-conductivity substrate, are presented. Experiments demonstrate that the power consumption has been reduced to less than 20 mW by glass reflow technology. In practice, the MEMS wind sensor can be packaged using DCA package, FC package, and Au-Au bonding package. Despite this, the wind sensor is easy to be affected by the surrounding temperature. Related model is established and verified by the experiments.

### Keywords

Wind sensor · Micromachined · MEMS · Thermal conductivity · Full-bridge technology · Package · Environment effect · Temperature difference · Sensitivity · Power consumption

## Introduction

Wind sensors are playing an important role in various fields such as agriculture, weather forecast, and transportation. Before, the wind information was obtained by the wind cups and indicators. After that, the ultrasound anemometers were used to realize wind speed measurement based on Doppler effect (Kato et al. 1992). However, these devices cannot meet the requirements of Internet of Things due to big size and high price. Fortunately, the rise of micro-electro-mechanical systems (MEMS) technology pushed the development of the wind sensor and made the devices with small size and low cost possible. There are two kinds of MEMS wind sensors: thermal wind sensors and nonthermal wind sensors. Because of small size and low cost, the thermal wind sensors have been widely researched around the world for a long time. This chapter reviews the progress of the thermal wind sensors and is arranged as following. In section “[Principle](#),” the principle of the micromachined thermal wind sensor is introduced. Then, different operation modes are categorized in section “[Operation Mode](#).” In section “[MEMS Thermal](#)

Wind Sensors,” a variety of wind sensors including 1D and 2D sensors are presented. Following, several methods are proposed in order to improve the sensitivity and reduce power consumption in section “Sensitivity Improvement and Power Consumption Reduction.” The wind sensors are packaged in section “Package” by different ways. In addition, the environmental effect on the wind sensors is also investigated in section “Environmental Effect.” Finally, the conclusion is drawn in section “Conclusion.”

## Principle

The MEMS thermal wind sensor is based on the effect of wind on the thermal phenomenon. There are three different sensing principles of MEMS thermal wind sensor: hot-wire principle, calorimetric principle, and time-of-flight principle (Kuo et al. 2012).

### Hot-Wire Principle

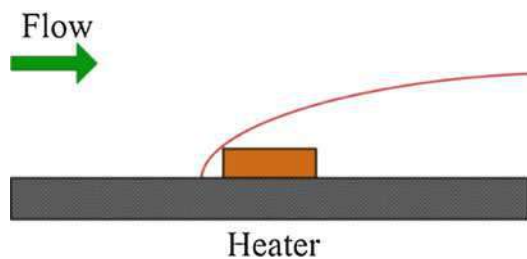
Figure 1 shows the schematic of the wind sensor based on hot-wire principle. An electrical signal is applied on the thermistor and it is heated based on Joule effect. When the wind blows, the heat is taken away. Therefore, the temperature of the thermistor decreases and the wind speed can be resolved by the resistance measurement of the thermistor (Oudheusden 1988; Giani et al. 2002).

The relationship between the heat taken away and the wind speed can be expressed by King’s law (1914).

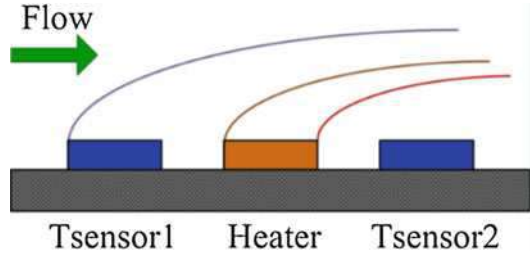
$$P = (A + BU^{0.5})\Delta T \quad (1)$$

where  $P$  is the heat taken away by the wind,  $U$  is the wind speed,  $\Delta T$  is the temperature difference between the chip and the environment,  $A$  and  $B$  are determined by the size of the sensor and fluid property. Clearly, the higher the wind speed is, the more the power is taken away. The hot-wire sensor can only measure wind speed while wind direction cannot be obtained.

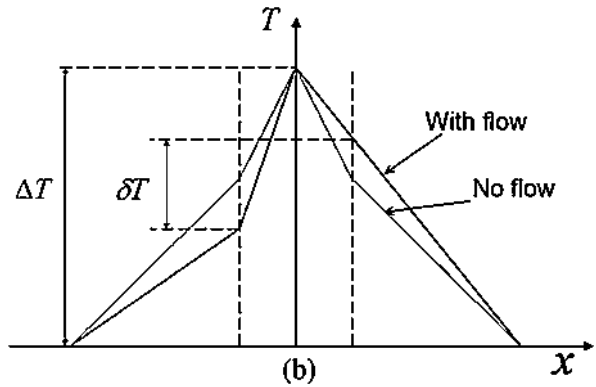
**Fig. 1** Schematic of the wind sensor based on hot-wire principle



**Fig. 2** Schematic of the wind sensor based on calorimetric principle



**Fig. 3** Temperature distribution of the sensor by calorimetric principle



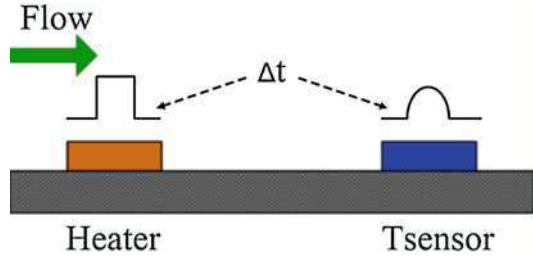
### Calorimetric Principle

As shown in Fig. 2, when an electrical signal is applied, the resistor is heated and the temperature field is formed. Two temperature sensors are placed upstream and downstream with the same distance to the heater. In case of no wind, the temperature field is symmetrical and the temperature  $T_1$  and  $T_2$  is equal. If the wind blows, the temperature field is destroyed and the heat is taken from upstream to downstream. As a result, the temperature upstream decreases while the temperature downstream increases, as shown in Fig. 3. Hence, the wind speed can be deduced by the temperature difference measurement of  $T_1$  and  $T_2$  (Makinwa 2004; Huijsing et al. 1984). The temperature sensor is usually used by thermistor, thermopile, diode, and transistor (Zhu et al. 2014).

In one-dimension (1D) wind sensor, the temperature difference  $\delta T$  between  $T_1$  and  $T_2$  is given by Oudheusden (1990):

$$\frac{\delta T}{\Delta T} = c_0 \frac{k_f W}{k_c D} \left( \frac{L^2}{\nu \rho c \alpha} \right)^{\frac{1}{3}} \sqrt{U} \tag{2}$$

**Fig. 4** Schematic of the wind sensor based on time-of-flight principle



where  $\Delta T$  is the overheated temperature with respect to ambient temperature,  $W$  is the chip size of the sensor,  $k_c$  and  $D$  are the thermal conductivity and thickness of the substrate, respectively.  $U$ ,  $k_f$ ,  $\nu$ , and  $\alpha$  are the wind speed, the heat conductivity, the kinematic viscosity, and the thermal diffusivity of the flow, respectively.  $\rho_c$  is the density of the sensor material.  $c_0$  is the constant which equals 0.08076 when the sensing parts are located on the opposite edges of the sensor.

### Time-of-Flight (TOF) Principle

As shown in Fig. 4, for the time-of-flight principle, an electrical pulse is applied to the heater and a thermal pulse is produced. The thermal pulse is taken away by the wind and is detected by the temperature sensor located downstream. Thus, the wind speed can be calculated and given by Wu and Sansen (2002)

$$U = \frac{L}{\Delta t} \quad (3)$$

Where  $L$  is the distance between the heater and the temperature sensor, and  $\Delta t$  is the traveling time of the thermal pulse.

In fact, when arriving at the temperature sensor, the thermal pulse is widen in bandwidth and reduced in amplitude due to thermal diffusion. In order to solve this problem, a novel wind sensor by TOF principle is reported (Durst et al. 2003). A continuous periodic current is applied to the heater and the wind speed is extracted by the time lag measurement between the transmitted and received signal.

### Operation Mode

For MEMS thermal wind sensors, five operation modes are possible: constant power (CP) mode, constant temperature difference (CTD) mode, temperature balance (TB) mode, constant voltage (CV) mode, and constant current (CC) mode.

### **Constant Power Mode**

In CP mode, the power dissipated in the heater is always kept as a constant. For the hot-wire principle, the heat is taken away by the wind and the temperature of the heater decreases. Finally, the wind speed can be detected by the resistance measurement of the heater. For the calorimetric principle, the sensitivity is high at low speed. However, when the speed increases, the output tends saturation and the measurement range is limited. The reason is that the temperature upstream is no less than that in environment while the temperature downstream is no more than that of the heater. Hence, the temperature difference between upstream and downstream is easy saturation.

### **Constant Temperature Difference Mode**

In order to enhance the dynamic range of the measurement, the CTD mode is presented for thermal wind sensor. In this mode, the power applied to the heater is controlled by the feedback circuits, which make the temperature difference between the heater and the surrounding as a constant. For the hot-wire principle, the heat taken away increases with the wind speed, which decreases the temperature of the heater. In order to realize CTD mode, the heating power is increased, which can represent the wind speed. For the sensor based on calorimetric principle, the saturation speed is increased and the measurement range is enhanced compared with the CP mode.

### **Temperature Balance Mode**

In the TB mode, different parts are heated separately and their temperature is kept balance. Consequently, the information of the wind speed and direction can be extracted by the heated power of all parts. In application, the simplest approach is to divide the chip into four symmetrical parts and heat them to a same temperature. Then, the wind information including speed and direction can be obtained by the four different heating power.

### **Constant Voltage Mode**

Obviously, the CV mode means the applied voltage to the heater is a constant. Sometimes, this mode replaces the CP mode to realize wind speed measurement.

### **Constant Current Mode**

The constant current mode, which make the current through the heater constant, is similar to the CV mode and can also be a substitution for CP mode.

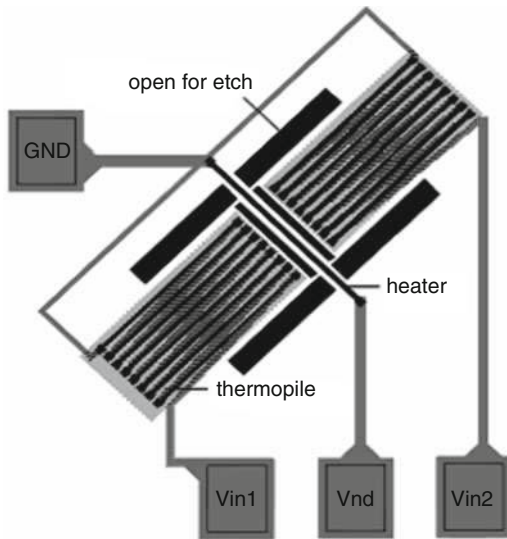
## MEMS Thermal Wind Sensors

MEMS thermal wind sensors are categorized into one-dimension (1D) wind sensors and two-dimension (2D) wind sensors. The 1D wind sensors can only measure bidirectional wind speed, while the 2D can realize speed and direction measurement at the same time.

### 1D Wind Sensors

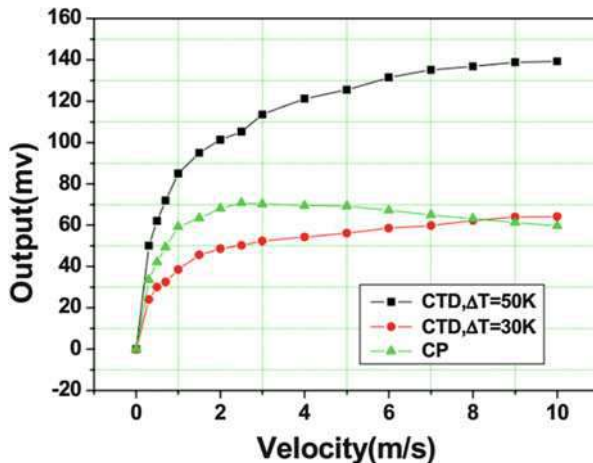
Gao proposed a 1D wind sensor based on calorimetric principle in Fig. 5 (Gao et al. 2004). The polysilicon line is fabricated as the heater. Polysilicon and Al are connected in series as the thermopile to measure temperature difference based on Seebeck effect. In case of no wind, the temperature distribution is symmetrical and the output voltage upstream equals to that downstream. When the wind blows, the output voltage of the thermopile downstream increases and the voltage of the thermopile upstream decreases. Thus, the wind speed can be resolved by two voltage difference measurement. Different from the hot-wire wind sensor, this design can realize bidirectional measurement. Moreover, anisotropic etching is performed on the front surface of the wafer to reduce thermal loss. The experiment shows that the measurement range is about 0–10 m/s in CTD mode, as shown in Fig. 6.

**Fig. 5** SEM photograph of the 1D wind sensor





**Fig. 6** SEM photograph of the 1D wind sensor



## 2D Wind Sensors

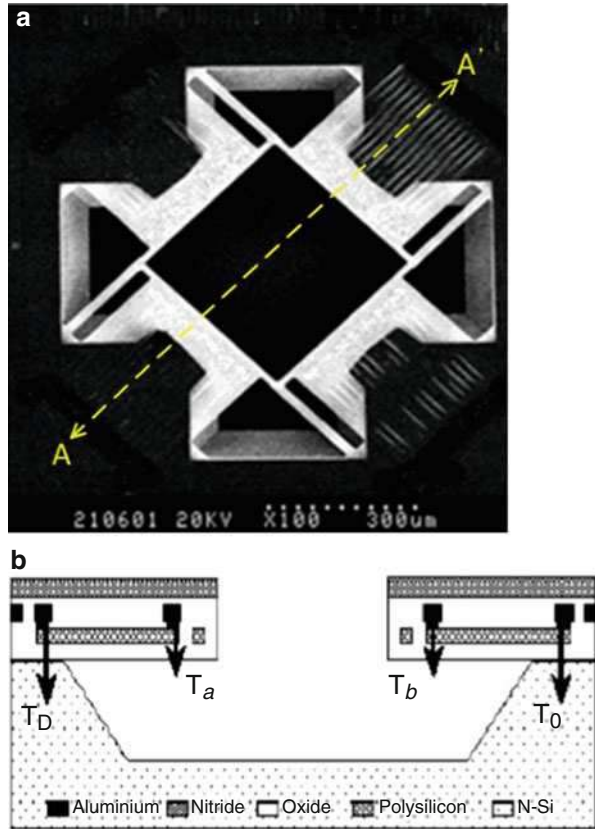
The 2D wind sensors are categorized into three types depending on different physical transduction methods. First, thermoelectric wind sensor uses thermopile to measure temperature profile based on Seebeck effect. Second, thermoresistive wind sensor makes use of thermistors to detect thermal change. In addition, diodes and transistors are applied to measure temperature in thermoelectronic wind sensor. Trigonometric function method and Gauss fitting method are the two most common ways to realize wind direction measurement.

### Thermoelectric Wind Sensor

In 2003, Zhang presented a 2D MEMS wind sensor fabricated by CMOS process (Zhang et al. 2002), as shown in Fig. 7. The heater made of  $N^+$  polysilicon is placed in the center of the sensor chip. Four thermopiles made up of  $P^+$  polysilicon and Al are located in a square around the heater for temperature detection. Both the heater and the thermopiles are fabricated in the silicon dioxide on the surface of the silicon substrate. In order to reduce thermal loss further, the silicon underneath the heater and hot junctions of the thermopiles is removed by front surface etching technology. The operational amplifier circuits are also integrated and the chip area is about  $2 \times 1.6$  mm. As shown in Fig. 8a, the experiment demonstrates that the output voltage is saturated at the speed of 33 m/s in CP mode. A good sine relationship between the voltage and the direction can be obtained in Fig. 8b. The response time is close to 16 ms when the blowing direction changes.

In 2004, Cheng et al. proposed a 2D wind sensor by CMOS technology based on the work of Zhang in Fig. 9 (Cheng et al. 2004). Instead of front surface etching, back etching technology is applied to reduce thermal loss. Moreover, Al strip is fabricated on the silicon dioxide over the heater in order to increase the temperature of the hot junction of the thermopile. By these two methods, the sensitivity can be

**Fig. 7** (a) SEM photograph of the wind sensor by Zhang et al. and (b) Schematic of the cross-section of AA'



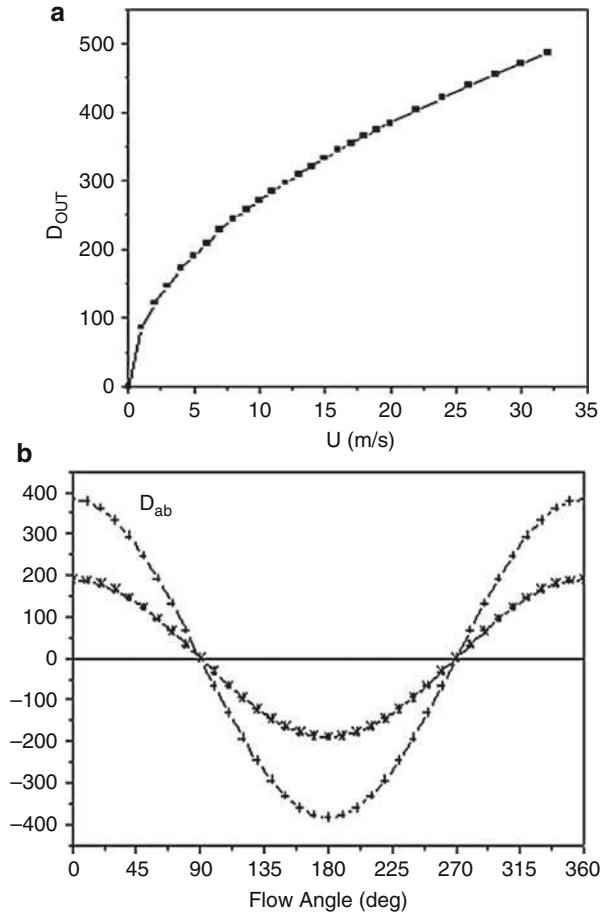
improved. Experiments in Fig. 10 show that this sensor can realize wind speed measurement from 0 to 23 m/s and wind direction measurement of 0–360°.

In 2009, Dong et al. proposed a 2D MEMS wind sensor by standard CMOS MEMS process (Dong 2012). In Fig. 11, polysilicon is fabricated as heater and the thermopiles formed by Al and P<sup>+</sup> polysilicon are used to measure temperature profile resulted by wind. In order to realize CTD mode, a PNP transistor is designed in the center to detect the chip temperature. In Fig. 12a, the wind sensor is tested and the maximum speed is measured to be 40 m/s with the relative error less than 3%. As shown in Fig. 12b, the direction spans from 0° to 360° and the absolute error is less than 3°.

### Thermoresistive Wind Sensors

Figure 13a shows the thermoresistive wind sensor fabricated on silicon substrate (Ye et al. 2017). Ni is used to fabricate heater and thermistors because of high temperature coefficient of resistance (TCR). Al is applied for contacting pads fabrication. In case of no wind, the temperature field is symmetrical and the resistance R1, R2, R3, and R4 is same. When the temperature field is changed by the wind, a resistance of R1 and R3 increases while the other decreases. A Wheatstone bridge including R1 and

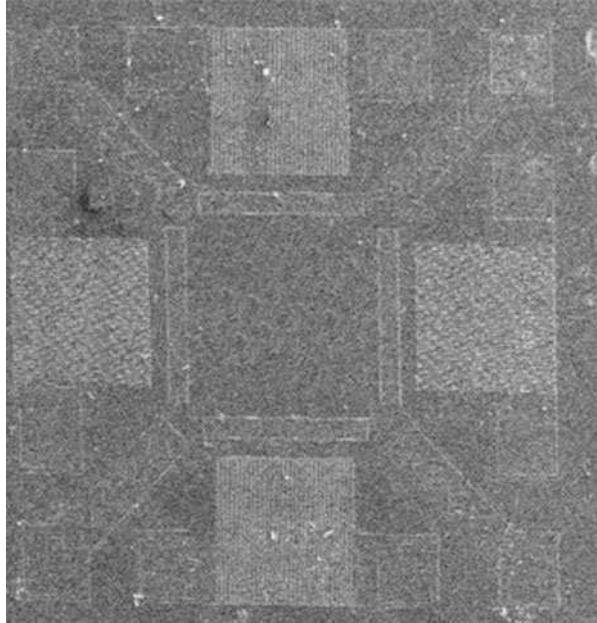
**Fig. 8** Experiments of the wind sensor by Zhang et al.: (a) wind speed and (b) wind direction



R3 is used and the resistance change can be expressed by the output voltage  $V_1$ , as shown in Fig. 13b. Therefore, the wind speed in north-south direction is obtained. Similarly, the resistance change between R2 and R4 can be extracted by  $V_2$ , which shows the wind speed in east-west direction. Hence, the total wind speed is resolved by  $V_1^2 + V_2^2$  and the wind direction can be deduced by  $\tan^{-1}(V_1/V_2)$  based on trigonometric function method. As shown in Fig. 14, this thermoresistive wind sensor exhibits a speed range of 0–30 m/s and direction range of 0–360° in CP mode.

In addition, Shen et al. also presented a thermoresistive wind sensor with different structures on ceramic substrate (Shen et al. 2007). As shown in Fig. 15, one is a traditional square sensor and the other is a circular structure. Pt is fabricated as heater and thermistors on ceramic substrate due to ease of fabrication. The proposed sensor can measure the wind speed from 0 to 10 m/s with the precision of 0.5 m/s. Due to use of ceramic substrate, the power consumption is reduced to less than 50 mW. It should be noticed that the resistance deviation between thermistors cannot be ignored due to poor flatness of ceramic substrate.

**Fig. 9** SEM photograph of the wind sensor by CMOS process



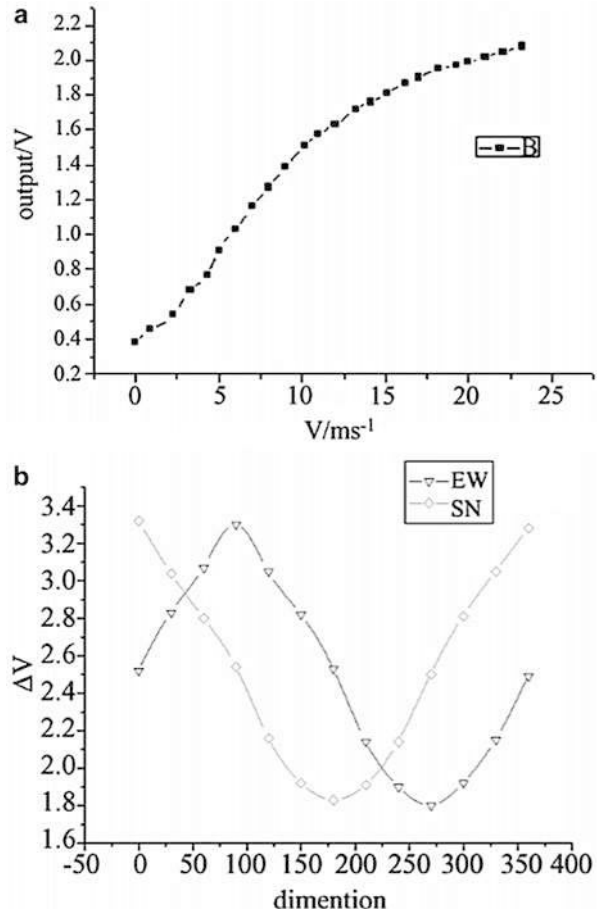
### Thermoelectronic Wind Sensors

For diodes and transistors with a fixed current, the breakover voltage decreases with the temperature linearly. This phenomenon can be used to realize temperature detection, and the diodes and transistors are often applied in thermoelectronic wind sensors. In 2001, Qin et al. reported an octagon-structure wind sensor by CMOS process (Qin et al. 2001), as shown in Fig. 16. A ring made of polysilicon is fabricated as heater. Eight laterally polysilicon diodes (LPD) as four pairs are located around the heater symmetrically. The temperature difference between two LDPs in a pair relies on the wind speed. The experiment shows that the temperature coefficient of LPD is about  $-1.8$  mV/K. The presented sensor can realize measurement range of 0 m/s to about 9.5 m/s from eight different directions, as shown in Fig. 17.

In 2006, Wu et al. improved the design of Qin in Fig. 18 (Wu and Qin 2006). The resistors made of Al serve as the heater and temperature sensor of the chip. Hot-wire principle is applied in CTD mode to measure the wind speed. Moreover, eight resistors made of polysilicon are fabricated around the heater for wind direction measurement. In addition, operational amplifiers and three to eight decoders are included in the chip by CMOS process.

In 2007, Sun et al. proposed a new wind sensor operating on CTD mode (Sun 2006), as shown in Fig. 19. The ring heater is made of polysilicon, and eight transistors are applied to measure temperature distribution. The novelty is that a transistor is fabricated in the center to measure the chip temperature. Therefore, the CTD mode is easily realized and the measuring range is up to 30 m/s, as shown in Fig. 20. At the same time, the temperature of eight transistors is measured and the

**Fig. 10** Measurement of the wind sensor by Cheng: (a) wind speed and (b) wind direction

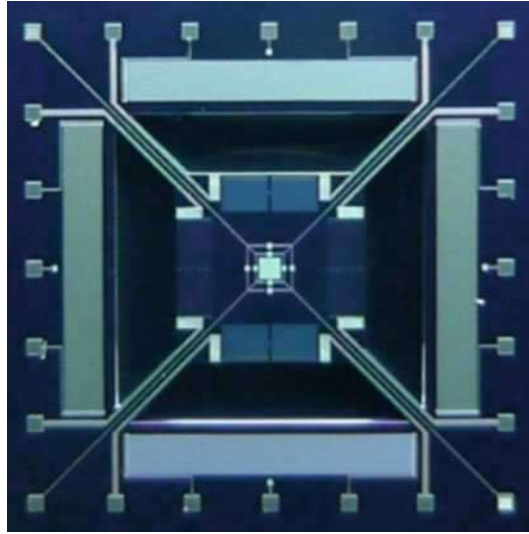


wind direction information can be obtained by fitting based on the Gauss function. Figure 21 shows that the wind direction solution is about 10° and the power consumption is no more than 100 mW.

### Other Wind Sensors

In 2008, Shen et al. proposed a cross-type wind sensor to realize self-test function (Shen et al. 2008, 2010). As shown in Fig. 22, it is composed of cross-structure heaters and four symmetrically located sensing parts. When the flow angle changes clockwise, the temperature differences among the four parts, namely  $T_{13}$  ( $T_{24}$ ) and  $T_{12}$  ( $T_{14}$ ), will be sine (cosine) functions of the wind direction. Further study shows that the amplitudes of  $T_{13}$  and  $T_{24}$  are two times of  $T_{12}$  and  $T_{14}$ , and the phases are 45° before the latter. Then two wind direction values can be obtained for self-test without additional components. As shown in Fig. 23a, the bridge output voltage of the sensor on glass substrate with heating power of 100 mW decrease quickly as the

**Fig. 11** Micro photograph of the wind sensor presented by Dong et al.



wind velocity increases. As shown in Fig. 23b, the measured values at the wind speed of 2 m/s and 5 m/s are well matched with the fitted curve with error less than  $5^\circ$ .

In 2011, Dong et al. developed a 2D wind sensor operating in TB mode (Dong et al. 2011). As shown in Fig. 24, the sensor is composed of four CTD elements. Each element includes one heater and one temperature sensor. The four CTD elements have been designed and implemented to heat the chip and keep itself at a constant overheat temperature higher than the airflow. Meanwhile, the elements are used to balance the temperature difference between the up-stream and down-stream CTD elements by controlling the heating power of the four heaters. The difference of the heating power between the up-stream and down-stream is a measure for the wind. As shown in Fig. 25, the measured results indicate that the accuracies of the wind sensor are in the range of 0.5 m/s and  $\pm 2^\circ$  at wind speed ranging from 0 to 40 m/s and wind direction from  $0^\circ$  to  $360^\circ$ , respectively.

---

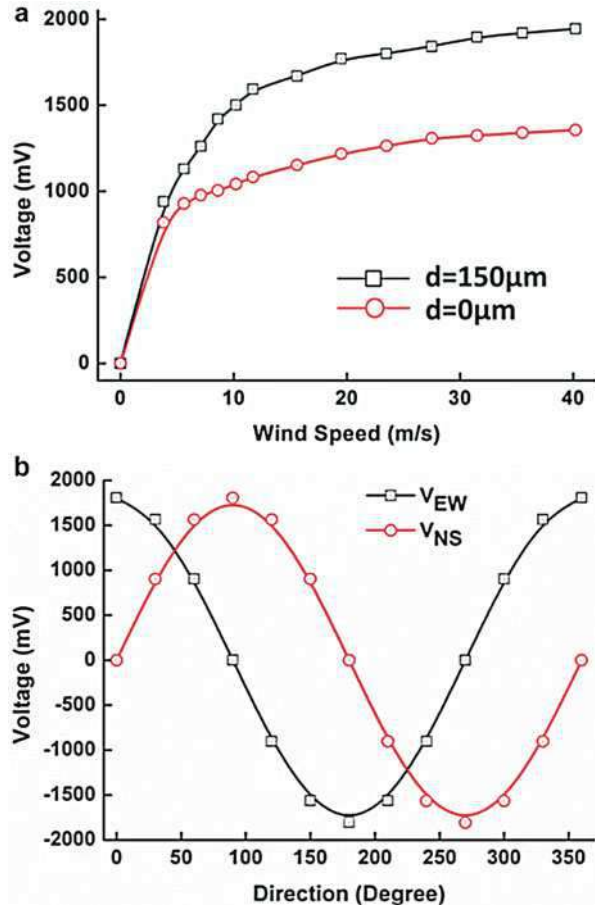
## Sensitivity Improvement and Power Consumption Reduction

In order to obtain high sensitivity and low power consumption, several different methods are always investigated and researched.

### Substrate Etching Technology

For the MEMS wind sensor, the heater was fabricated on the silicon substrate before. It is found that most of the produced heat losses in the substrate due to high thermal conductivity of silicon (about  $150 \text{ W}/(\text{m}\cdot\text{K})$ ). Therefore, the power consumption is large and the sensitivity is low. From Eq. 2, it can be observed that the temperature

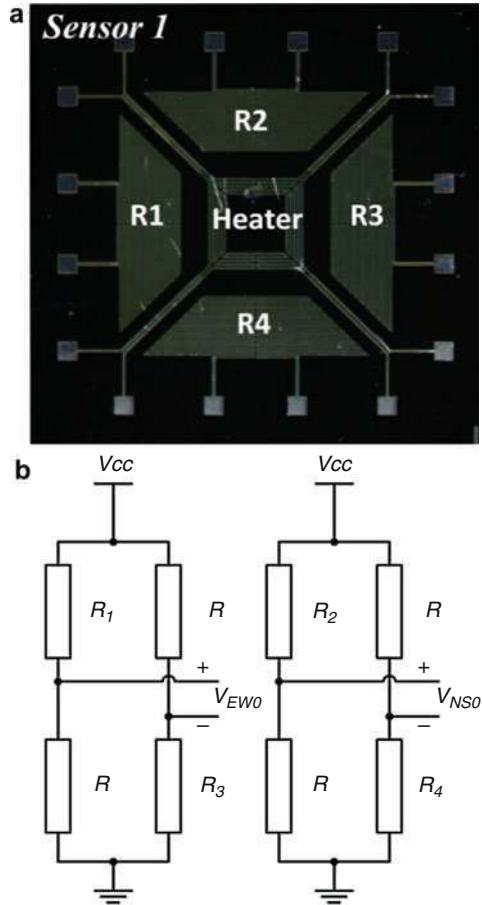
**Fig. 12** Tests of the wind sensor presented by Dong et al.: (a) wind speed test and (b) wind direction test



difference between upstream and downstream is inverse proportional to the substrate thickness. Consequently, for the front surface sensing, the simplest method to solve this problem is to put the wind sensor on the several-micrometer-thickness silicon dioxide or silicon nitride, which is on the surface of the silicon substrate. Then, the substrate is etched to a MEMS membrane and the thermal loss is reduced, as shown in Fig. 26. As a result, most of the heat is interacted with the wind and the sensitivity is improved.

For the back surface sensing, sometimes, trenches are fabricated between the heater and the thermistors, which can reduce the lateral thermal loss, as shown in Fig. 27 (Ye et al. 2017). Figure 28 gives an example of wind sensor with improved sensitivity by trenches technology. The experiments demonstrate that the sensitivity is increased from 9.01 to 16.37 mV/(m/s) at the wind of 3 m/s in CP mode.

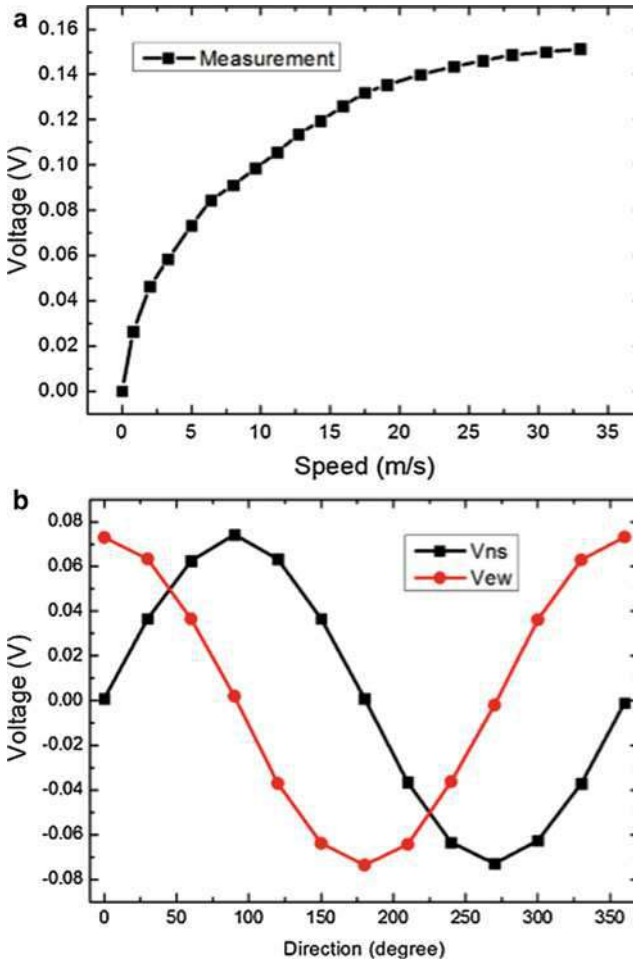
**Fig. 13** (a) Microphotograph of the thermoresistive wind sensor and (b) Wheatstone bridges to measure resistance change



**Substrate with Low Thermal Conductivity**

In Eq. 2, the smaller the thermal conductivity of the substrate is, the larger the temperature difference is. As a result, in addition to substrate etching technology and trench technology, materials with low thermal conductivity are also chosen as the substrate of wind sensor in order to achieve high sensitivity and low power consumption. Except for silicon dioxide and silicon nitride, ceramic and glass are two common candidates as substrate in micromachined thermal wind sensor. The thermal conductivity is about 20 W/(m•K) and 1.4 W/(m•K) for ceramic and glass, respectively. In 2008, 2D wind thermoresistive sensors are fabricated on ceramic and glass substrate (Shen 2009), respectively. As shown in Fig. 29, Pt is lifted off to form the heating and sensing elements. The measured results are recorded in Fig. 30. It is seen that the power consumption is less than 50 mW and the measurement range is

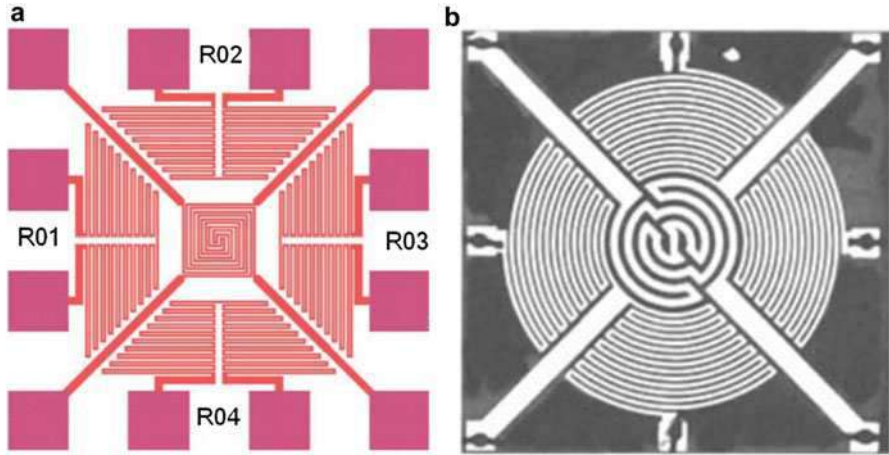




**Fig. 14** Measured results of the thermoresistive wind sensor: (a) wind speed and (b) wind direction

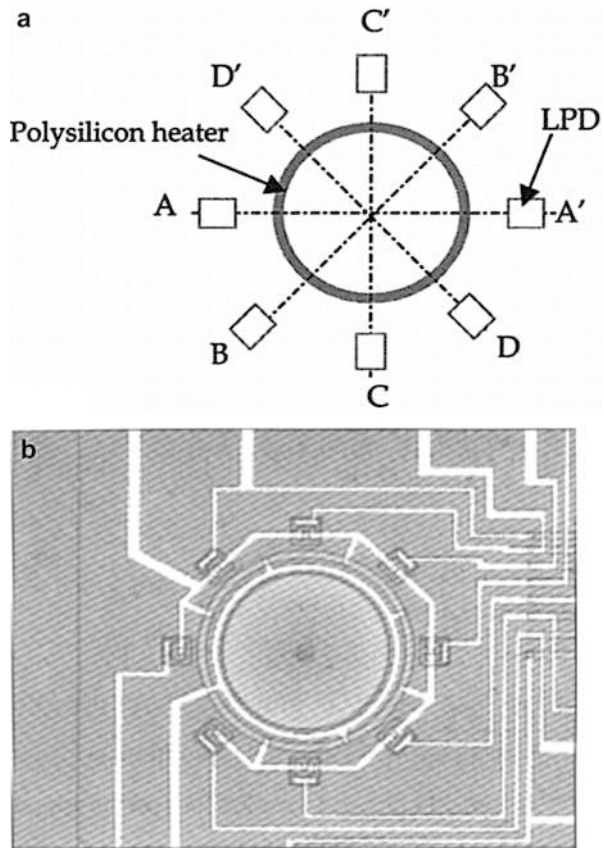
from 0 to 10 m/s. Above all, the sensitivity is improved by ten times compared with the traditional silicon wind sensor. However, the flatness in the surface of ceramic substrate limits its application.

In Fig. 31, in 2014, Gu et al. fabricated a thermoelectric wind sensor on ceramic substrate for the first time (Gu 2012; Zhu et al. 2014). Nickel heaters and four Ni/Al thermopiles are directly fabricated on the front surface of the ceramic plate while the back side is exposed to the wind as the sensing surface. The whole process is simple and the sensor substrate is used directly as a packaging board for sensing wind. The sensor is tested in CP mode. Experimental results in Fig. 32 show that this wind sensor can measure wind speeds up to 30 m/s and wind direction in a full range of 360°.

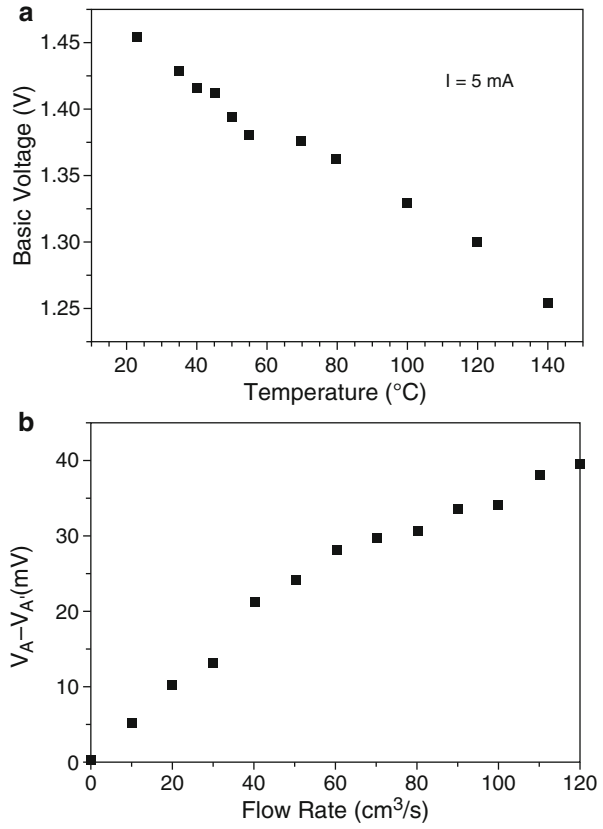


**Fig. 15** (a) Layout of the thermoresistive wind sensor with square structure and (b) micrograph of the wind sensor with circular structure

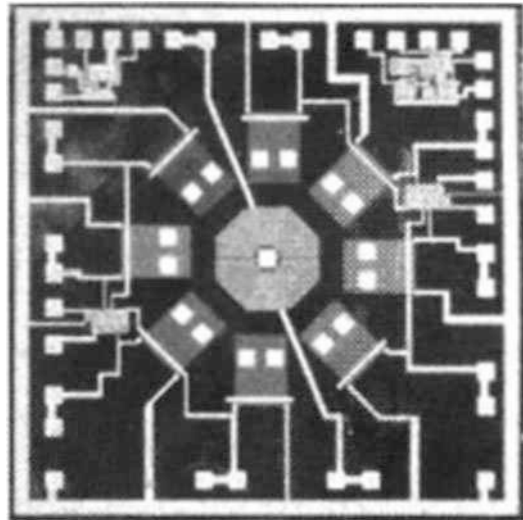
**Fig. 16** Thermoelectronic wind sensor proposed by Qin: (a) schematic overview and (b) micro photograph



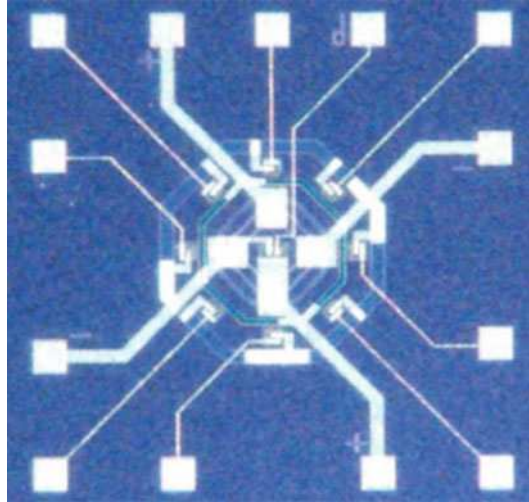
**Fig. 17** Measured results of the thermoelectronic wind sensor: (a) bias voltage change with temperature and (b) voltage difference between LCP A and LCP A'



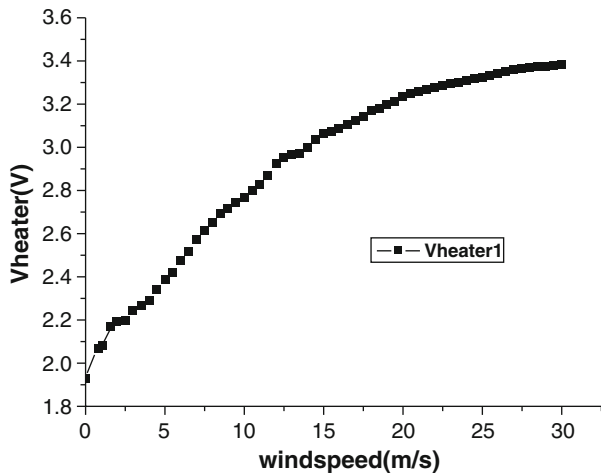
**Fig. 18** Micrograph of the thermoelectronic wind sensor by Wu et al.



**Fig. 19** Micro photograph of the thermoelectronic wind sensor by Sun et al.



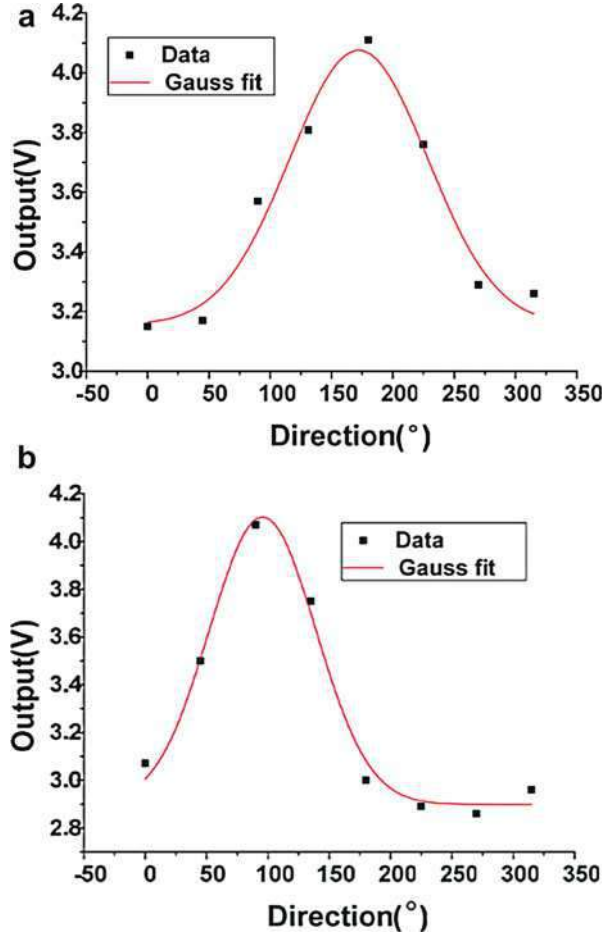
**Fig. 20** Measured heating voltage changing with wind speed up to 30 m/s



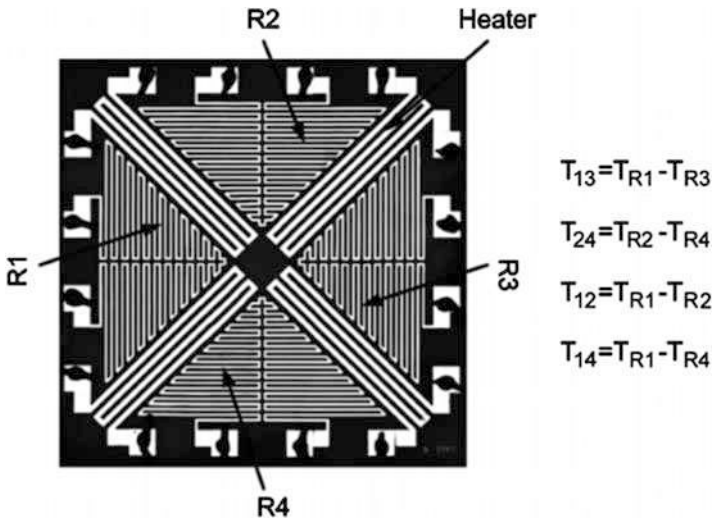
### Full-Bridge Technology

As shown in Fig. 33, in the traditional wind sensor, four thermistors are fabricated to a square around the heater symmetrically (Shen et al. 2008). The thermistor R1 and R3 both act as one of four resistors in a Wheatstone bridge and the voltage is proportional to wind speed. Figure 33b shows this structure named half-bridge structure. Shen et al. designed two thermistors located at different distances from

**Fig. 21** Wind direction fitting by Gauss function: (a) the direction is  $180^\circ$  and (b) the direction is  $90^\circ$



the central heater in each direction, as shown in Fig. 34a. Therefore, four thermistors including R11, R12, R31, and R32 constitute a Wheatstone full bridge. When the wind blows, R31 and R32 increase while R11 and R12 decrease. By the connection way in Fig. 34b, the output voltage can be increased and the sensitivity is improved. Moreover, the measurement range of the wind sensor is also expanded because two thermistors are connected in series in different location to the heater. Figure 35 shows the micro photograph of the wind sensor with full bridge technology. The measured results of wind speed and direction are recorded in Fig. 36a, b, respectively. The measurements demonstrate that the sensitivity is close to  $10 \text{ mV}/(\text{m/s})$  and the range is up to  $35 \text{ m/s}$  based on calorimetric principle.



**Fig. 22** Micro photograph of the wind sensor to realize self-test function by Shen et al.

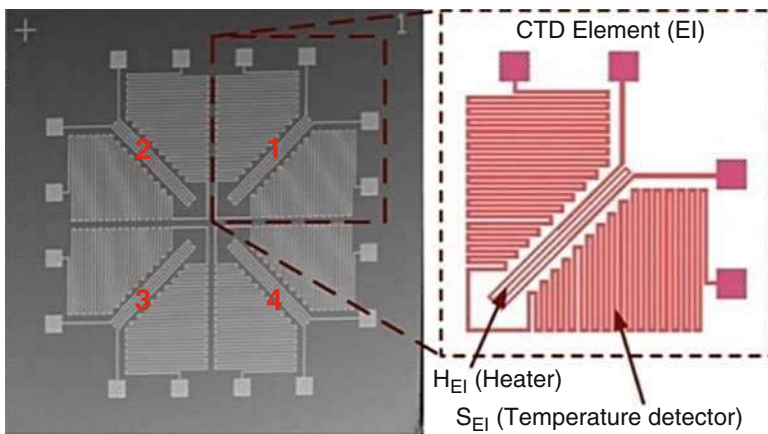
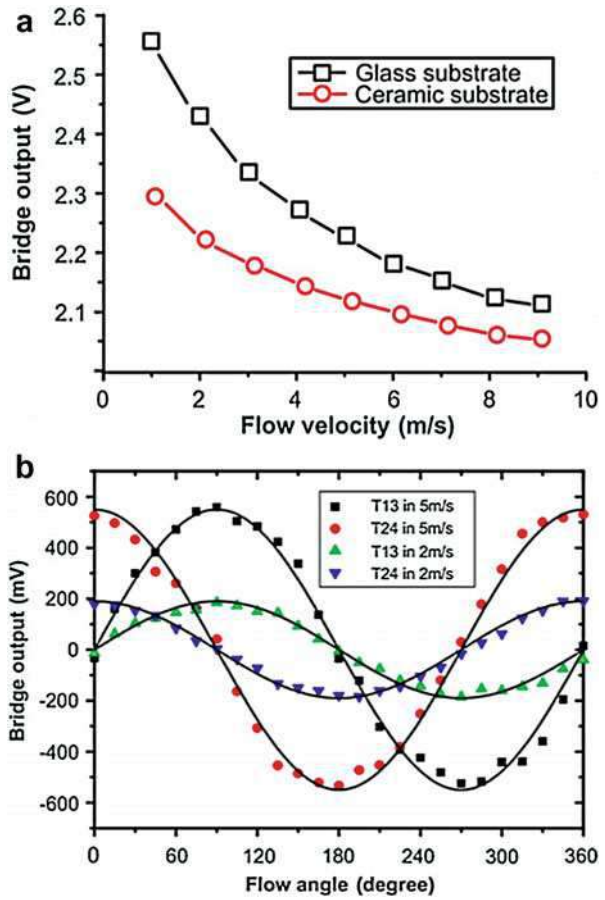
**Other Methods**

Except the methods above, Zhu et al. proposed a novel wind sensor with improved sensitivity and low consumption using back surface sensing (Zhu et al. 2015), as shown in Fig. 37. The silicon with holes is anodic bonded to the glass and reflow is performed with high temperature. Thus, the holes are filled with glass and the silicon is thinned until the glass appears. Heating and sensing resistors are fabricated on the silicon substrate. As a result, the heat conduction is suppressed and the heat convection is enhanced. Figure 38 gives the microphotograph of front surface and back surface of the fabricated sensor. The experimental results are recorded and the power consumption can be reduced to less than 20 mW. As shown in Fig. 39, this wind sensor can measure the speed from 0 to 30 m/s and the direction from 0° to 360°.

On this basis, the power consumption is further reduced by using HF wet etching (Zhu et al. 2016). After wet etching, the thickness of the glass substrate decreases, so that the sensor’s thermal vias become more exposed to the wind, as shown in Fig. 40a. As a result, the conductive heat transfer is weakened and the convective heat transfer is enhanced in sensor operation. Before and after wet etching for 7 and 14 min in Fig. 40b, at the wind speed of 5 m/s, the measured sensitivities of the sensor with a total power consumption of 24.5 mW are 77.2, 98.6, and 164.1 mK/(m/s).

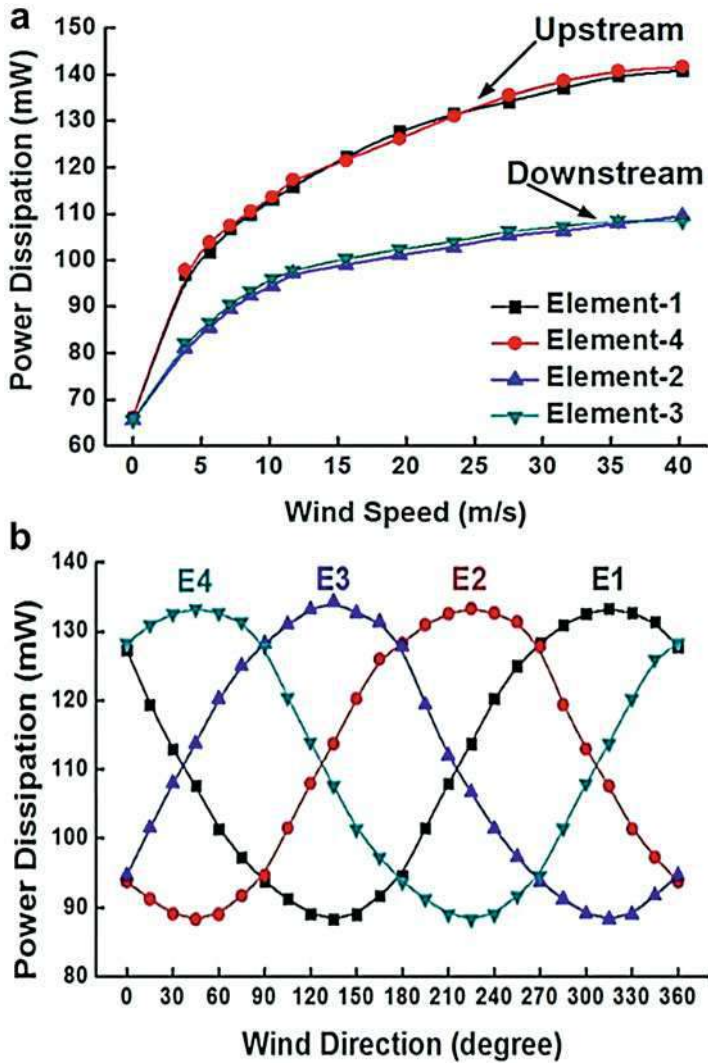
In addition to the methods above, for the thermoelectric wind sensor based on Seebeck effect, the two materials with large difference of Seebeck coefficient will also increase the output voltage. For example, P+ polysilicon reveals positive

**Fig. 23** Experiments of the wind sensor with self-test function: (a) wind speed experiment and (b) wind direction experiment



**Fig. 24** Micro photograph of the wind sensor working in TB mode

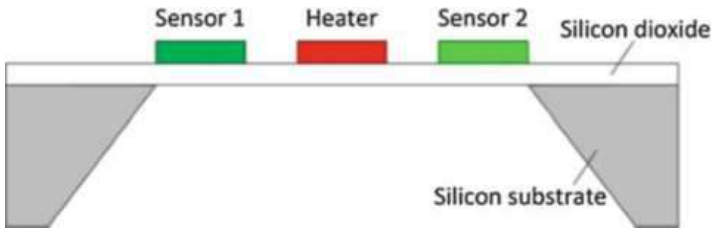




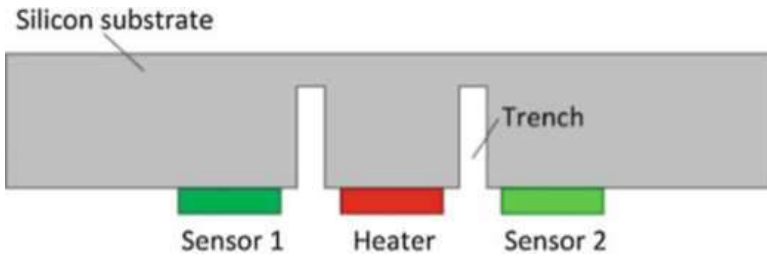
**Fig. 25** (a) Wind speed measurement and (b) direction measurement of the wind sensor in TB mode

Seebeck coefficient while N+ polysilicon has negative Seebeck coefficient. These two materials are combined and a large voltage can be obtained. However, this will increase the difficulty in fabrication. For the thermoresistive wind sensor, the thermistors with high temperature coefficient of resistance, such as Ge, can help to increase the output signal and improve the sensitivity of the sensor. Moreover, square wave signal with a certain duty cycle instead of DC signal can be applied to the heater and low power consumption is achieved.

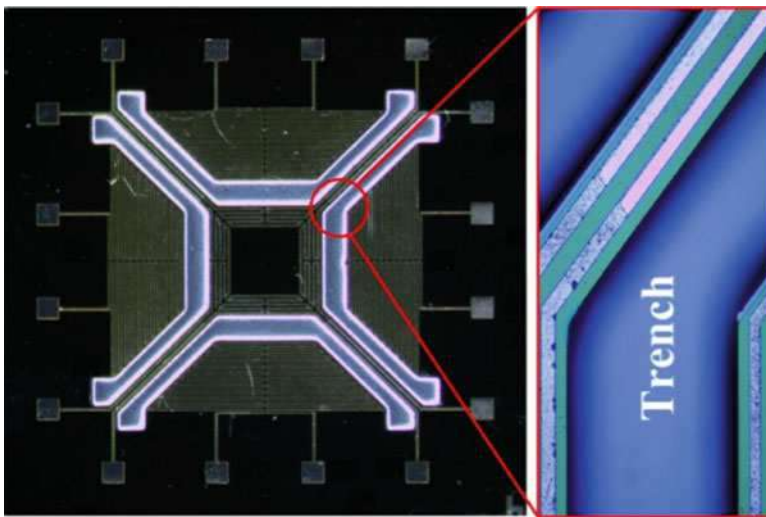




**Fig. 26** Schematic of the wind sensor fabricated on a MEMS membrane by substrate etching technology

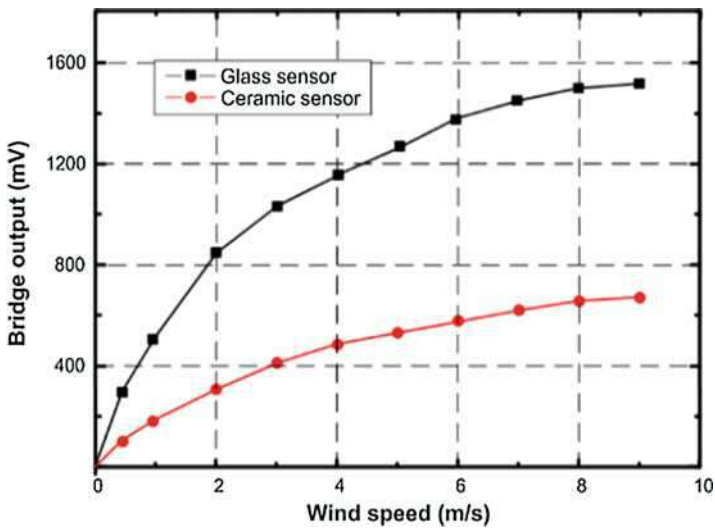


**Fig. 27** Schematic of the wind sensor with trenches between the heater and the thermistors

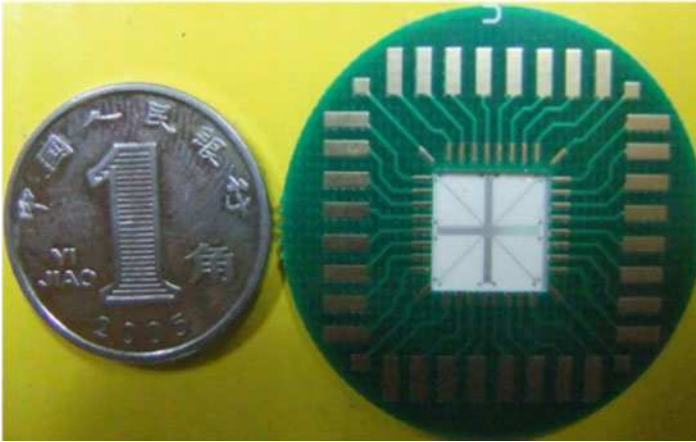


**Fig. 28** Micro photograph of the wind sensor with trenches between the heater and the thermistors

**Fig. 29** Micrograph of the wind sensor fabricated on ceramic substrate

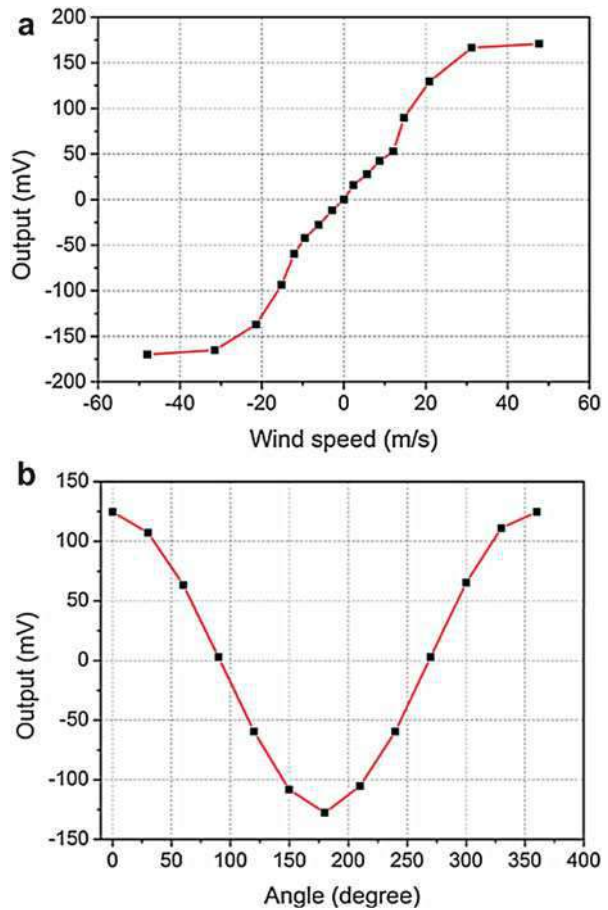


**Fig. 30** Speed measurement of the thermoresistive wind sensor fabricated on ceramic and glass substrate



**Fig. 31** Micrograph of the thermoelectric wind sensor fabricated on ceramic substrate

**Fig. 32** Experiments of the thermoelectric wind sensor fabricated on ceramic substrate: (a) wind speed experiment and (b) wind direction experiment



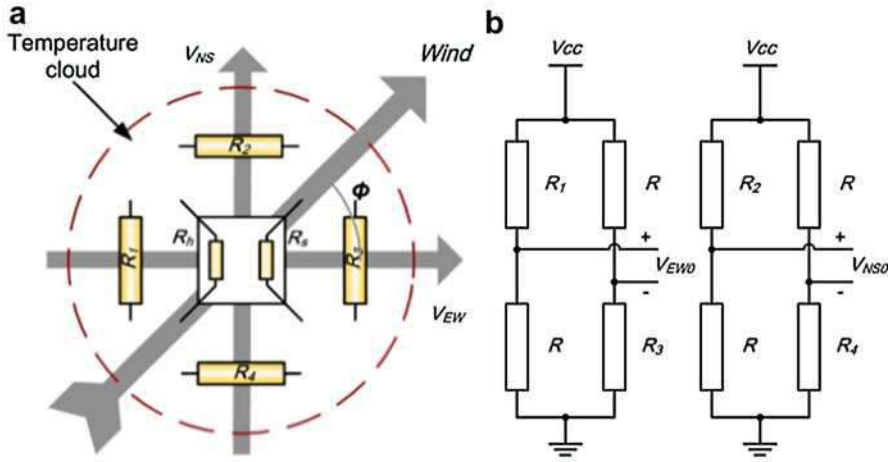


Fig. 33 (a) Schematic diagram of the wind sensor and (b) circuits with half bridge technology

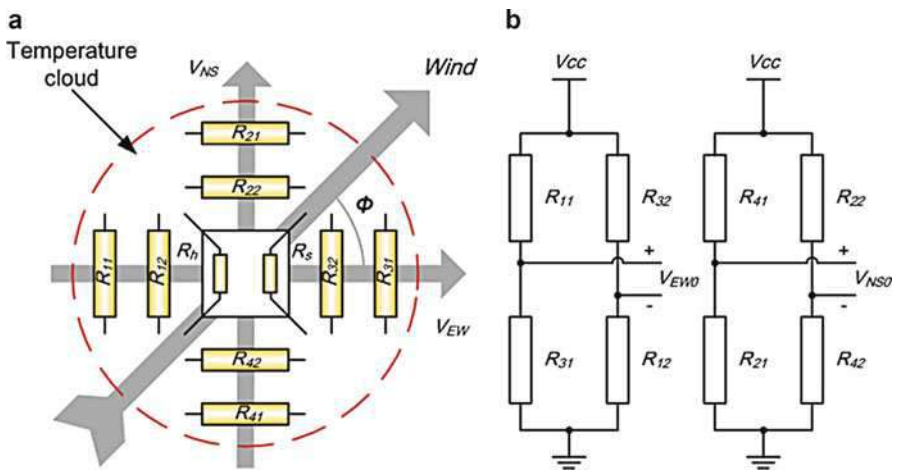
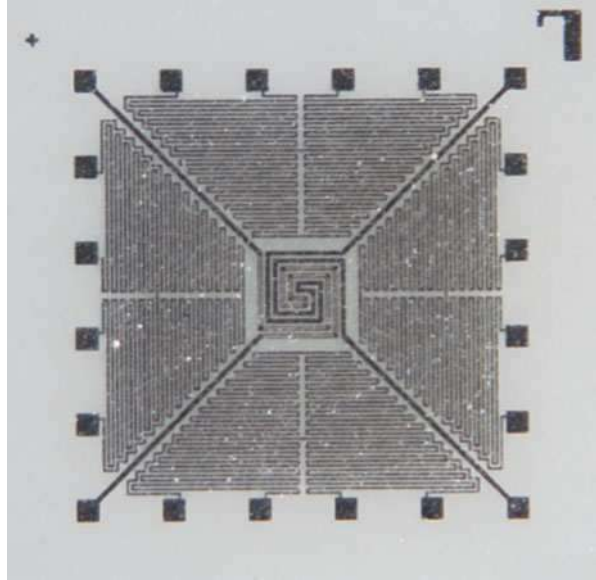


Fig. 34 (a) Schematic diagram of the wind sensor and (b) circuits with full bridge technology

## Package

Different from other flow sensors, the wind sensor is placed outside and is easy to be polluted by particles and rain in the air. Therefore, package is necessary and three different package methods are widely used in MEMS thermal wind sensor.

**Fig. 35** Micro photograph of the wind sensor with full bridge technology on ceramic substrate



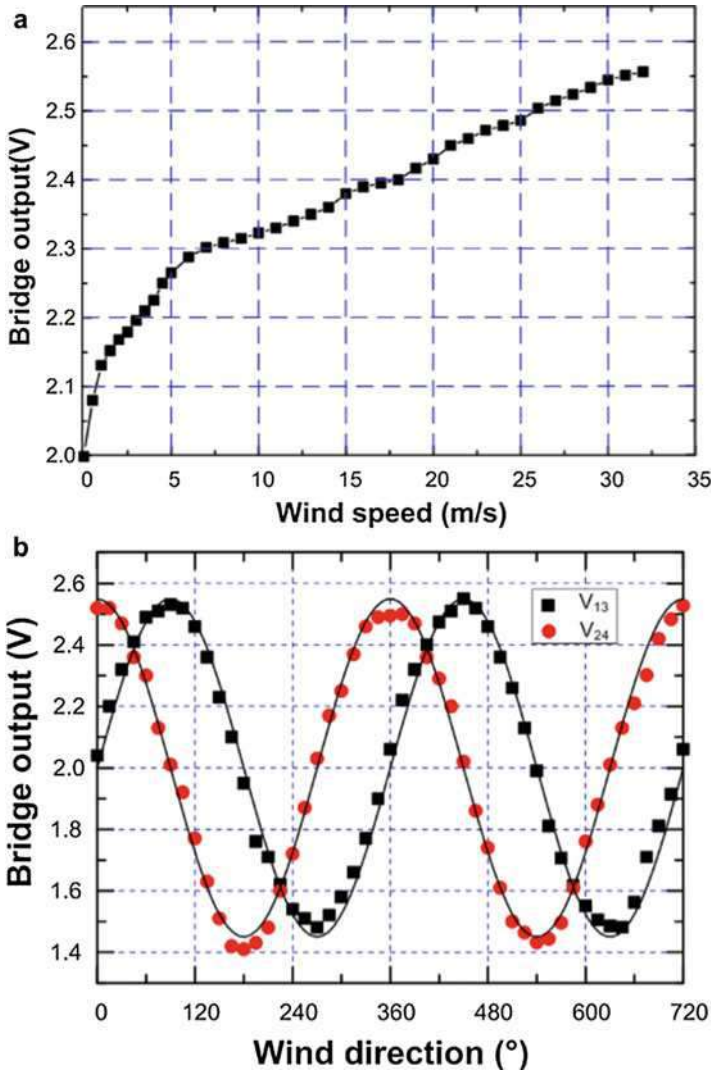
### Direct Chip Attach Package

The simplest method is direct chip attach (DCA) package. As shown in Fig. 41a, the front surface of the chip is glued to one side of a FR4 substrate while the back surface is used to sense wind change (Shen et al. 2008). The contacting pads in the chip are bonded by wires and connected to processing circuits in the FR4 substrate. Finally, encapsulant such as epoxy resin is used to protect the sensor and the bonding wires. Figure 41b shows the schematic and microphotograph of the wind sensor by DCA package, respectively.

### Flip Chip Package

The second package is flip chip (FC) package. As shown in Fig. 42a, Au and Pb/Sn bumps are fabricated on chip and FR4 substrate (Sun et al. 2010), respectively. Then, these two kinds of bumps are welded using a vacuum chuck at about 240 °C. In order to enhance the mechanical strength, epoxy-based underfill is applied. Figure 42b gives the wind sensor with Au bumps and FR4 substrate with Pb/Sn bumps.

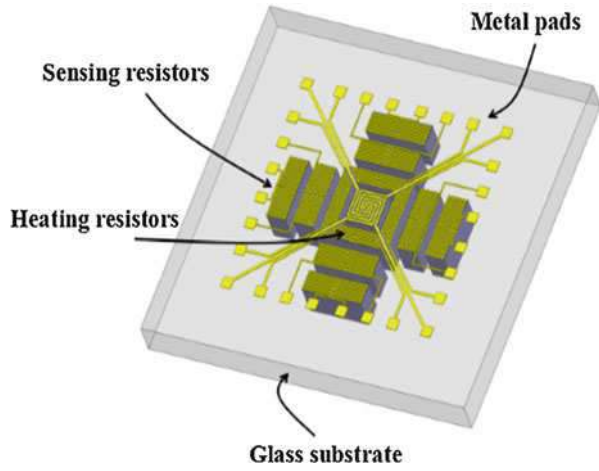
Additionally, Sun et al. proposed a FC package to ceramic substrate (Sun et al. 2005, 2007), as shown in Fig. 43a. The sensor chip was fabricated by standard CMOS technology, and it is FC packaged on a thin ceramic substrate using copper pillar bump technology. In Fig. 43b, heat transfer is performed between the heaters



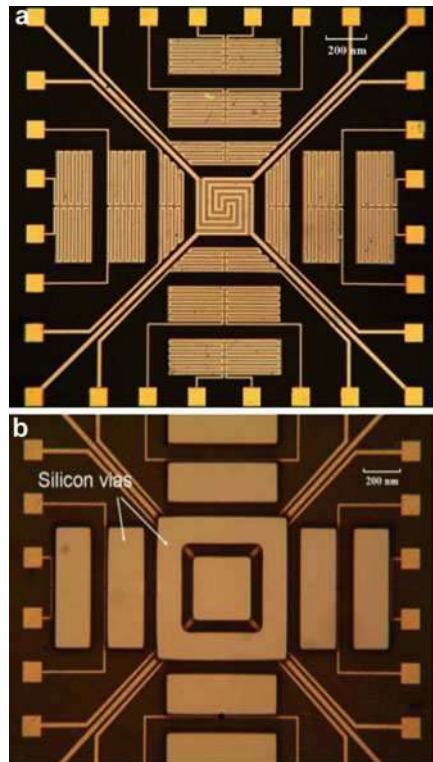
**Fig. 36** Measurement of the wind sensor with full bridge technology: (a) speed measurement and (b) direction measurement

and the ceramic substrate via the pillar bump. The backside of the ceramic substrate provides a smooth surface for the sensor to contact with the flow. The change of flow-induced temperature distribution on the flow sensing surface is measured by temperature sensors. Meanwhile, the ceramic substrate holds the sensor chip and protects it from being contaminated or even destroyed in harsh environment.

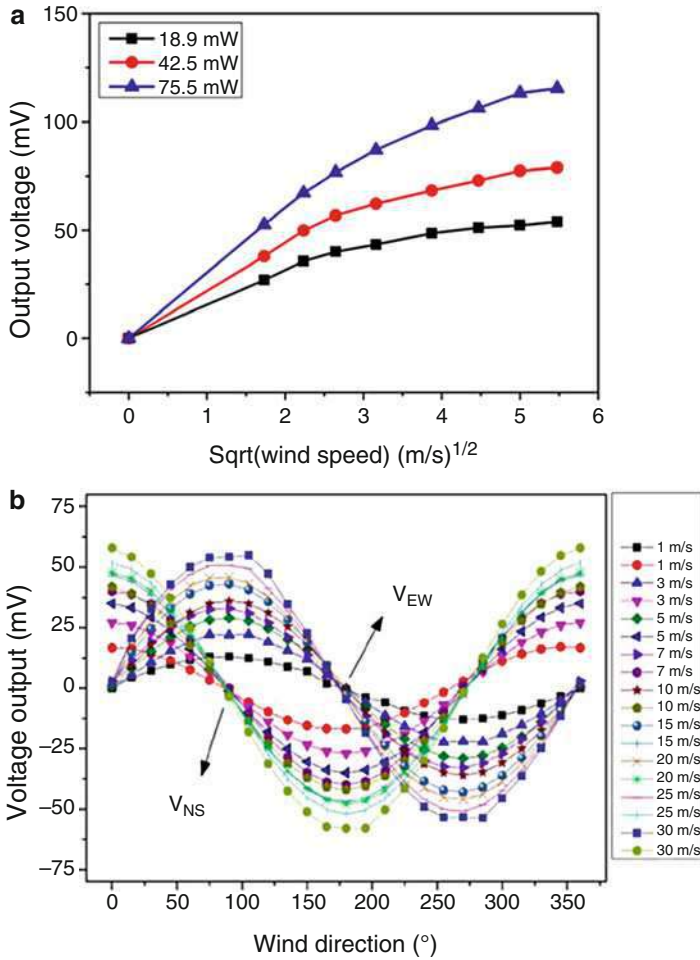
**Fig. 37** Schematic diagram of the wind sensor by Zhu et al.



**Fig. 38** Microphotograph of the wind sensor by Zhu: (a) front surface and (b) back surface







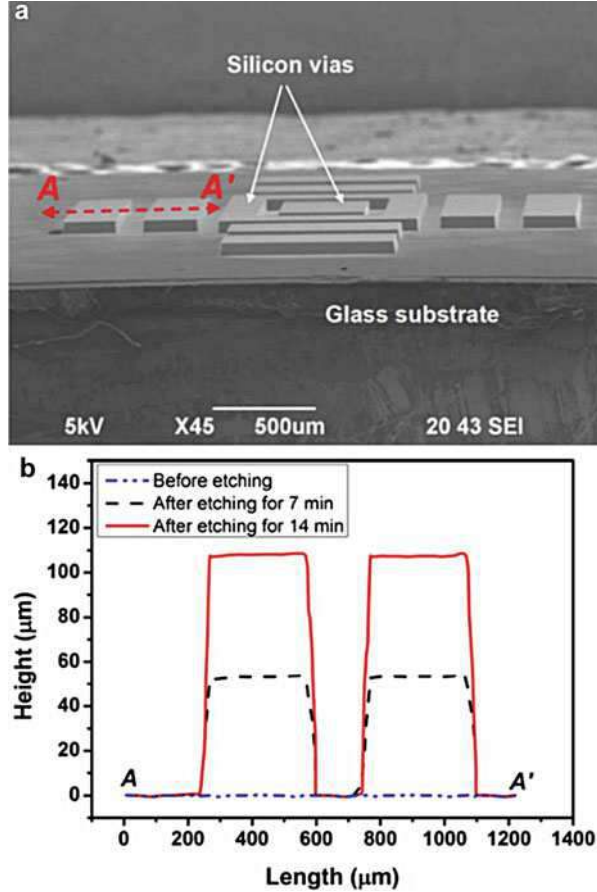
**Fig. 39** Measured results of the wind sensor by Zhu: (a) wind speed and (b) wind direction

### Au-Au Bonding Package

In 2012, Dong et al. proposed a novel package method by Au-Au wafer bonding technology (Dong et al. 2011, 2012). As shown in Fig. 44a, the sensing parts are fabricated on the silicon substrate using standard IC process, while the heating parts are designed on the ceramic substrate. Au bumps are fabricated on these two substrates and the thickness is about 5  $\mu\text{m}$ . The Au-Au bonding is performed at 380  $^{\circ}\text{C}$



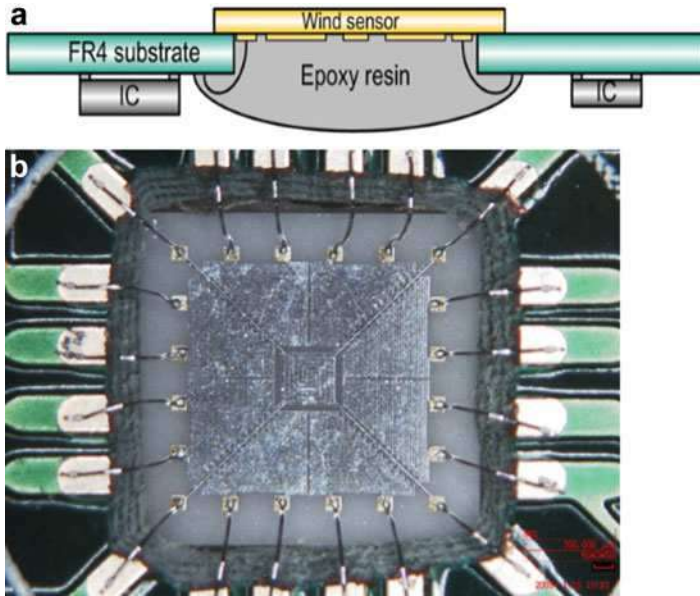
**Fig. 40** (a) The SEM photograph of the wind sensor after wet etching. (b) Step height lines of the sensor surface in different etching time by step profiler



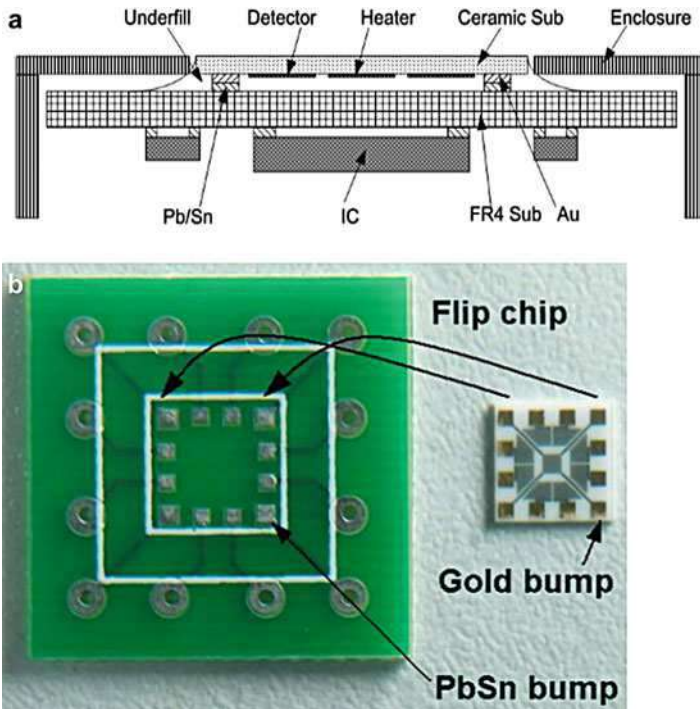
for about 30 min. One kind of the Au bumps is used for electric connection between silicon substrate and ceramic substrate while the other kind is used for thermal connection between two substrates. After that, the redundant silicon is removed by the wet etching, as shown in Fig. 44b. Figure 45 presents the microphotograph of the wind sensor before and after Au-Au bonding package.

## Environmental Effect

The wind sensor is placed outside and easily affected by the environment although the package is performed. Obviously, ambient temperature has an effect on the wind sensor due to its thermal principle.

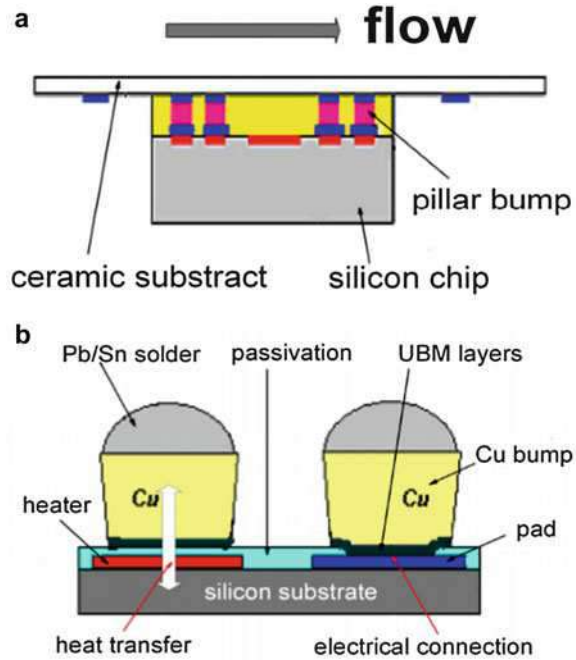


**Fig. 41** (a) Schematic of the DCA package and (b) microphotograph of the wind sensor after DCA package



**Fig. 42** (a) Schematic overview of the FC package and (b) photo of the wind sensor and FR4 substrate by FC package

**Fig. 43** (a) Schematic overview of the FC package to ceramic and (b) copper pillar bump structure



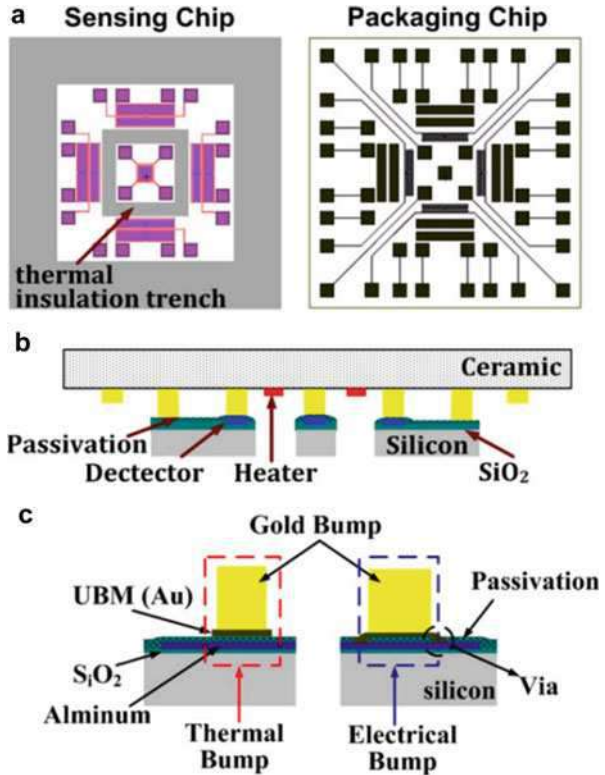
## Temperature Effect

Huang et al. researched the temperature effect on the micromachined thermal wind sensor (Huang et al. 2015; Chen et al. 2015). Figure 46 displays the measured output voltage change in 10 m/s when the ambient temperature increases from  $-20\text{ }^{\circ}\text{C}$  to  $40\text{ }^{\circ}\text{C}$ . A linear curve is fitted and the slope is close to  $1.25\text{ mV/k}$ . In addition, a theoretical model is established and temperature compensation is performed. Figure 47 shows the wind speed measurement result with temperature compensation using the proposed model. As can be observed, after compensation, the relative error of the wind speed measurement at different ambient temperatures is less than  $\pm 3.5\%$ .

## Humidity Effect

Humidity effect is also investigated by Chen et al. in 2014 (Chen et al. 2014). The experimental results are obtained by testing the sensor in sunny day and rainy day. Figure 48 shows the output voltage for the relative humidity of 52% and 70% at  $23\text{ }^{\circ}\text{C}$ . The experimental results indicate that the output voltage decreases with the

**Fig. 44** (a) Schematic overview of layout of two chips, (b) cross-section of the Au-Au bonding technology and (c) gold bumps for thermal exchange and electrical connection

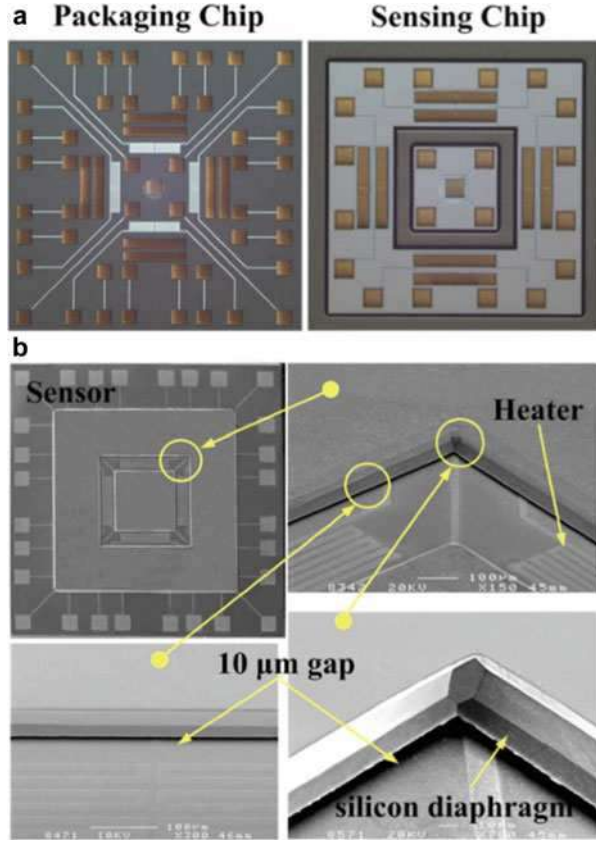


increase of the humidity. The effect of relative humidity on the output is less than 1%, which is not obvious compared with the temperature effect.

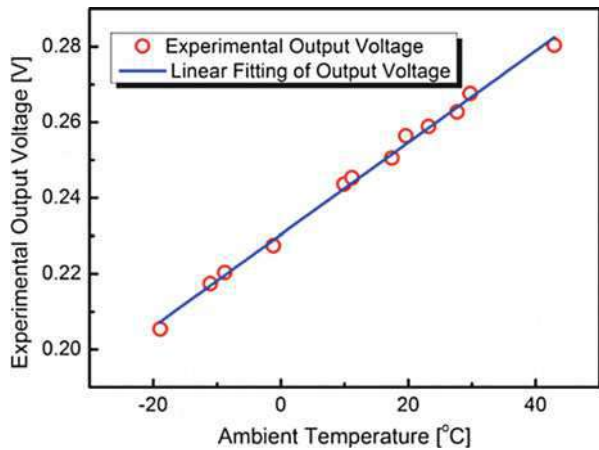
## Conclusion

Micromachined thermal wind sensor is introduced in this chapter. Different principles and operation modes are addressed. Several typical wind sensors including 1D and 2D devices are given and the related measurements are also presented. The wind speed measuring dynamic range exceeds 0–40 m/s with the direction from 0° to 360°. Several methods are considered and analyzed, and the power consumption can be reduced to less than 20 mW. In application, the MEMS wind sensor can be packaged using various packages such as DCA package, FC package, and Au-Au bonding package. Despite this, the wind sensor is easy to be affected by the surrounding temperature. Related measurements are given and the temperature effect is modeled.

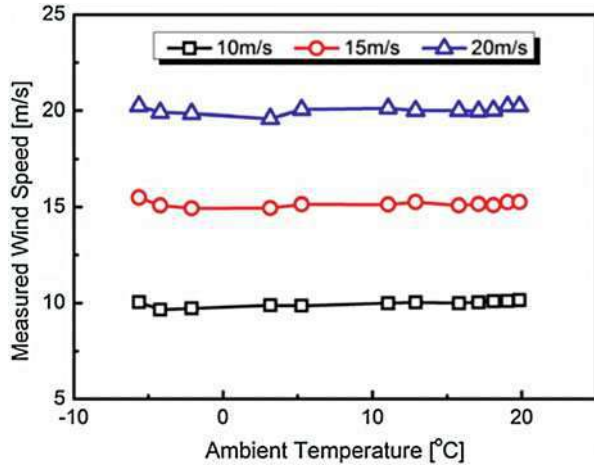
**Fig. 45** Microphotograph of the wind sensor of (a) before and (b) after Au-Au bonding package



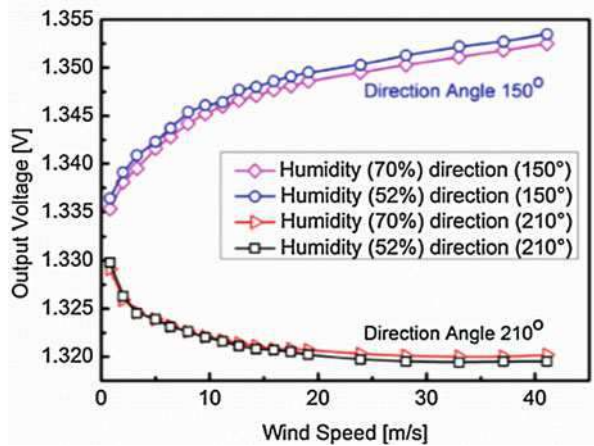
**Fig. 46** Measurement of temperature effect on MEMS wind sensor in the wind speed of 10 m/s



**Fig. 47** Measured wind speed of the sensor in different ambient temperatures with temperature compensation



**Fig. 48** Measurement of humidity effect on the MEMS wind sensor



**References**

Chen B, Zhu Y, Qin M et al (2014) Effects of ambient humidity on a micromachined silicon thermal wind sensor. *J Microelectromech Syst* 23(2):253–255

Chen B, Zhu Y, Yi Z et al (2015) Temperature effects on the wind direction measurement of 2D solid thermal wind sensors. *Sensors* 15(12):29871–29881

Cheng H, Qin M, Gao D (2004) Thermal-film temperature-difference-based CMOS wind sensor design and realization. *Chin J Electron Devices* 27(3):486–489

Dong Z (2012) Study on MEMS wind sensors and control system. PhD dissertation, Southeast University

Dong Z, Chen J, Qin Y et al (2011) A hot film wind sensor with four constant temperature difference elements fabricated on ceramic substrate. In: *Proceeding of IEEE sensors*. Limerick. 28–31 Oct 2011

Dong Z, Chen J, Qin Y et al (2012) Fabrication of a micromachined two-dimensional wind sensor by Au-Au wafer bonding technology. *J Microelectromech Syst* 21(2):467–475

- Durst F, Al-Salaymeh A, Bradshaw P (2003) The development of a pulsed-wire probe for measuring flow velocity with a wide bandwidth. *Int J Heat Fluid Flow* 24:1–13
- Gao D, Qin M, Cheng H (2004) Optimization and test of a calorimetric silicon flow sensor. In: International conference on information acquisition, Hefei. 21–25 June 2004
- Giani A, Mailly F, Bonnot R (2002) Thermal model of thin film anemometer. *Microelectron J* 33(8):619–625
- Gu F (2012) A study to the thermopile ceramic wind sensor. Master's thesis, Southeast University
- Huang Q, Chen B, Zhu Y et al (2015) Modeling of temperature effects on micromachined silicon thermal wind sensors. *J Microelectromech Syst* 24(6):2033–2039
- Huijsing JH, Schuddemat JP, Verhoef W (1984) Monolithic integrated direction-sensitive flow sensor. *IEEE Trans Electron Devices* 29(1):57–60
- Kato N, Ohkuma T, Kim JR (1992) Full scale measurements of wind velocity in two urban areas using an ultrasonic anemometer. *J Wind Eng Ind Aerodyn* 41(1):67–78
- King LV (1914) On the convection of heat from small cylinders in a stream of fluid. *Philos Trans R Soc Lond Ser A* 214:373–432
- Kuo JTK, Yu L, Meng E (2012) Micromachined thermal flow sensors—a review. *Micromachines* 3:550–573
- Makinwa KKA (2004) Flow sensing with thermal sigma-delta modulators. PhD dissertation, Delft University of Technology
- Oudheusden BWV (1988) Silicon flow sensors. *Control Theory Appl IEE Proc DIET* 135(5):373–380
- Oudheusden BWV (1990) Silicon thermal flow sensor with a two-dimensional direction sensitivity. *Meas Sci Technol* 1(7):565–575
- Qin M, Huang Q, Zhang Z et al (2001) A novel micromachined thermal anemometer based on laterally polysilicon diode flow sensor. In: Proceedings of the SPIE 4601, micromachining and microfabrication process technology and devices, 15 Oct 2001
- Shen G (2009) System integration and packaging of thermal wind sensor. PhD dissertation, Southeast University
- Shen G, Wu J, Zhang H et al (2007) Design and test of 2-D MEMS wind speed and direction sensor. *Micronanoelectron Technol* 7(8):285–287
- Shen G, Qin M, Huang Q (2008) A smart 2-D wind sensor with self-test function. In: Proceeding of IEEE sensors 2008, Lecce. 26–29 Oct 2008
- Shen G, Qin M, Huang Q (2010) A cross-type thermal wind sensor with self-testing function. *IEEE Sensors J* 10(2):340–346
- Sun J (2006) Design and package of CMOS thermal flow sensors. Master's thesis, Southeast University
- Sun J, Qin M, Huang Q (2005) A flip-chip packaged CMOS thermal flow sensor. In: Proceeding of IEEE sensors, Irvine. 30 Oct–3 Nov 2005
- Sun J, Qin M, Huang Q (2007) A flip-chip packaged CMOS thermal flow sensor. *IEEE Sensors J* 7(7):990–995
- Wu J, Qin M (2006) Study on integrated controlling and measuring circuits of the two-dimensional CMOS silicon wind sensor. *Chin J Electron Devices* 29(1):95–101
- Wu J, Sansen W (2002) Electrochemical time of flight flow sensor. *Sensors Actuators A Phys* 97:68–74
- Ye Y, Yi Z, Qin M et al (2017) DRIE trenches and full-bridges design for sensitivity improvement of MEMS silicon thermal wind sensor. In: Proceeding of 30th IEEE international conference on micro electro mechanical systems, Las Vegas, 22–26 Jan 2017
- Zhang Z, Qin M, Huang Q (2002) Novel CMOS two-dimensional integrated gas flow sensor. In: Proceedings of SPIE, micromachining and microfabrication process technology and devices, Oct 2001
- Zhu Y, Chen B, Qin M et al (2014) 2D micromachined thermal wind sensors—a review. *IEEE Internet Things J* 1(3):216–230
- Zhu Y, Chen B, Qin M et al (2015) Development of a self-packaged 2D MEMS thermal wind sensor for low power applications. *J Micromech Microeng* 25:055011
- Zhu Y, Qin M, Huang J et al (2016) Sensitivity improvement of a 2D MEMS thermal wind sensor for low-power applications. *IEEE Sensors J* 16(10):4300–4308



# Integrated Vacuum Microsensor Systems in CMOS Technology

Jiaqi Wang and Zhenan Tang

## Contents

Fundamentals of Vacuum Sensing .....	578
Vacuum Technology and Its Applications .....	578
Development of Vacuum Sensor .....	579
Microfabricated Vacuum Sensors .....	579
Pirani Vacuum Sensors .....	580
Ionization Vacuum Sensors .....	581
Key Fabrication Process for Micro-Vacuum Sensors .....	582
CMOS Platform Technology for Vacuum Sensors .....	584
CMOS Pirani Vacuum Sensors .....	584
CMOS Ionization Vacuum Sensors .....	588
CMOS Vacuum Multisensor Systems .....	591
Key Fabrication Process for CMOS Vacuum Sensors .....	591
Outlook and Future Development .....	593
References .....	593

## Abstract

Vacuum sensor is a type of pressure sensor and widely used in the vacuum measurement. In this chapter, besides the traditional vacuum sensors, the micro-vacuum sensors based on silicon micromachining technique are introduced. What is more, the CMOS-compatible Pirani and thermal ionization vacuum sensors are discussed in details. The CMOS-compatible process is important for sensor fabrication, including suspending structure, wet etching, dry etching, etc. The influence of sensor to the circuit should also be considered, such as the thermal influence. The

J. Wang (✉)

Institute of Microelectronics, Dalian University of Technology, Dalian, China  
e-mail: [wjq@dlut.edu.cn](mailto:wjq@dlut.edu.cn)

Z. Tang

School of Microelectronics, Dalian University of Technology, Dalian, China  
e-mail: [Tangza@dlut.edu.cn](mailto:Tangza@dlut.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,  
[https://doi.org/10.1007/978-981-10-5945-2\\_10](https://doi.org/10.1007/978-981-10-5945-2_10)

577



sensor system can integrate the sensor with voltage reference, signal sampling circuit, A/D converter (ADC), central processing unit (CPU), and interface circuit. This technique will generate the vacuum sensor system with sensor and its control system monolithically, which is convenient for applications. The integrated vacuum sensor system has the advantages of small size, volume, and weight, which would have its potential applications, such as the space exploration, micro-packaging, etc.

### Keywords

Vacuum · Sensor · Vacuum sensor · Pirani sensor · Ionization sensor · Sensor system · CMOS sensor · Micromachining · Mems · CMOS compatible

## Fundamentals of Vacuum Sensing

### Vacuum Technology and Its Applications

Vacuum is defined as the physical environment below the atmospheric pressure. It also represents the gas molecular and gas pressure. Vacuum technology adopts some physical or chemical methods to establish, measure, and utilize vacuum. Vacuum technology has developed with a long history. Modern vacuum technology has been 400 years since Torricelli discovered the vacuum phenomenon in 1644. Especially in the twentieth century, the vacuum technology developed rapidly and is applied widely in industry. Since a long history of vacuum technology in different countries, many units for vacuum are produced. The SI unit of pressure is the pascal, which is defined as one newton per square meter ( $1 \text{ Pa} = 1 \text{ N/m}^2$ ). There are many non-SI pressure units, including the torr and mbar ( $1 \text{ torr} = 133.322 \text{ Pa}$ ,  $1 \text{ mbar} = 100 \text{ Pa}$ ). This chapter retains the units of vacuum in the references and converts them to the international unit pascal for convenience of reading. Vacuum is divided into low, medium, high, and ultrahigh vacuum with different gas pressures, as shown in Table 1 (Abbott and Jabour 2011). Perfect vacuum is an ideal state of no particles at all. Even if all particles of matter were removed, there would still be photons and gravitons, as well as dark energy, virtual particles, and other aspects of the quantum vacuum.

**Table 1** Pressure ranges of each quality of vacuum

Vacuum quality	Pa
Atmospheric pressure	$10^5$
Low vacuum	$1 \times 10^5$ to $3.3 \times 10^3$
Medium vacuum	$3.3 \times 10^3$ to $1 \times 10^{-1}$
High vacuum	$1 \times 10^{-1}$ to $1 \times 10^{-7}$
Ultrahigh vacuum	$1 \times 10^{-7}$ to $1 \times 10^{-10}$
Extremely high vacuum	0 to $1 \times 10^{-10}$
Perfect vacuum	0

The most important application of vacuum is in various industries. It can be widely used in vacuum metallurgy, crystal growth, chemical industry, food packaging, semiconductor technology, etc. For example, in semiconductor industry, many fabrication processes must be implemented in vacuum environment, such as plasma etching, chemical and physical vapor deposition (CVD and PVD), ion implantation, and EUV lithography.

As we know, outer space of the Earth is a vacuum; if human wants to go out of the Earth to explore the universe, vacuum technology should be developed first, which prompts the application of vacuum technology in aerospace industry (Johnson et al. 2011). Another important application of vacuum technology is in scientific research area, such as life science, analysis of materials, accelerator, and LIGO (Laser Interferometer Gravitational Wave Observatory) detector.

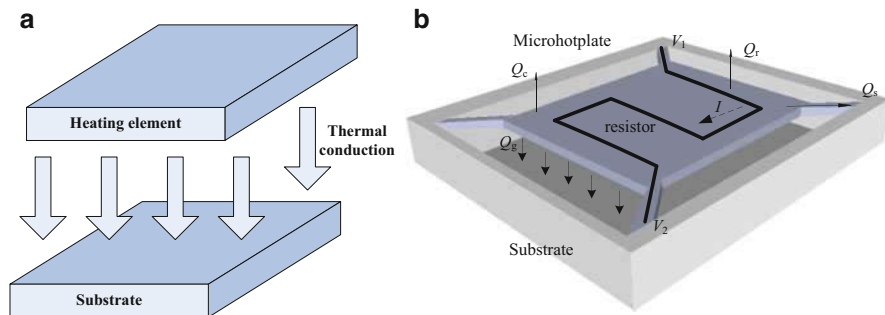
## Development of Vacuum Sensor

Vacuum can be measured either directly by measuring the force applied on the area by gas pressure or indirectly by measuring the physical quantity proportional to the gas pressure, such as thermal conductivity of gas, particle number density, particle impingement rate in gas, etc. Thus, there are mainly three kinds of vacuum sensor, including mechanical vacuum sensor, thermal conductivity vacuum sensor, and ionization vacuum sensor. Based on these three kinds of vacuum sensors, many different vacuum sensors are developed and applied on different vacuum ranges. The mechanical vacuum gauges include liquid manometer, McLeod sensor, rotary piston sensor, elastic element sensor, (Quartz) bourdon tube vacuum sensor, diaphragm vacuum sensor, piezoelectric vacuum sensor, and spinning rotor sensor, covering from  $10^5$  to  $10^{-4}$  Pa. Pirani sensor is one of the thermal conductivity vacuum sensor from  $10^5$  to  $10^{-1}$  Pa. For the high or ultrahigh vacuum measurement from 10 to  $10^{-10}$  Pa, the ionization vacuum sensors (IVS) are adopted, including Penning vacuum sensor, inverted magnetron, triode IVS, Bayard-Alpert IVS, ion spectroscopy IVS, and partial pressure sensor (Jousten 2008). In applications, many different types of vacuum sensors work in sequence to measure gas pressure. For example, if the high vacuum is measured, some types of diaphragm or thermal conductivity vacuum sensor are used to measure the pressure from  $10^5$  to  $10^{-1}$  Pa, and then the ionization vacuum sensor starts to measure the gas pressure below  $10^{-1}$  Pa. Or else, if the vacuum is not low enough to turn on the ionization vacuum sensor, the sensor will be damaged.

---

## Microfabricated Vacuum Sensors

Among different kinds of vacuum sensors, some of them can be fabricated by micromachining technology. In this part, the micro-Pirani and thermal ionization vacuum sensors are introduced.



**Fig. 1** The basic working principle of the Pirani sensor

### Pirani Vacuum Sensors

In 1906, Dr. Marcello Pirani, working in Siemens & Halske, Berlin, tried to melt the metal in vacuum environment. During the experiment, it was found that as the gas pressure decreased, the heating current for melting the metal decreased. This phenomenon was later named as Pirani effect, and the vacuum sensor based on this principle to detect gas pressure was called the Pirani vacuum sensor, which is also the thermal conductivity vacuum sensor (Ellett and Zabel 1931). The Pirani sensor typically contains a heating element which is exposed to the gas environment. The heating element is heated by the current through it and cooled by the surrounding gas heat dissipation, mainly the thermal conduction if neglecting the heat convection and radiation, as shown in Fig. 1a. If the gas pressure is reduced, the heat dissipation of the gas decreases so that the temperature of the heating element increased in the case of constant heating current and vice versa. Thus, gas pressure can be obtained by measuring of temperature of heating element (Wilfert 2004).

Microhotplate is one of the typical heating elements, and Fig. 1b is the simple structure of the microhotplate. The microhotplate is heated by the heating resistor, and the temperature is measured at same time. The heat of the microhotplate can be dissipated by the following heat loss mechanisms: solid thermal conduction through the supporting beams ( $Q_s$ ), gaseous conduction above and below the microhotplate ( $Q_g$ ), convection through gas ( $Q_c$ ), and thermal radiation ( $Q_r$ ). Due to small size and usually low operating temperature in microhotplate, the last two ones can always be neglected (Zhang et al. 2006).  $Q_s$  can be considered to be a constant and  $Q_g$  is dependent on the gas pressure (Topalli et al. 2009), which can be simply described as (Klaassen 1996)

$$Q_g = Q_0 \frac{PP_t}{P + P_t} = \begin{cases} Q_0 P_t & P > P_t \\ Q_0 P & P \leq P_t \end{cases} \quad (1)$$

$Q_0$  is the gaseous conduction in 1 Pa and  $P_t$  is the empirical transition parameter, which is dependent on the gap between microhotplate and substrate. According to

Eq. 1, only the gas pressure below  $P_t$  can be measured by the Pirani sensor. The lower-limit pressure of the dynamic range is determined by the solid thermal conduction through the supporting beam. Reducing the gap between the microhotplate and the substrate is the most effective method of increasing the high-pressure limit of the dynamic range. The sensitivity to the pressure is proportional to the area of the microhotplate (Chae et al. 2005).

The suspended gap filled with gas between the high-temperature element and low-temperature element is essential for Pirani vacuum sensor. Thus, there are mainly three structures of Pirani sensor, including microhotbridge (Khosraviani and Leung 2009), microhotplate (Zhang et al. 2006), and horizontal heat sink (Topalli et al. 2009).

To fabricate the suspended gap of microhotbridge and microhotplate, the most effective method is to adopt the wet etching to remove the thin layer below the structure. For the horizontal heat sink, the dry etching is the good choice to have deep etching with better vertical uniformity.

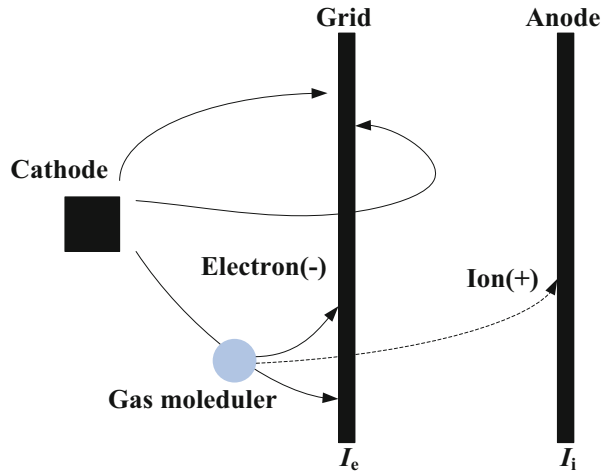
The heating and temperature sensing material of Pirani sensor is very important to its performance, which determines its sensitivity, power consumption, and service life. These materials include metal and metal compounds (Shie et al. 1995), polysilicon (Mitchell et al. 2008), doped silicon (Lee et al. 2007), and nanotubes (Kawano et al. 2007). They should have the good thermal stability. They also need to have higher-temperature coefficient because the heating resistor could be used as the temperature measurement resistor as well.

## Ionization Vacuum Sensors

The ionization vacuum sensor is used for the vacuum measurement below  $10^{-1}$  Pa. It consists of the cathode, grid, and anode. Electrons emitted from the cathode move several times in back and forth around the grid before finally entering the grid. During these movements, some electrons collide with a gaseous molecule to form a pair of ion and electron. The number of these ions is proportional to the gaseous molecule density multiplied by the electron current emitted from the cathode, and these ions are poured into the anode to form an ion current. Since the gaseous molecule density is proportional to the pressure, the gas pressure is estimated by measuring the ion current (Abbott et al. 2005).

Electrons are emitted from cathode and accelerated toward the grid (or electron collector) which has a large positive voltage applied to it, as seen in Fig. 2. Electrons initially miss the grid circle back until they are collected, forming the electron current  $I_e$ . If the gas molecules are present, some of the energetic electrons collide with some of the molecules, ionizing them. The ions are collected by the anode, which has the negative potential, forming the ion current  $I_i$ . The electrons generated by the ionization are also collected by the grid, adding to  $I_e$ . Theoretically, the  $I_i$  is proportional to the electron current  $I_e$ , the gas pressure  $p$ , the ionization rate  $\alpha$  of the ambient gas, and the electron path  $L$

**Fig. 2** The operation of the hot-filament vacuum gauge



$$I_i = I_e p \alpha L \quad (2)$$

Thus

$$(I_i/I_e)/p = \alpha L \quad (3)$$

The gauge sensitivity can be defined as

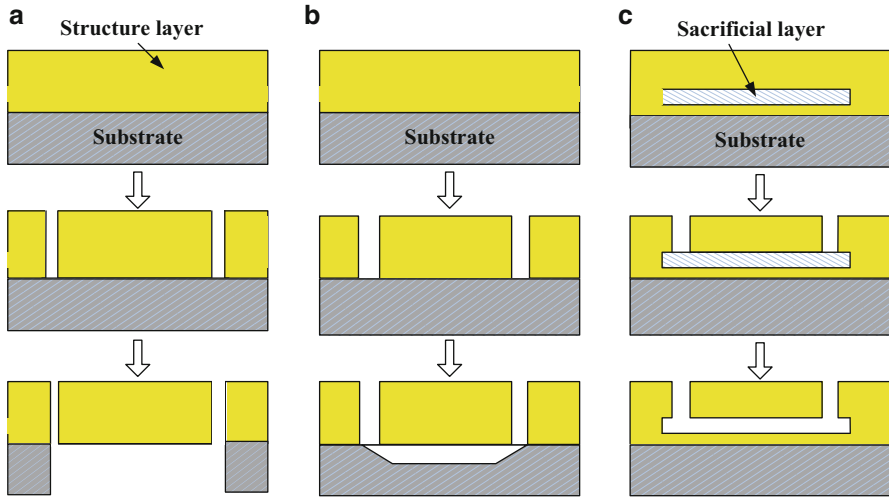
$$S = \alpha L \quad (4)$$

The sensitivity  $S$  given in Eq. 4 is a function of  $\alpha$  and the average electron-path length  $L$ , which depends on the biasing voltage, dimension, and the relative position of the cathode, grid, and anode (Williams 1997).

Based on the mechanism of electron emission from cathode, the ionization vacuum sensor can be divided into thermal emission and field emission ionization sensor. The key point for the ionization vacuum sensor is the materials of cathode. The cathode materials of the gauge include the tungsten (Williams and Muller 1997), the nanotube (Dong and Myneni 2004) (Bower et al. 2007), and silicon tip (Sun 1999), all of which can emit electron. Tungsten is the ideal material for the thermal ionization vacuum sensor due to its lower function work and high melting point. Nanotube and silicon/metal tip are used for the field emission ionization vacuum sensor because it is much easier with the sharp tip to emit the electron when applied to a voltage.

### Key Fabrication Process for Micro-Vacuum Sensors

For vacuum sensors including Pirani sensor and thermal ionization vacuum sensor, the key process is how to fabricate suspending structure, such as microhotbridge or



**Fig. 3** Fabrication process for suspending structures, including (a) back bulk-micromachining technology, (b) front bulk-micromachining technology, and (c) surface sacrificial layer technology

microhotplate for Pirani sensor and microbridge cathode for the ionization vacuum sensor. The most effective methods to form the suspending structure included the bulk micromachining and surface micromachining, as shown in Fig. 3. In the case of bulk micromachining, the microstructure is formed by etching the relative thick bulk substrate material, including the back and front bulk micromachining technology. For the surface micromachining, the thin-film layers deposited on top of a substrate are selectively removed as a sacrificial layer in a defined sequence to form the suspended structures.

There are two types of bulk-micromachining technology: back bulk micromachining and front bulk micromachining. For back bulk micromachining, the substrate is etched from backside by some etchant, including the wet etching or some kind of dry etching. The disadvantage is that it would take much time and cost to etch the thick layer of substrate materials. Thus, the front bulk micromachining is used to solve this problem. The etchant is dipped into the surface area of the bulk to remove the surface material to suspend the structure. The problem will focus on the difficulties to control the thickness of etching. In order to solve that, the surface micromachining technology with sacrificial layer is developed to control the thickness of etching and has better etching selectivity.

A sacrificial layer is deposited and patterned. After that, a structural thin film is deposited and patterned, which will perform the mechanical or electrical functions for the final device. A selective etchant then removes the sacrificial layer, leaving the structural layer suspended. The thickness of the sacrificial layer determines the suspended gap of the device. Common sacrificial layer materials includes polysilicon, silicon oxide, and aluminum.

For the field ionization vacuum sensors with cathode of nanotube, many fabrication processes are used. Traditionally, the cathode part should be first evaporated a layer as catalyst, such as iron or nickel, then carbon nanotube (CNT) was grown using chemical vapor deposition (CVD) (Dong and Myneni 2004). Recently, printing CNT paste directly to the cathode is a more convenient method for the fabrication of the field ionization vacuum sensor.

For silicon tip fabrication, oxidation is an effective method to sharpen the tip. During oxidation, the preliminary structure is sharpened due to the slower rate of surface reaction at the tip than at the sidewall. The faster surface reaction rate at the sidewall results in more silicon being consumed than at the tip, and consequently the tip is sharpened (Sun 1999).

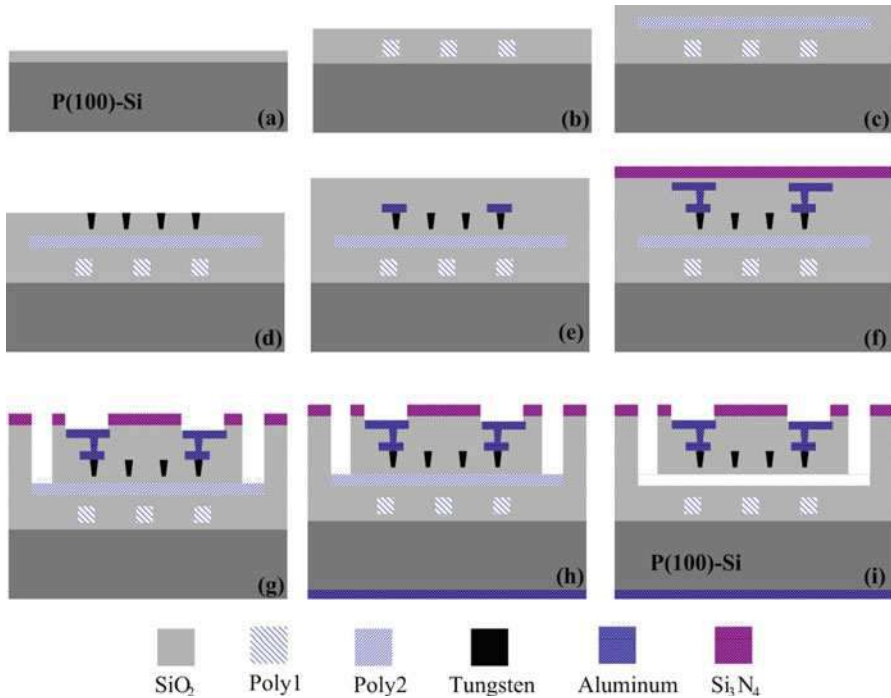
---

## CMOS Platform Technology for Vacuum Sensors

In measurement of vacuum, different types of vacuum sensors should work sequentially to measure the high vacuum. The sensor control systems are very complex, mainly including the driven circuit for the Pirani sensor and ionization sensor, signal sample circuit to obtain the signal representing the vacuum value, the A/D converter to convert the analog signal to digital for calculation, and the CPU to decide on the state of vacuum threshold to change to another type of vacuum sensor. Thus, if the sensors and all the control systems can be integrated on a chip, it will be very convenient for the application of vacuum sensor. Since this sensor system is very small in volume and weight, it can be useful in some special applications, such as micro-packaging or space application. In this part, we focus on integrating the Pirani and ionization vacuum sensor on a chip to meet the demands of vacuum test from the atmosphere to high vacuum.

### CMOS Pirani Vacuum Sensors

Although the MEMS-based micro-Pirani pressure sensor has been reported for over two decades, the promising approach is micro-Pirani sensor system combining the micromachined sensing structures and microelectronic building blocks on a single chip. Some micro-Pirani sensor systems are reported with the microbridge/micro-hotplate as the sensory component, the constant temperature bias circuit, and the on-chip A/D converter (Mastrangelo and Muller 1991) (Klaassen and Kovacs 1997). The digitized gas pressure can be obtained to have lower noise and make it easier for users than the traditional Pirani sensor. In this part, a micro-Pirani vacuum sensor system will be introduced. It consists of the tungsten microhotplate array with four same microhotplates in series, the constant current circuit, the 8-bit A/D converter, and the digital interface. With the external EEPROM as the look-up table circuit, this sensor system can not only sample the gas pressure between  $10^{-1}$  and  $10^5$  Pa but also real time display the gas pressure in LED and transfer it to the external process

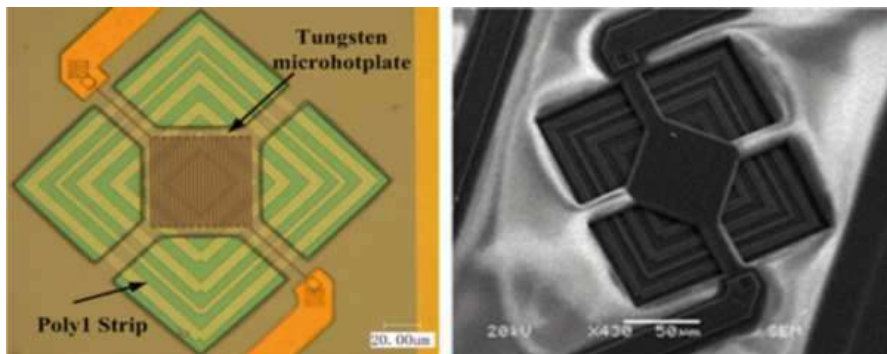


**Fig. 4** The fabrication process of the tungsten microhotplate

unit. All of them make it more suitable for the application in the gas pressure measurement and monitor systems (Wang et al. 2011).

The sensor system is implemented in 0.5  $\mu\text{m}$  CMOS process with two polysilicon layers (Poly1 and Poly2) and three metal layers (Metal1, Metal2, and Metal3). The metal plug between Metal1 and Metal2 is tungsten, and the one between Metal2 and Metal3 is aluminum due to its larger dimension. Tungsten is employed as the heater for microhotplate in the form of serpentine resistor instead of via plug in the standard CMOS process. Figure 4 describes the cross section of the fabrication process in details, from step (a) to step (i). The steps (a)–(g) are implemented in the CMOS foundry. (a) The  $\text{SiO}_2$  is deposited on P-type (100) wafer. (b) Poly1 is deposited and etched to the strip. (c) A 0.34- $\mu\text{m}$ -thick Poly2 is deposited as a sacrificial layer below the tungsten microhotplate area. After etching the  $\text{SiO}_2$ , the tungsten is grown by CVD process and polished by the chemical mechanical planarization (CMP) process, forming the tungsten resistor. (e) The Metal2 is sputtered to connect to the anchors of the tungsten resistor, leaving Metal1 unconnected. (f) The Metal3 is sputtered and  $\text{SiO}_2/\text{Si}_3\text{N}_4$  is used as the passivation layer. (g) The etch windows for the microhotplate and the bonding pads for the circuit are etched simultaneously during the bonding-pad etching process, and Poly2 is exposed due to overetch. (h) After dicing the wafer, the wet-etching process is prepared. Before this process, the new aluminum layer is sputtered at the backside of the die to prevent etching





**Fig. 5** The optical microscopy and SEM pictures of the tungsten microhotplate

from the backside. (i) The microhotplate is suspended with the removal of the polysilicon sacrificial layer by an improved tetramethyl ammonium hydroxide (TMAH) etching method which does not etch the exposed aluminum pads. About 8 h is required to remove the sacrificial layer in TMAH. The end of the wet-etching process can be identified according to the color of the Poly1 strip before and after etching.

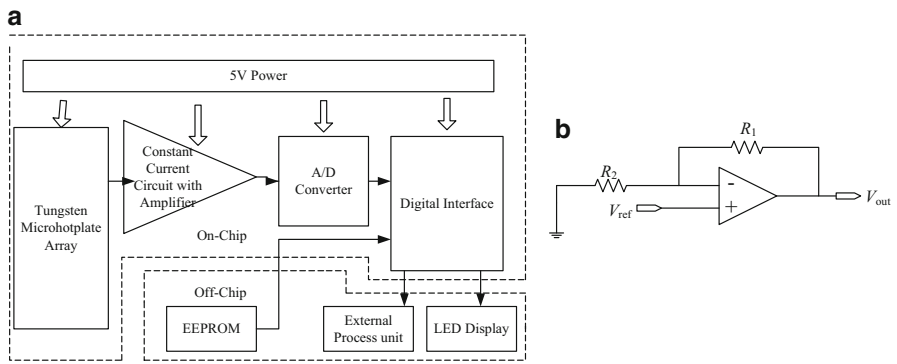
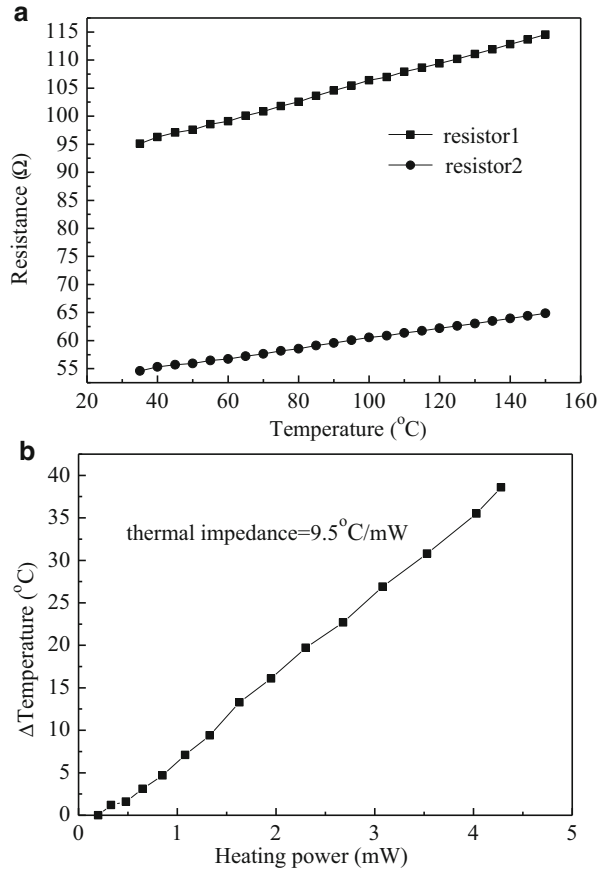
Figure 5 shows the optical microscopy and SEM pictures of the tungsten microhotplate, respectively. The tungsten microhotplate with a square area of  $60 \times 60 \mu\text{m}$  is suspended by four beams which have a length of  $30 \mu\text{m}$  and a width of  $15 \mu\text{m}$ . The thickness of the microhotplate is about  $4 \mu\text{m}$ . The width of the tungsten resistor is  $0.8 \mu\text{m}$ . The suspended gap between the tungsten microhotplate and the substrate is  $0.34 \mu\text{m}$  which is the thickness of Poly2.

Tungsten resistors have been exposed to the temperature-controlled oven with temperatures ranging from  $35$  to  $150 \text{ }^\circ\text{C}$ . The resistance and the corresponding temperature are recorded every  $5 \text{ }^\circ\text{C}$ . The results are shown in Fig. 6a. The temperature coefficient of tungsten is about  $1.5 \times 10^{-3}/^\circ\text{C}$ .

A current source (Keithley 2400) provides a constant current to the resistor of the tungsten microhotplate, and the voltage drop across the resistor is measured. Temperature is determined from the measured resistance with the temperature coefficient of the resistance. Heating power can be calculated from the current and voltage measured. Then a linear curve fit is applied to the power versus temperature data to extract thermal impedance that is the slope of the linear curve (Chae et al. 2005), as shown in Fig. 6b. The value of the thermal impedance of the tungsten microhotplate is  $9.5 \text{ }^\circ\text{C}/\text{mW}$  (Wang et al. 2009).

Based on the tungsten microhotplate, the Pirani vacuum sensor system is implemented. It consists of tungsten microhotplate array, the constant current circuit with amplifier, A/D converter, EEPROM, and digital interface. With these modules, the sensor system can not only sample the sensor output in analog signal, but also convert it to digital signal to display in LED and interface with the external central process unit (CPU), as shown in Fig. 7a.

**Fig. 6** Calibration of microhotplate. (a) Temperature coefficient of the tungsten resistors, (b) thermal impedance of the tungsten microhotplate



**Fig. 7** (a) Block diagram of the Pirani sensor system, (b) constant current bias circuit for Pirani sensor

The constant current circuit is adopted to bias the microhotplate. It is sensitive to the low gas pressure due to the high temperature of the microhotplate. Figure 7b is the constant current circuit.  $R_1$  is the tungsten resistors in microhotplate array, and  $R_2$  is also the tungsten resistors in substrate to compensate the influence of the temperature variation of substrate.

In Fig. 7b,  $V_{\text{ref}}$  is the external reference voltage, and the output voltage is

$$V_{\text{out}} = V_{\text{ref}} \left( 1 + \frac{R_1}{R_2} \right) \quad (5)$$

When the  $V_{\text{ref}}$  is constant, the voltage in the minus input of amplifier is also a constant and the current through  $R_2$  and  $R_1$  is constant.  $R_1$  increases when the gas pressure decreases, and  $V_{\text{out}}$  increases when the gas pressure decreases (Wang and Tang 2009).

The output voltage of the constant current circuit is connected to an 8-bit successive approximation register (SAR) A/D converter, which digitized the analog output voltage of the sensor to the digital signal.

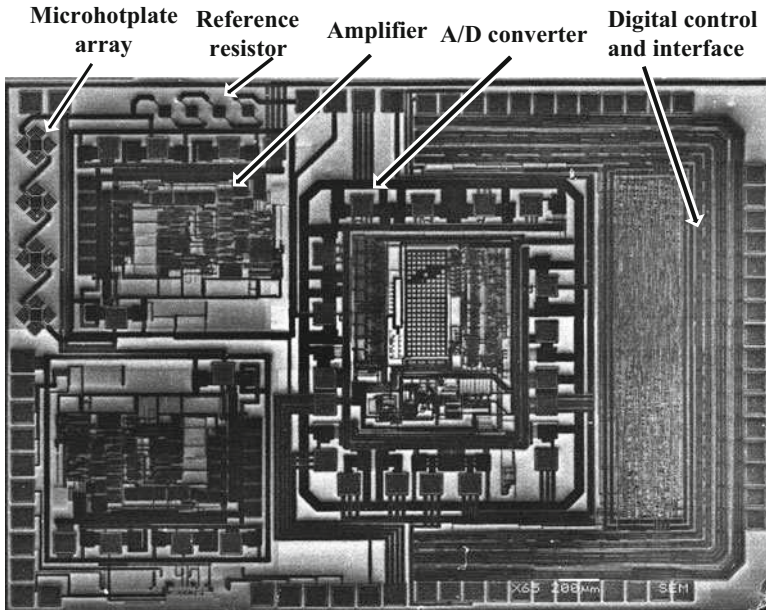
The digital signal from A/D converter is transferred to the EEPROM as the look-up table circuit. The output of the A/D converter is connected to the address of EEPROM, and the corresponding value in this address is the digital code for LED display. The manufacturing process of EEPROM cannot be compatible with our CMOS process, so the commercial EEPROM product is chosen as the external memory component.

The digital control and interface circuit is implemented with some functions, including reading the digital data from A/D converter, transferring it to the address of the EEPROM, reading the corresponding data from this address, encoding it to display in LED, and exchanging the data with the external processor. The sensor system is fabricated in 0.5  $\mu\text{m}$  CMOS process, and the sacrificial layer is etched by TMAH solutions. Figure 8 is the SEM picture of the sensor system including the tungsten microhotplate array, operational amplifier, A/D converter, and the digital interface. The dimension of the sensor system is 2.5  $\times$  3.6 mm.

The sensor system is packaged in PGA84, and Fig. 9a is the test board for the sensor system with the external components such as EEPROM and crystal. The LED is driven to display the gas pressure. With this vacuum system in Fig. 9b, the gas pressure between  $10^{-1}$  and  $10^5$  Pa can be sampled, digitized, and real time displayed by the LED, as shown in Fig. 9c, d. It can also be transferred to the external process unit, which makes it suitable for the application in the gas pressure control system.

## CMOS Ionization Vacuum Sensors

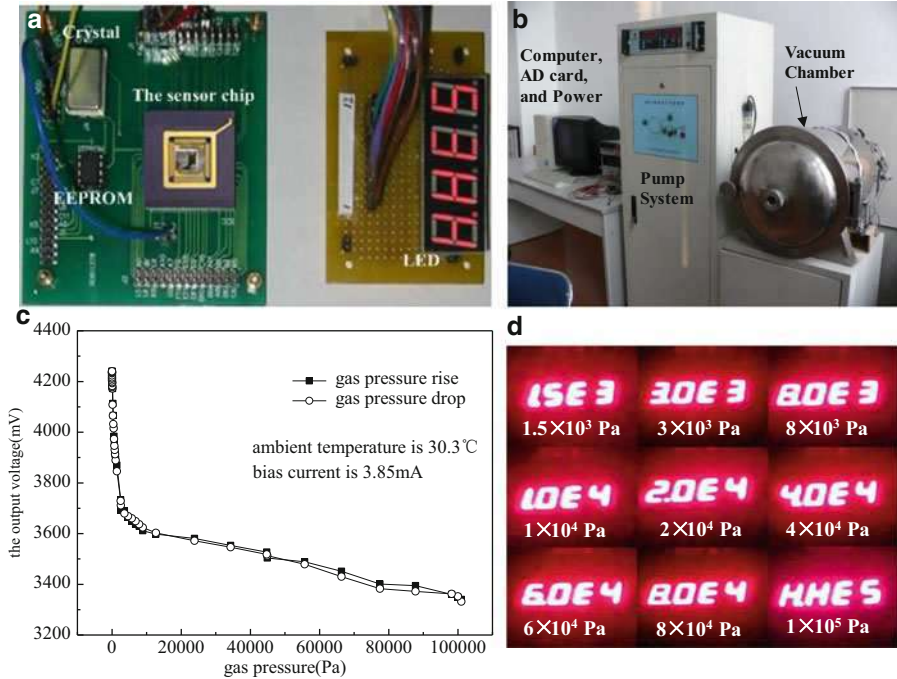
Also in this part, a CMOS-compatible hot-filament vacuum sensor is introduced. With it, the vacuum sensor and its control circuit can be fabricated monolithically on a chip which will decrease the noise and be easy to use. The feasibility is that the



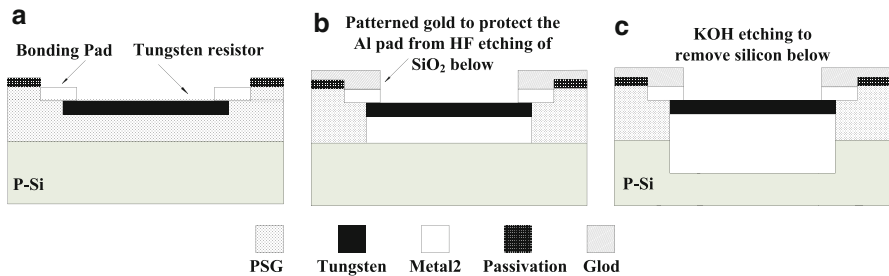
**Fig. 8** SEM picture of the Pirani sensor system

common material for the cathode of the hot-filament vacuum gauge is tungsten which is also used as the plug material in the standard CMOS process (Wang and Yu 2015). In the design, the hot-filament gauge is implemented in  $0.5\ \mu\text{m}$  CMOS process that features two metal layers (Metal1, Metal2). The main fabrication process consists of CMOS process and post-CMOS process, as shown in Fig. 10. In CMOS process, the tungsten resistor is fabricated with its anchors connected only to Metal2, leaving the Metal1 unconnected, as seen in Fig. 10a. In the bonding-pad opening process, the surface of the tungsten resistor is also exposed by overetching the PSG and  $\text{Si}_3\text{N}_4$  passivation layer. In post-CMOS process, the Al pad is protected by gold, and HF is used to remove the PSG below the tungsten resistor, as seen in Fig. 10b. In Fig. 10c, the silicon substrate below the tungsten resistor is etched by KOH to further suspend the tungsten resistor as the cathode of the hot-filament sensor. The picture of the cathode structure is shown in Fig. 11a. Also, the spring structure of cathode can improve stability of the cathode during heating, as shown in Fig. 11b.

When vacuum chamber is less than  $0.5\ \text{Pa}$ , the electron emitter was applied a current, and the voltage for grid and anode is assigned  $100$  and  $-30\ \text{V}$ , respectively. The test circuit and boundary condition are shown in Fig. 12. Reference resistor was used to adjust the bias current for cathode. During testing, the sensor would be burned if the current is large, while the sensitivity was low if the current is small.  $2$ ,  $4$ , and  $6\ \text{mA}$  current was chosen. For  $2\ \text{mA}$  current, when the gas pressure is higher than  $0.5\ \text{Pa}$ , the sensor is burned; for  $4$  and  $6\ \text{mA}$  of heating current, when the

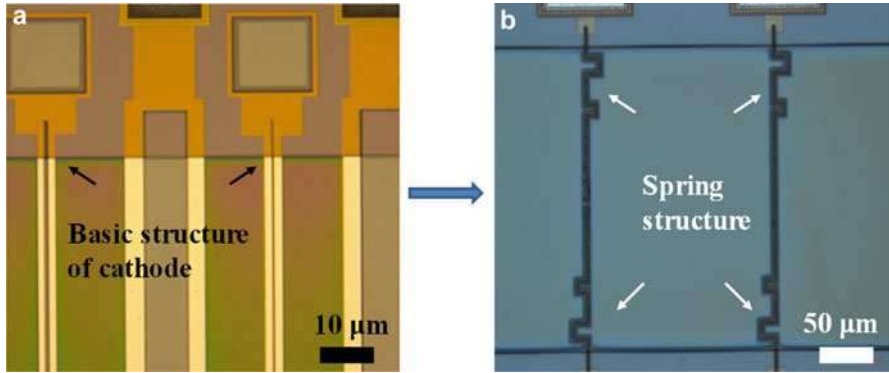


**Fig. 9** (a) Packaging and test circuit board of Pirani sensor system, (b) calibration system for vacuum sensor, (c) calibration results of Pirani sensor, (d) LED display of vacuum value

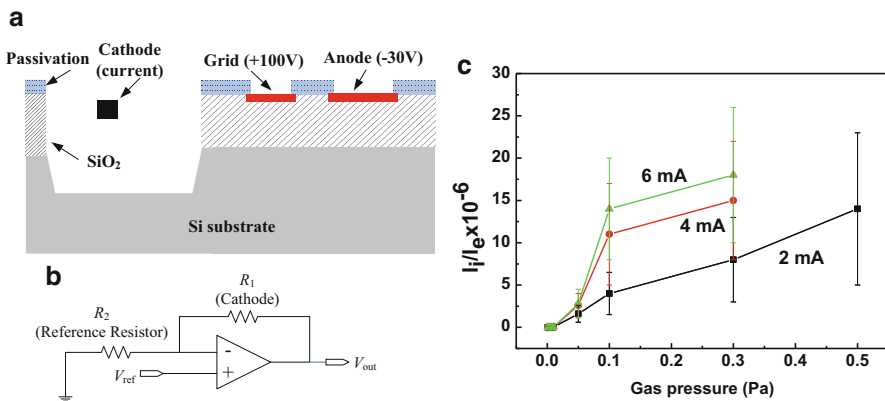


**Fig. 10** Fabrication process for the cathode of hot-filament vacuum sensor

pressure is higher than 0.3 Pa, the sensor is burned. From the test results, the response range of the sensor is 0.05–0.5 Pa, which was lower than the thermal ionization sensor with large volume. The main reason is that the electron trajectory is limited, and the probability of the ionized gas molecule is small for the microthermal ionization sensor (Williams and Muller 1997). However, the sensor has the advantages of small size, volume, and weight, which would have its potential applications, such as the space exploration, micro-packaging, etc.



**Fig. 11** Microscope picture of cathode of ionization gauge



**Fig. 12** Structure, bias circuit, and results of thermal ionization vacuum sensor

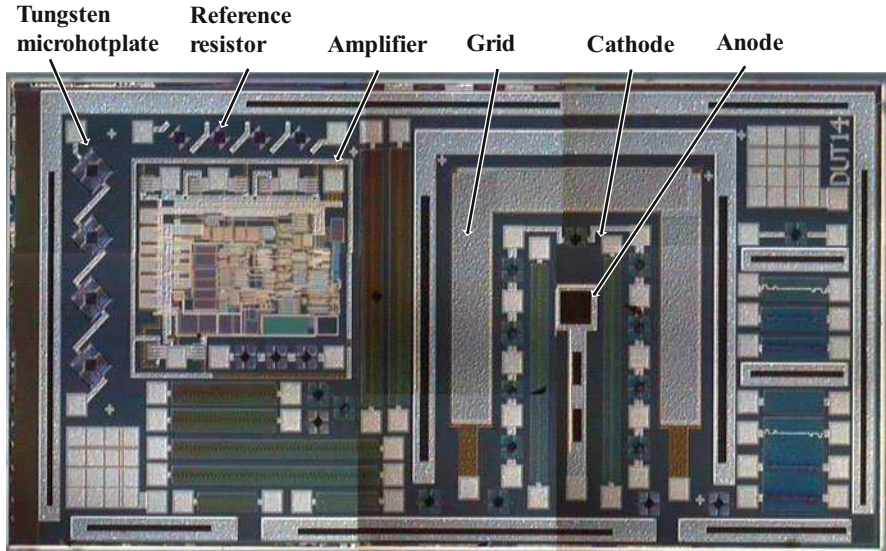
### CMOS Vacuum Multisensor Systems

Since the CMOS-compatible Pirani sensor and ionization sensor are implemented individually, the future work should focus on how to combine two of them on a chip. Figure 13 consists of the Pirani sensor, constant current circuit, and thermal ionization vacuum sensor.

### Key Fabrication Process for CMOS Vacuum Sensors

For the vacuum sensor fabricated by the CMOS process, the compatible materials should be considered first. In Pirani sensors, the material of the heating resistor is important. Usually, the platinum is the best choice, but it is not the common material used in CMOS process. The CMOS-compatible materials included polysilicon,





**Fig. 13** Fully integrated vacuum sensor system, including Pirani sensor and thermal ionization vacuum sensor on a chip

aluminum, and tungsten. With the developing manufacturing technology of the integrated circuits, the polysilicon will have a smaller temperature coefficient to meet the demands of the integrated circuit technology. The worst situation is that the polysilicon may be unsuitable for the design of the microhotplate if more advanced manufacturing technologies are adopted, thus greatly limiting the development of the CMOS-compatible micro-Pirani pressure sensors. Paul presented a micro-Pirani pressure sensor with the microhotplate whose heating resistor was the meander-shaped aluminum thin film (Paul and Baltes 1995). However, the aluminum is prone to suffer from a reliability problem known as electromigration. Furthermore, the melting point of aluminum is low. Tungsten is not only resistant to the electromigration failure but also has a high melting point. The temperature coefficient of tungsten is comparable to the platinum. For the thermal ionization vacuum sensor, the best material is tungsten due to the high melting point and low work function.

The key process of the CMOS-compatible vacuum sensor is the etching process. During etching, the aluminum electrode pads of CMOS circuit should be protected. When the alkaline solution is used to etch the bulk or sacrificial layer of silicon or polysilicon, the buffered TMAH is an alternative solution, instead of KOH. Etch rate of aluminum is much lower than that of silicon or polysilicon. When the acidic solution is used to remove the  $\text{SiO}_2$ , the aluminum electrode should be covered with the gold or platinum to prevent etching by the HF. In this process, the gold or platinum should be sputtered and patterned after the CMOS process in foundry.

Thermal influence is another consideration of the vacuum sensor system. Since the Pirani and thermal ionization vacuum sensors are the thermally driven, the heat

of the sensor can also be conducted from the substrate to the electrical circuit part although the sensors are suspended. This kind of heat can potentially influence the function of the circuit. Thus, in the design of the sensor system, it is better to place the sensor at the corner of the die, instead of at the center of the die. The enough distance between the sensor and the circuit should be as large as possible.

---

## Outlook and Future Development

With the development of CMOS-compatible technology and the decreasing cost, the integrated vacuum sensor system will have more functions. Until now, the reported integrated vacuum sensor systems included vacuum sensors, signal sampling, processing, and analog-digital conversion (ADC) circuit. The next generation of sensor systems will integrate more functional modules, including vacuum sensor with different working principles, the self-power systems, CPU, and wireless transmission module (Bluetooth, wifi), which have a great convenience to users and the development of vacuum test system.

Although there are a large number of academic researches for micro-vacuum sensor systems, there is no product of integrated vacuum sensor system in the market, mainly because the traditional vacuum sensor products are very mature and widely used in industrial equipment, leading to less demand of products of miniaturized vacuum sensor systems. Therefore, the most likely application of integrated vacuum sensor will include small space vacuum measurement, such as microelectronic device packages of infrared thermal imager and resonator which require high vacuum sealing. Another potential application is space industry, which has a strict demand of the size and weight of the system. In space vehicles, in order to control the temperature rising during the transportation of liquid hydrogen, liquid oxygen, liquid nitrogen, and other low-temperature mediums, vacuum multilayer insulation pipes with high vacuum between different layers are used to reduce heat dissipation; micro-vacuum sensor is needed due to its small size. For the Mar rover, the main test components are placed in a sealed vacuum chamber to ensure the accuracy and effectiveness of detection; the vacuum sensor system has to be installed in the vacuum chamber. Due to the strict limitation of the size, weight, and power of the vacuum sensor system, the integrated vacuum sensor system will be the only option.

---

## References

- Abbott P, Jabour Z (2011) Vacuum technology considerations for mass metrology. *J Res Nat Inst Stand Technol* 116(4):689
- Abbott P, Looney J, Mohan P (2005) The effect of ambient temperature on the sensitivity of hot-cathode ionization gauges. *Vacuum* 77(2):217–222
- Bower CA, Gilchrist KH, Piascik JR, Stoner BR, Natarajan S, Parker CB, Wolter SD, Glass JT (2007) On-chip electron-impact ion source using carbon nanotube field emitters. *Appl Phys Lett* 90(12):124102



- Chae J, Stark BH, Najafi K (2005) A micromachined Pirani gauge with dual heat sinks. *IEEE Trans Adv Packag* 28(4):619–625
- Dong C, Myneni GR (2004) Carbon nanotube electron source based ionization vacuum gauge. *Appl Phys Lett* 84(26):5443–5445
- Dong K-Y, Lee YD, Kang BH, Choi J, Ju B-K (2013) Design of a multi-walled carbon nanotube field emitter with micro vacuum gauge. *Nanoscale Res Lett* 8(1):1
- Ellett A, Zabel RM (1931) The pirani gauge for the measurement of small changes of pressure. *Phys Rev* 37:1102–1111
- Johnson L, Young R, Montgomery E, Alhorn D (2011) Status of solar sail technology within NASA. *Adv Space Res* 48(11):1687–1694
- Jousten K (2008) *Handbook of vacuum technology*. Wiley, Weinheim
- Kawano T, Suter M, Cho C, Chiamori H, Lin L (2007) Single carbon nanotube pirani gauge by local synthesis. In: *TRANSDUCERS 2007-2007 international solid-state sensors, actuators and microsystems conference*. IEEE, pp 1015–1018
- Khosraviani K, Leung AM (2009) The nanogap Pirani—a pressure sensor with superior linearity in an atmospheric pressure range. *J Micromech Microeng* 19(4):045007
- Klaassen EH (1996) *Micromachined instrumentation systems*. PhD dissertation, Stanford University
- Klaassen EH, Kovacs GT (1997) Integrated thermal-conductivity vacuum sensor. *Sens Actuators A* 58(1):37–42
- Lee J, Wright TL, Abel MR, Sunden EO, Marchenkov A, Graham S, King WP (2007) Thermal conduction from microcantilever heaters in partial vacuum. *J Appl Phys* 101(1):014906
- Mastrangelo CH, Muller RS (1991) Microfabricated thermal absolute-pressure sensor with on-chip digital front-end processor. *IEEE J Solid State Circuits* 26(12):1998–2007
- Mitchell J, Lahiji GR, Najafi K (2008) An improved performance poly-Si Pirani vacuum gauge using heat-distributing structural supports. *J Microelectromech Syst* 17(1):93–102
- Paul O, Baltes H (1995) Novel fully CMOS-compatible vacuum sensor. *Sens Actuators A* 1(46):143–146
- Shie JS, Chou BC, Chen YM (1995) High performance Pirani vacuum gauge. *J Vac Sci Technol A* 13(6):2972–2979
- Sun C (1999). *Microengineered sensor devices with field emission electron sources*. PhD dissertation, New Jersey Institute of Technology
- Topalli ES, Topalli K, Alper SE, Serin T, Akin T (2009) Pirani vacuum gauges using silicon-on-glass and dissolved-wafer processes for the characterization of MEMS vacuum packaging. *IEEE Sensors J* 9(3):263–270
- Wang J, Tang Z (2009) A fully CMOS-compatible micro-Pirani gauge based on a constant current. In: *IEEE 22nd international conference on, Micro electro mechanical systems, MEMS*. IEEE, pp 829–832
- Wang JQ, Yu J (2015) Fabrication process and electro-thermal modeling for the cathode of the CMOS-compatible hot-filament vacuum gauge. *Key Eng Mater* 645:836–840
- Wang JQ, Tang ZA, Li JF, Zhang FT (2009) A MicroPirani pressure sensor based on the tungsten microhotplate in a standard CMOS process. *Ieee T Ind Electron* 56(4):1086–1091
- Wang J, Tang Z, Li J (2011) Tungsten-microhotplate-array-based pirani vacuum sensor system with on-chip digital front-end processor. *J Microelectromech Syst* 20(4):834–841
- Wilfert S (2004) Miniaturized vacuum gauges. *J Vac Sci Technol A* 22(2):309–320
- Williams KR (1997) *Micromachined hot-filament vacuum devices*. PhD dissertation, UC Berkeley
- Williams KR, Muller RS (1997) Micromachined hot-filament ionization pressure sensor and magnetometer. In: *1997 International conference on solid-state sensors and actuators (Transducers 97)*, vol 1000. Chicago, pp 1249–1252
- Zhang F, Tang Z, Yu J, Jin R (2006) A micro-Pirani vacuum gauge based on micro-hotplate technology. *Sens Actuators A* 126(2):300–305



# Micro Thermal Flow Sensor

Rong Zhu

## Contents

Overview .....	596
Design and Fabrication of Sensors .....	597
Principle of Thermal Flow Sensing .....	597
Fabrication of Hot-Film Sensors .....	600
Geometric Design of Hot-Film Elements .....	602
Measurement of Thermal Flow Sensor .....	606
Working Operation of Flow Sensor .....	606
Temperature Compensation .....	607
Applications .....	610
Surface Airflow Detection for Micro Air Vehicles .....	610
Micro anemometer for 2D Wind Detection .....	615
Respiratory Airflow Monitor .....	617
Concluding Remarks .....	621
References .....	622

## Abstract

Measurement of fluid mechanics is very important in various fields, and flow sensors have been widely applied to execute accurate and efficient measurements. Compared with other sensing principle, thermal flow sensors are based on convective heat transfer and take merits of simple structure and easy use and thus offer a practical solution for various fluidics applications. In this chapter, we describe mainly hot-film anemometer fabricated on polyimide substrate. Hot-film or hot-wire anemometer utilizes a thermal element that serves as both a joule

---

R. Zhu (✉)

State Key Laboratory of Precision Measurement Technology and Instrument, Department of Precision Instrument, Tsinghua University, Beijing, China  
e-mail: [zr\\_gloria@mail.tsinghua.edu.cn](mailto:zr_gloria@mail.tsinghua.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,  
[https://doi.org/10.1007/978-981-10-5945-2\\_19](https://doi.org/10.1007/978-981-10-5945-2_19)

595

heater and a temperature sensor. We introduce the principle of thermal flow sensing, design and fabrication of the micro hot-film flow sensor, the measurement methodology, and application cases by using the micro hot-film flow sensors.

---

**Keywords**

Flow sensor · Hot film · Micro anemometer · Measurement principle · Temperature compensation

---

**Overview**

Measurement of fluid mechanics is very important in various fields, and flow sensors have been widely applied to execute accurate and efficient measurements. Of these sensors, micromachined flow sensors have been developed for several decades, which can be classified as either thermal or nonthermal. Silicon-based flow sensor was firstly demonstrated in 1974 (Van Putten and Middelhoek 1974), and then many application-oriented flow sensors have been developed based on various sensing principles, such as thermal anemometry, Doppler frequency shift, and indirect inference from pressure differences (Chen et al. 2003). Among these sensors, thermal flow sensors take merits of simple structure and easy use and thus offer a practical solution for various fluidics applications (Van Baar et al. 2001).

The principle of thermal flow sensors is based on convective heat transfer. According to the first law of thermodynamics, King's law of convection, and Newton's law of cooling, a heat balance can be maintained between heat generation and heat loss (Bruun 1995). Changes in the fluidic flow that affect the heat transfer around the heated element are transduced into electrical signals by thermometers.

There are three types of thermal flow sensing: hot wire/hot film, calorimetric, and time of flight (Nguyen 1997). Hot-wire/hot-film anemometer utilizes a thermal element that serves as both a joule heater and a temperature sensor. Under a constant power and zero flow rate, the temperature of the electrically heated thermal element reaches a steady-state value. If an external fluid flow passes through the thermal element, the thermal element experiences forced convective cooling. Accordingly, the temperature of the thermal element varies and provides a means to gauge the cooling rate and the flow velocity (Chen and Liu 2003). Calorimetric sensing involves at least one temperature sensor upstream and downstream, a heating element that detects the thermal profile around the heater due to fluid flow. The forward and backward directions of fluid flow can be detected with flow velocity measurements. Time-of-flight sensing involves at least one heater and one downstream temperature sensor. The transit time of a thermal pulse generated by the heater and transferred to the downstream sensor is tracked to extract flow rate information.

In this chapter, we focus on the hot-film anemometer fabricated on flexible substrate due to its simple structure, easy implementation, and wide usages.

## Design and Fabrication of Sensors

### Principle of Thermal Flow Sensing

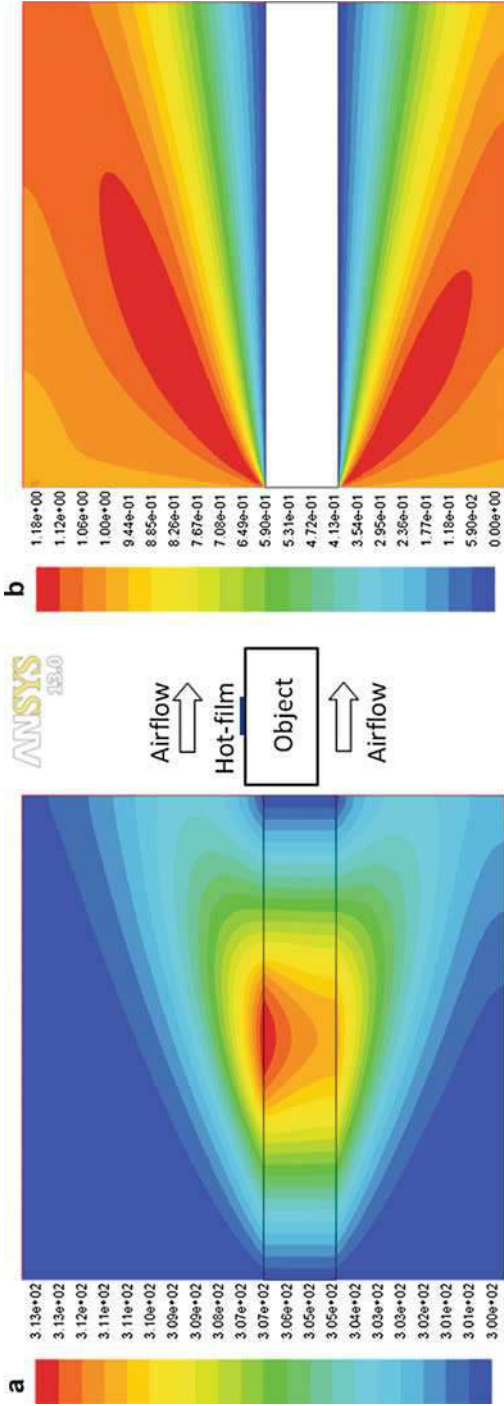
The working principle of the hot-film flow sensors relies on the detection of the convective heat transfer between an electrically heated resistive sensing element and fluidic flow. The hot film experiences cooling due to its heat transfer given that the hot-film element is heated to a higher temperature than that of the surrounding. Heat transfer depends on the flow velocity. Thus, a flow sensing is accomplished without using any moving parts and requires only the hot-film element that serves as both joule heater and temperature sensor. The hot-film flow sensor is small, with high spatial resolution, simple to be implemented, and easy to be arrayed locally. Especially, hot-film sensor can be fabricated on flexible substrate to form a flexible sensor, which can be adhered on the surface of an object to detect surface flow. The flow field information around the surface of an object can be divided into two components, that is, normal pressure exerted on the surface and shear stress along the surface. Both of the components are useful for inferring the flow parameters, such as flow velocity and wall shear stress in fluids. Hot-film sensors are located in the boundary layer of a flow environment and experience heat loss when subjected to fluid flow. Flow sensing is accomplished by monitoring the resistance change of the hot-film element with temperature.

When sensing a surface flow, the boundary layer close to the object needs to be emphatically considered. The sensing principle of the thermal sensor is based on the heat transfer that is related with the thermal boundary layer. In heat transfer, Prandtl number  $Pr$  controls relative thickness of momentum and thermal boundary layers.

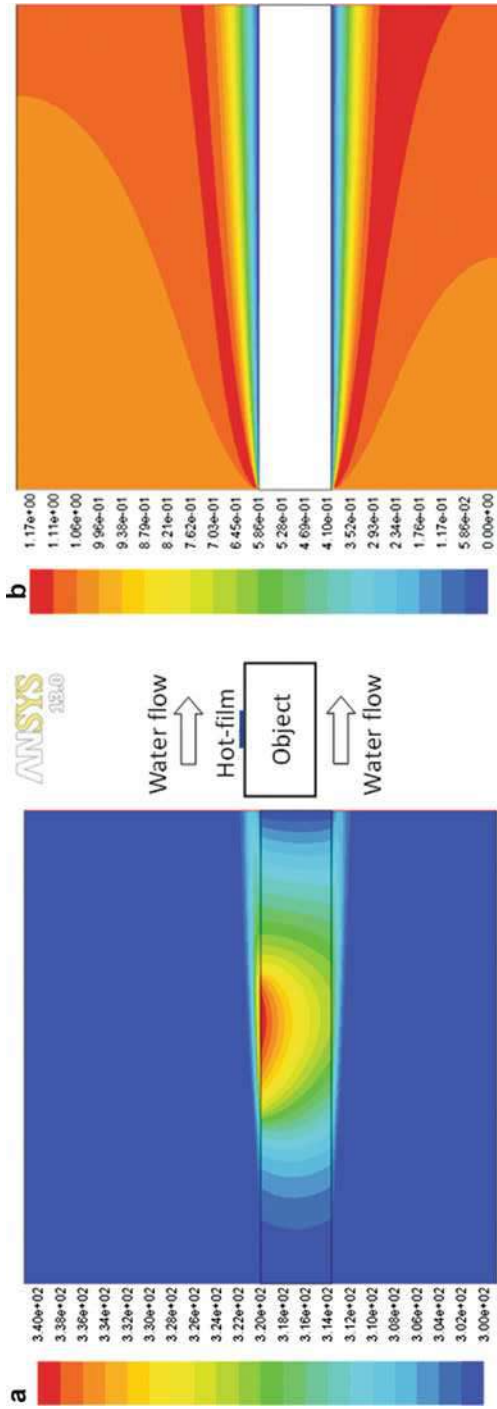
$$Pr = \frac{\nu}{\alpha} = \frac{\text{kinematic viscosity}}{\text{thermal diffusivity}} \approx \left( \frac{\delta}{\delta_t} \right)^3 \quad (1)$$

where  $\delta$  and  $\delta_t$  are thicknesses of kinematic boundary and thermal boundary layers, respectively, and  $\nu$  and  $\alpha$  are kinematic viscosity and thermal diffusivity, respectively.

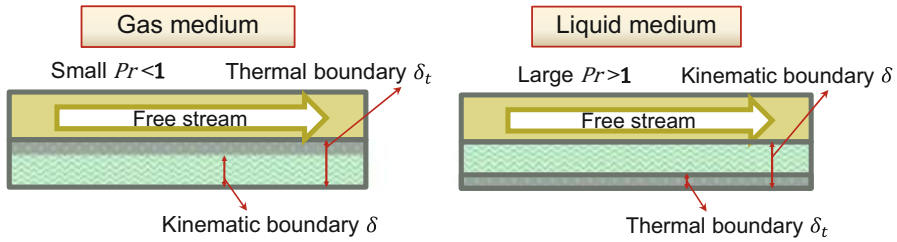
When  $Pr$  is smaller than 1, it means that heat diffuses very quickly compared to velocity (momentum). This means for air and many other gases ( $Pr$  is around 0.7–0.8), thickness of thermal boundary layer is higher than that of velocity boundary layer. An example of the thermal boundary layer and corresponding momentum boundary layer under an airflow blowing through a heated hot film on a substrate simulated by using ANSYS Fluent is shown in Fig. 1. Vice versa, when  $Pr$  is larger than 1, for water (around 7) and oil (around 100–40,000), thickness of thermal boundary layer is much smaller than velocity boundary layer. Simulated results of thermal boundary layer and fluid boundary layer under water flow are shown in Fig. 2. Figure 3 shows the relationship of thermal and kinematic boundary layers in different fluid medium with  $Pr < 1$  and  $Pr > 1$ , respectively. In case of smaller  $Pr$ , the thermal layer of a hot-film sensor is dependent on both of the velocity boundary



**Fig. 1** The simulated thermal boundary layer and corresponding momentum boundary layer on condition that airflow of 1 m/s blows through a hot film (dimension of viewing area is  $1 \times 1 \text{ mm}^2$ ) (Zhu et al. 2013). (a) Contours of static temperature (K). (b) Contours of velocity magnitude (m/s)



**Fig. 2** The simulated thermal boundary layer and corresponding momentum boundary layer on condition that a water flow of 1 m/s goes through the hot film on a substrate (dimension of viewing area is  $1 \times 1 \text{ mm}^2$ ) (Zhu et al. 2013). (a) Contours of static temperature (K). (b) Contours of velocity magnitude (m/s)



**Fig. 3** Relationship between the thermal and kinematic boundary layers in different fluid medium

layer and free-stream velocity, which means that the sensor temperature is dominated by both shear stresses and free-stream flow velocity. But in case of bigger  $Pr$ , the thermal layer of the sensor is relatively thin and completely falls into velocity boundary layer, which means the sensor temperature is dominated by wall shear stress of the flow. Accordingly, thermal flow sensor is appropriate for detecting both shear stress and flow velocity in airflow but wall shear stress in water flow. Generally, boundary layer is dependent on fluidic properties, i.e.,  $\delta \propto \sqrt{\gamma}$ , where  $\gamma$  is kinematic coefficient of viscosity. Due to the ratio of  $\gamma_{\text{air}}/\gamma_{\text{water}} \approx 10^4$ , thickness of boundary layer in water is much smaller than that in air (smaller in about two orders) as shown in Figs. 1 and 2, which turns out a rigorous demand to the sensor thickness for detecting wall shear stress in water.

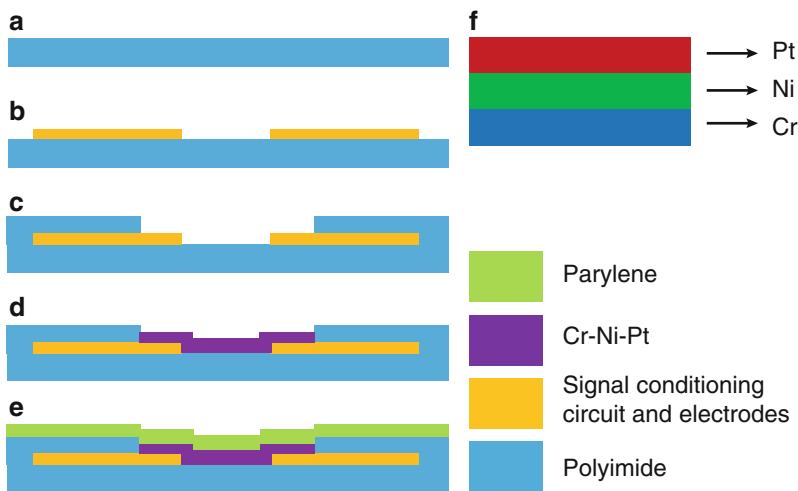
## Fabrication of Hot-Film Sensors

Currently, most micromachined sensors are developed using single-crystal silicon substrates. An important reason for making sensors out of silicon lies in the fact that piezo-resistive element can be realized in silicon by selective doping. However, silicon devices are relatively brittle compared with polymer and metal-based devices. A silicon beam may fracture easily in the presence of shock. Besides, lead bonding between sensing elements and signal conditioning circuit is complicated. To breakthrough these limitations, a sort of micro flow sensor is developed using a flexible polyimide substrate fabricated by integrating a standard flexible printing technique with thin-film sputtering/deposition (Liu et al. 2009b).

Polyimide (PI) is chosen as substrate because thermal conductivity of polyimide is low, almost two orders of magnitude lower than that of silicon. Also, polyimide has good mechanical properties and is chemically inert. In addition, polyimide is a standard material of flexible circuit boards, and thus sensor electrodes and connecting wires can be manufactured by using a standard flexible printing technology. For the above considerations, the selection of polyimide substrate provides both technical and economic advantages. Thickness of polyimide substrate can be selected from standard series (e.g., 50 or 150  $\mu\text{m}$ ) in accordance to practical requirement.

Material selection for hot-film sensing elements needs to consider about its capability of thermal sensing. A temperature coefficient of resistance (TCR) is an index to assess thermos-sensibility of material. TCR describes relative change of resistance that is associated with a given change in temperature. Metal film resistors (e.g., tungsten, platinum, or metal alloys) are commonly adopted as sensing elements for hot-film sensors. Tungsten has a modest TCR, but with a high-melting point, and thus is hard to be deposited. Platinum (Pt) does not have as high TCR as nickel (Ni). However, the oxidation resistance behavior of Ni is not satisfactory, while Pt has good antioxidation. Considering about both sensitivity and stability, sensing elements of hot-film flow sensors can be made of homogenous material of Pt or composite material of Ni and Pt. In composite material, Ni is used as main material and Pt as a covering layer, in addition to chromium (Cr) or titanium (Ti) as an adhesion between Ni and substrate PI for enhancing adhesiveness of sensing elements.

The hot-film sensing elements can be fabricated by incorporating printing and film depositing processes. One example of fabrication process is shown in Fig. 4. The process is started with standard printing technique on a PI substrate (Fig. 4a–c). The sensor electrodes and connecting wires made of Cu/Au are printed on the PI substrate. The prepared electrodes and wires help to simplify the process of lead bonding for the sensors. Then the metal films are deposited layer by layer (Fig. 4d), orderly from bottom to top. The metal films are patterned to be a predefined shape and lie across electrodes by using a prepared solid mask or photoetching. Finally, a parylene film is optionally deposited on the sensor and served as an encapsulation for waterproofs and dustproofs (Fig. 4e).



**Fig. 4** Fabrication process of hot-film sensor. (a–c) Standard printing technology on flexible polyimide substrate. (d) Deposition of Cr-Ni-Pt films. (e) Optionally coating parylene. (f) The structure of the composite metal layers (Liu et al. 2009b)



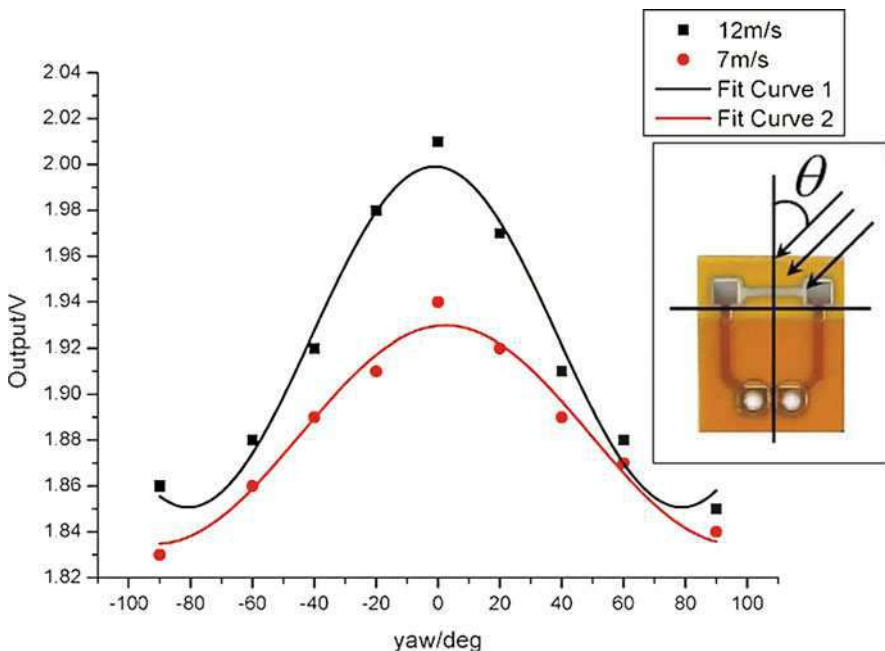
After the fabrication process of the sensing elements, a posttreatment is performed in order to make the sensors more stable and durable. In general, the hot-film sensors are annealed at a high temperature above recrystallization temperature of sensing element material while considering heat resistance of PI substrate. The sensors are further electrified for aging at a constant power for several hours. By the aid of the posttreatment, structural defects and interior stress are eliminated, and therefore stabilities and thermos-sensitivities of sensors are enhanced.

## Geometric Design of Hot-Film Elements

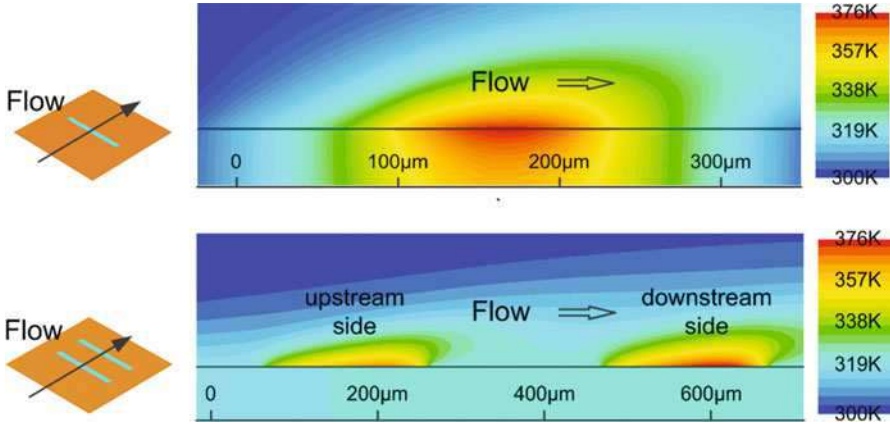
Geometric design of sensing elements is very important to fulfill various application demands. A simple geometric shape of one-component sensing element is strip profile. Under a sufficiently large aspect ratio, the flow angle response of a strip-shape sensing element follows a cosine relation (Bruun 1995)

$$U(\theta) = U_0 + U_v \cos(\theta) \quad (2)$$

where  $U$  and  $\theta$  are sensor output and flow angle, respectively, and  $U_0$  and  $U_v$  are zero output and amplitude, respectively, which are all correlated with flow



**Fig. 5** Angular responses under different flow velocities (Liu et al. 2009b)

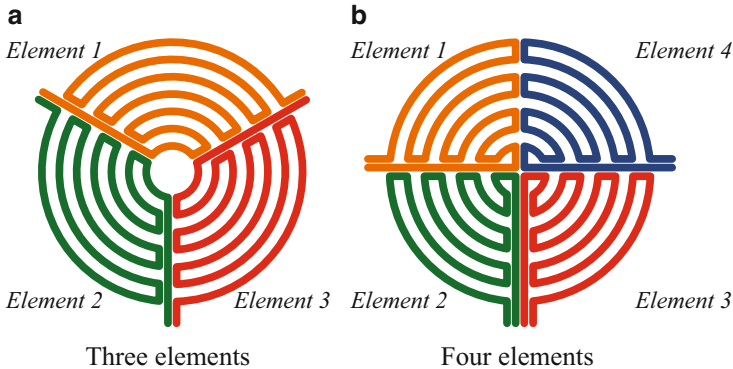


**Fig. 6** Simulated temperature field above one or two thermal elements

velocity. Figure 5 shows actual measurements of a strip thermal flow sensor corresponding to different flow velocity and flow angle. The angular response of the sensor slightly deviates from the cosine law due to the low aspect ratio of sensing elements. Under a low aspect ratio, hot-film flow sensor is suffered by tangential cooling effect. However measurement results still exhibit a directional dependence, which can be applied to gauge two-dimensional (2D) flow vector by using a combination of thermal element arrays (Nguyen 2005; Liu et al. 2009a).

Hot-film element of flow sensor is electrically heated to a higher temperature than that of the surrounding and thus forms a temperature distribution above the thermal element. When an external flow passes through the sensor, the temperature field will be deflected in the direction of the flow that results in the temperature difference among upstream or downstream elements as shown in Fig. 6.

Combination of multiple thermal elements can measure 2D flow vector over a range of 360°. To implement this, both heating and sensing elements need to follow specific criterions. Firstly, the entire heat generation area needs to be central symmetry so as to avoid coupling of total heat generation with flow direction. Secondly, the sensing area needs to be divided into electrically insulated sectors for detecting flow-induced heat power differences among the sectors. According to both considerations above and as a matter of fact, heating and sensing elements are integrated into one element in a hot-film sensor, and a round shape is adopted and divided equally into insulated sectors serving as thermal sensing sections. In each section, a roundabout wire is used to serve as a thermal sensing element. The number of divided sectors implies the number of sensing elements, which also determines the number of conditioning circuits needed for the sensor. For simplifying operation system and reducing power consumption, the number of thermal elements needs to be minimized. The number of elements also refers to the number of sensor outputs. Experimental results indicate that output of each element in flow vector sensor



**Fig. 7** Sensor layout design. (a) Three elements. (b) Four elements

follows approximately a law of sine or cosine with respect to flow direction angle. Flow velocity is approximately proportional to zero output and amplitude of sine/cosine wave, and flow direction angle is approximately proportional to signal phase. Theoretically, deducing two unknown variables (velocity and angle) requires two algebraic equations to be solved. It seems only two thermal elements are necessary to deduce flow velocity and angle. However two elements are symmetrical about center line of round that can only deduce angle in a range of  $180^\circ$ . To achieve full-scale sensing range of  $360^\circ$ , at least three thermal elements are obligatorily necessary to constitute a flow vector sensor. On the other hand, the number of sensor outputs also determines the measurement accuracy of the sensor; more redundant data help to enhance sensor accuracy. Layout designs of three-element sensor and four-element sensor are shown in Fig. 7. For a flow vector sensor with a round-shaped heating area, if there is no flow passing through, temperature distribution over the heating area is circular symmetry. When an external flow passes through the sensor, the temperature distribution over the sensor will be deflected in the flow direction, which results in a temperature difference between upstream and downstream that can be detected by the thermal elements distributed in the round area of the sensor, as shown in Fig. 8. Data fusion algorithm to extract the parameters (velocity and angle) of 2D flow vector from outputs of thermal elements in the sensor will be introduced in later section.

Dimension of a flow vector sensor is determined in consideration of spatial resolution, sensitivity, and fabrication. Sensing area of thermal flow sensor needs to be as small as possible in order to achieve high spatial resolution for capably detecting flow vector at one point rather than an average flow over an area. In addition, dimension of sensor also depends on fabrication and the element resistance that relate to sensitivity of sensor. In general, smaller resistances of thermal elements result in a lower sensitivity. Figure 9 shows layout designs and fabricated prototypes of flow vector sensors.

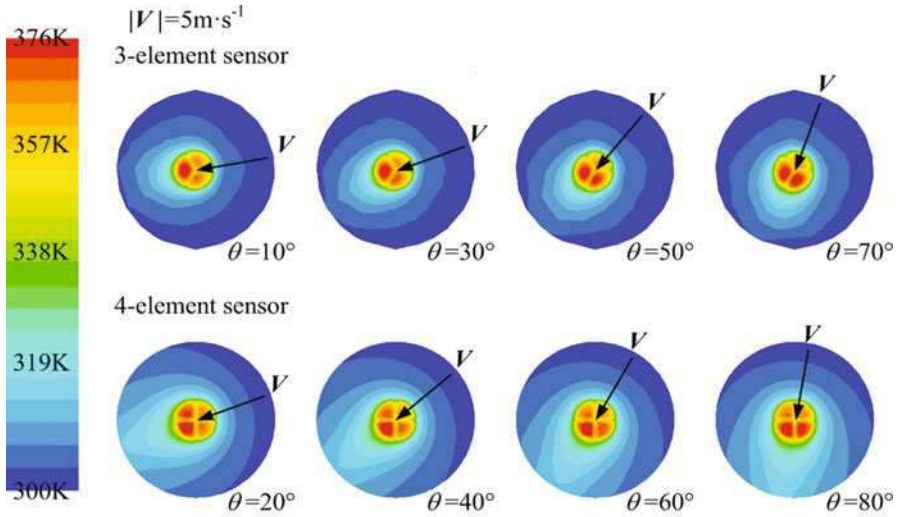


Fig. 8 Temperature distribution over the surface of a round heating area

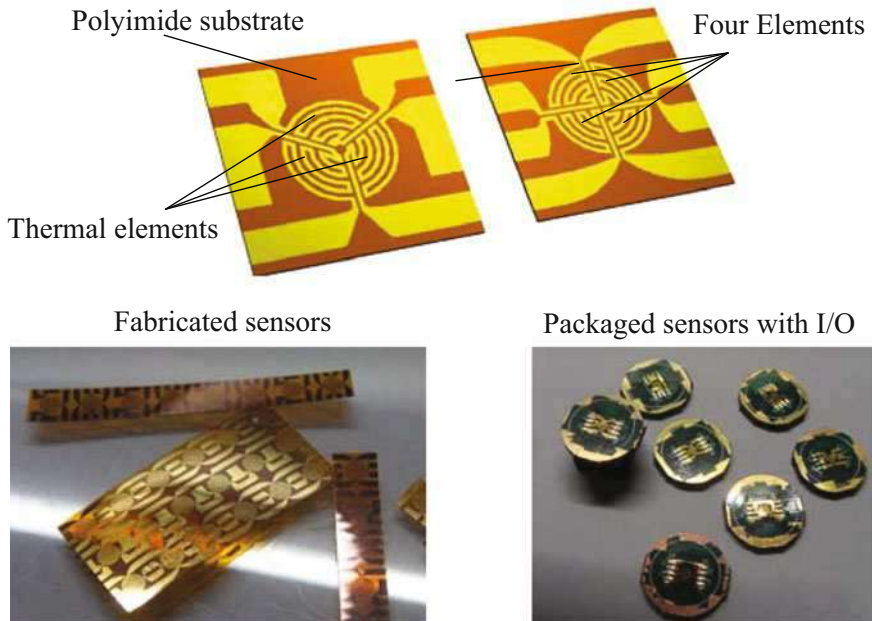


Fig. 9 Prototypes of fabricated flow vector sensors

## Measurement of Thermal Flow Sensor

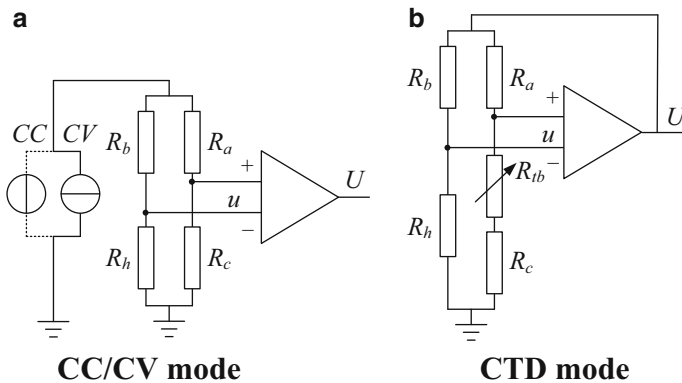
### Working Operation of Flow Sensor

A hot-film sensor can be operated either in constant voltage (CV), constant current (CC), and constant temperature difference (CTD) modes. In CV mode, heating voltage applied to thermal element is held constant, the circuit is easy to be realized, but the sensitivity is low. In CC mode, heating current is held constant, and its sensitivity is still low. Comparatively, in CTD mode, the temperature difference between sensing element and environment is held constant by feedback circuitry and the increase in heating power required for maintaining constant temperature is measured in response to flow. Feedback provides automatically adjusting electronic compensation for thermal inertia of sensing element when its operating point varies. Major advantage of CTD mode is that frequency response is quicker and sensitivity is higher than that of other modes (Bruun 1995). Figure 10 shows schematic views of CC, CV, and CTD modes, where a sensing element is placed on one leg of a Wheatstone bridge.  $R_h$  refers to the resistance of the hot-film sensor,  $R_a$ ,  $R_b$  are the reference resistances in the bridge,  $R_c$  is a temperature-compensating sensor, and  $R_{tb}$  is a resistance used to adjust joule heating level of thermal flow sensor.

Heat balance equation for a thermal element under electrical joule heating is

$$Q_s = Q_{\text{gen}} - Q_{\text{conv}} - Q_{\text{cond}} \quad (3)$$

where  $Q_s$  denotes rate of heat storage,  $Q_{\text{gen}}$  denotes generated power from Joule heating,  $Q_{\text{conv}}$  denotes rate of heat loss due to forced convection, and  $Q_{\text{cond}}$  denotes sum of conduction losses. Under a given  $Q_{\text{gen}}$  and fluid flow rate, it is important to minimize  $Q_{\text{cond}}$  in order to achieve a greater sensitivity against velocity changes. In a steady state, thermal flow sensor reaches a thermal equilibrium, and  $Q_{\text{cond}}$  is much smaller than  $Q_{\text{conv}}$ , approximately  $Q_{\text{gen}} = Q_{\text{conv}}$ .



**Fig. 10** Schematic views of CC, CV, and CTD modes for operating sensors. (a) CC/CV mode (b) CTD mode

Convective heat transfer of a finite long hot-film element can be expressed by the following equation according to King's law (King 1914):

$$Q_{\text{conv}} = \alpha \cdot \Delta T(A + BV^n) \quad (4)$$

where  $V$  is flow velocity,  $\alpha$  is TCR of the hot-film element,  $\Delta T$  is working temperature higher than ambient environment,  $\alpha \cdot \Delta T$  is defined as overheat ratio,  $A$  and  $B$  are constants determined by material and geometric features of sensing elements, and  $n$  is a geometric factor about 0.2 ~ 0.5. Further considering the power from joule heating, the output of hot-film flow sensor operated in a CTD circuit can be deduced by the following equation:

$$U^2 = R_h(1 + \beta)^2 \Delta T(A + BV^n) \quad (5)$$

It is shown that the sensitivity of the hot-film flow sensor is proportional to the resistance of  $R_h$ , working temperature  $\Delta T$ , bridge parameter  $\beta = R_b/R_h$ , and  $B \propto \lambda(\rho/\mu)^{0.5}(\mu C_p/\lambda)^{0.33}$ , where  $\rho$ ,  $\mu$ ,  $C_p$ ,  $\lambda$  are fluid density, viscosity coefficient, specific heat capacity, and coefficient of heat conduction. For different fluid medium, the hot-film flow sensor bears different sensitivity due to different fluid properties of the flow. Since water has much higher density and specific heat capacity than air and possesses much smaller viscosity coefficient, the flow sensitivity of the sensor in water is much higher than airflow sensitivity at the same operating condition.

## Temperature Compensation

Obviously, a thermal flow sensor is suffered from ambient temperature fluctuation, and a temperature compensation method needs to be considered for overcoming the sensor temperature drift. One general method is called numerical compensation, entails a complete calibration of the flow sensor over entire measuring range of flow velocities and temperatures, and corrects the sensor output using the temperature info based on the temperature calibration (Hultmark and Smits 2010; Lekakis 1996). Another approach to realize temperature compensation is called analogy compensation (Bruun 1995), where a temperature resistor is put directly into Wheatstone bridge to neutralize temperature drift as shown in Fig. 10b. Compared with the numerical method, this method is straightforward and automatic and thus with fast response. The analogy method requires the temperature resistor has an equal TCR with that of the hot-film element in order to realize temperature compensation, which implies that the temperature resistor and the hot-film flow element have to be made of the same material and with good uniformity in fabrication.

## Basic Principle for Analogy Compensation

When hot-film element and temperature-compensating resistor have equal TCRs and  $R_c \gg R_h$  in Fig. 10b, the working resistances of two thermal resistors can be expressed by

$$R_h = R_{h0}[1 + \alpha(T_a + \Delta T)] \quad (6)$$

$$R_c = R_{c0}(1 + \alpha T_a) \quad (7)$$

where  $R_{h0}$  denotes the resistance of the hot-film element  $R_h$  at  $0^\circ\text{C}$ ,  $T_a$  denotes ambient temperature,  $\alpha$  denotes TCR of the hot-film resistor and temperature-compensating resistor, and  $R_{c0}$  denotes the resistance of temperature-compensating resistor  $R_c$  at  $0^\circ\text{C}$ . Set  $R_{c0}/R_{h0} = R_a/R_b$ . When Wheatstone bridge circuit at CTD mode shown in Fig. 10b reaches balance, the following relationship comes into existence:

$$R_h \cdot R_a = (R_c + R_{tb}) \cdot R_b \quad (8)$$

Introducing Eqs. (6) and (7) into Eq. (8) gives

$$R_{tb} = \frac{R_a}{R_b} \cdot R_{h0} \cdot \alpha \cdot \Delta T \quad (9)$$

The above equation indicates that the working temperature difference between the hot-film element and ambient temperature can be held constant by setting a certain value of  $R_{tb}$ . The CTD mode ensures elimination of temperature drift for hot-film flow sensors. The above requirement  $R_c \gg R_h$  is necessary to ensure the temperature-compensating resistor in the bridge circuit serving as an ambient temperature sensor. The reason is that under an equal applied voltage  $U$ , the resistor  $R_h$  with a small resistance bears a higher electrical power and thereby is heated to be a hot film, while the resistor  $R_c$  with a large resistance bears a lower joule heating and thereby meet cold demands for ambient temperature sensor. In practical application, set  $R_c/R_h \geq 5$ .

### Temperature Compensation Using Temperature Sensors Independent of Thermal Flow Sensors

For the cases that temperature sensor and hot-film flow sensor are independent with each other, which means the two resistors have different TCRs, the above compensation method fails to deal with the temperature drift. Another compensation method by configuring Wheatstone bridge based on two TCRs of the temperature resistor and the thermal flow sensor is introduced below (Que et al. 2011).

Refer to Fig. 10b, assume that the flow sensor  $R_h$  and the temperature-compensating sensor  $R_c$  have different TCRs,  $\alpha_h$  and  $\alpha_c$ . Temperature compensation approach mainly involves the design of  $R_{tb}$ ,  $R_a$ , and  $R_b$ . The ratio  $\eta = u/U$  is defined as an objective function of the bridge. The goal of the temperature compensation is to keep the ratio  $\eta$  independent of temperature variation  $T_a$ .

For the circuit in Fig. 10b, the parameter  $\eta$  can be formulated as follows:

$$\eta = \frac{R_b}{R_h + R_b} - \frac{R_a}{R_a + R_c + R_{tb}} \quad (10)$$

$$\text{i.e. } \eta = \frac{R_b \cdot R_c + R_{tb} \cdot R_b + R_a \cdot R_h}{(R_h + R_b) \cdot (R_a + R_c + R_{tb})}$$

Define numerator ( $Num$ ) and denominator ( $Den$ ) of  $\eta$  as follows:

$$Num = R_c \cdot R_b + R_{tb} \cdot R_b - R_a \cdot R_h \quad (11)$$

$$Den = (R_h + R_b) \cdot (R_a + R_c + R_{tb}) \quad (12)$$

Perform partial derivative of  $\eta$  with respect to temperature  $T_a$  as follows:

$$\frac{\partial \eta}{\partial T_a} = \left[ \frac{\partial Num}{\partial T_a} \cdot Den - Num \cdot \frac{\partial Den}{\partial T_a} \right] \cdot Den^{-2} \quad (13)$$

According to Eqs. (8) and (9) and (11), (12), and (13), the following relationships shown in Eqs. (14) and (15) can keep  $\partial \eta / \partial T_a = 0$ , which meet the requirement for the temperature compensation.

$$\frac{\partial Num}{\partial T_a} = (R_{c0} \cdot \alpha_c \cdot R_b - R_a \cdot R_{h0} \cdot \alpha_h) = 0 \quad (14)$$

$$Num = (R_c \cdot R_b + R_{tb} \cdot R_b - R_a \cdot R_h) = 0 \quad (15)$$

According to Eq. (14), the first criterion for the temperature compensation is given by

$$\frac{R_a}{R_b} = \frac{R_{c0} \cdot \alpha_c}{R_{h0} \cdot \alpha_h} \quad (16)$$

Besides Eq. (16), the bridge ratio  $R_a/R_b$  generally needs to be set larger than 5 for fulfilling  $R_c/R_h \geq 5$  to realize cold demand for the temperature sensor.

According to Eq. (15), the second criterion for the temperature compensation is given by

$$\frac{R_c + R_{tb}}{R_h} = \frac{R_a}{R_b} \quad (17)$$

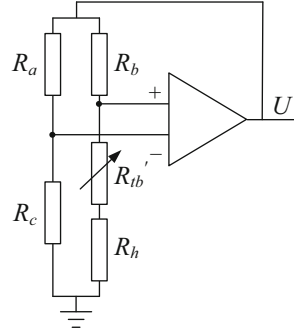
Introduce Eqs. (8) and (9) and (16) into Eq. (17);  $R_{tb}$  is determined as follows:

$$R_{tb} = R_{c0} \left( \frac{\alpha_c}{\alpha_h} + \alpha_c \cdot \Delta T - 1 \right) \quad (18)$$

On condition that the term  $\alpha_c/\alpha_h + \alpha_c \cdot \Delta T - 1$  is positive,  $R_{tb}$  can be availablely determined by using Eq. (18). In most cases, the temperature sensor has a higher TCR than that of the thermal flow sensors (i.e.,  $\alpha_c > \alpha_h$ ), giving  $\alpha_c/\alpha_h + \alpha_c \cdot \Delta T - 1 > 0$ , and then  $R_{tb}$  is set using Eq. (18). However, in an inverse case, where  $R_{tb}$  becomes negative, e.g., when the temperature sensor has a much lower TCR, the compensation scheme needs to be modified, and a Wheatstone circuit shown in Fig. 10b needs be changed to Fig. 11.



**Fig. 11** Modified circuit when  $R_{tb}$  calculated by Eq. (17) is negative (Que et al. 2011)



In Fig. 11, the resistor  $R_{tb}'$  is connected in series with  $R_h$  and is calculated by:

$$R_{tb}' = R_{h0} \left( \frac{\alpha_h}{\alpha_c} - \alpha_h \cdot \Delta T - 1 \right) \quad (19)$$

Equation (19) can be deduced following a similar procedure to that for Eq. (18).

One example is shown in Fig. 12, where the experimental results indicate that the sensor temperature drift can be reduced from 22% to less than 0.6% in a temperature range of 35 °C ~ 90 °C.

This compensation method provides a flexible means for a thermal flow sensor. The temperature sensor does not have to possess an equal TCR to that of the thermal flow sensor. Therefore, any inside and outside temperature sensor can be used and thus simplifies sensor operation and expands the range of usage.

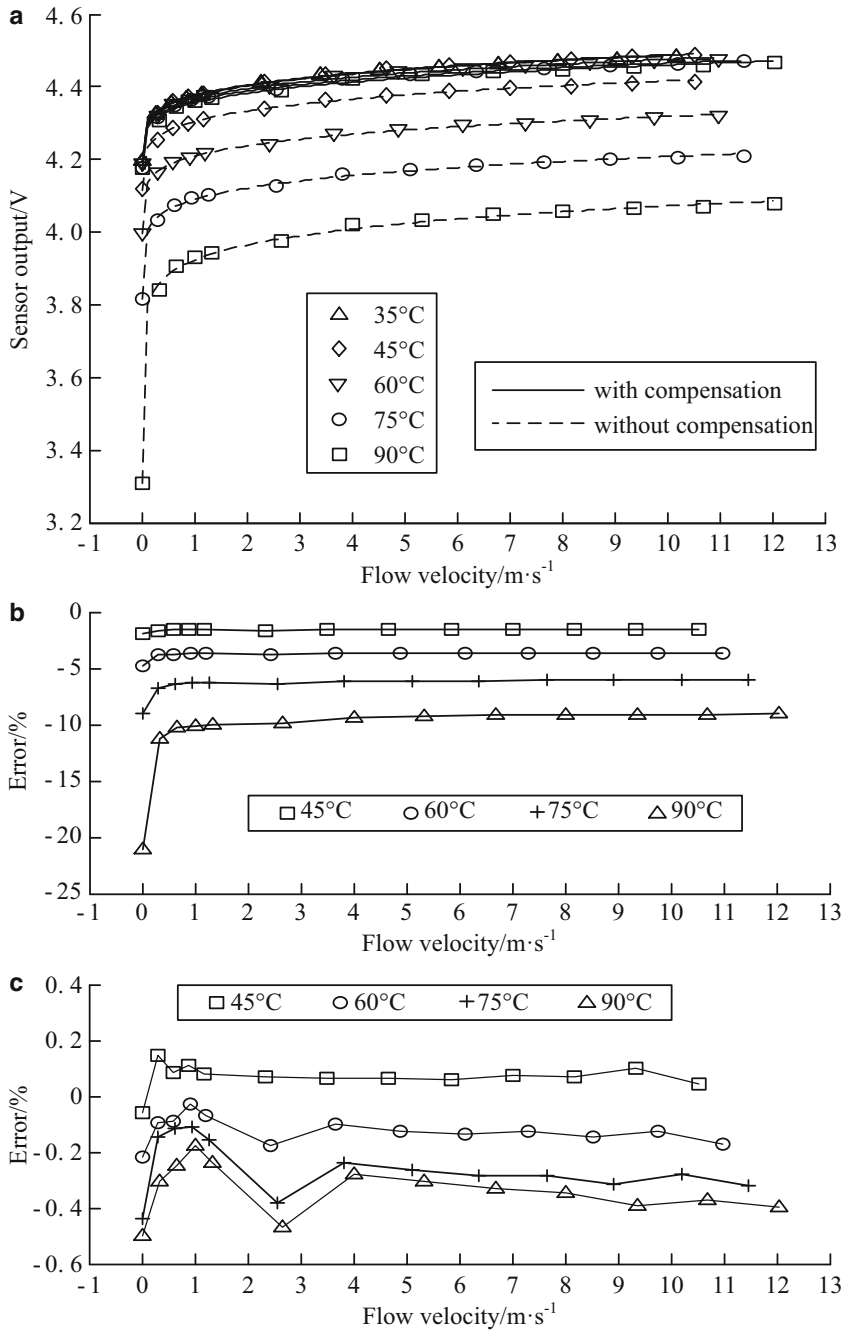
---

## Applications

In this section, some applications using hot-film flow sensors are presented. The application scenarios include surface airflow detection for micro air vehicles, micro-anemometer for wind detection, and respiratory flow monitor.

### Surface Airflow Detection for Micro Air Vehicles

Small unmanned air vehicles (UAVs) and micro air vehicles (MAVs) are attracting growing interests for both military and civilian applications in decades (Shyy et al. 1999). However, they suffer from particular characteristics of small Reynolds number, low inertia, low flight speed, weak stability or instability, and so on (Lian et al. 2003), which make flight poor and prone to wind. Wind disturbances, which are generally comparable to the vehicle's flight speed, make the instantaneous flight Reynolds number, air speed, angle of attack, and angle of sideslip varied substantially, which directly lead to the instability of the flight. To improve performances of MAVs/UAVs, their aerodynamic parameters need to be surveyed by using flow



**Fig. 12** Experimental results for thermal flow sensor with and without the temperature compensation. (a) Sensor outputs at different temperature under different flow rate. (b) Error of sensor without temperature compensation. (c) Error of sensor with temperature compensation

sensors and considered in flying control systems as complementary information to inertial guiding systems and autopilots. Conventional techniques for detecting these parameters, mostly based on pitot tubes or electromechanical self-orienting vanes (Fang 1999; Hagen and Seidel 1994), usually protrude outside the aircrafts (Callegari et al. 2004), need hard mechanical ties to the body and/or intrusive pneumatic links inside, and would be easily damaged while landing. What's more, distributed flow measurement, especially distributing sensors on flapping wings or flexible membrane wings for detecting multiple aerodynamic parameters, calls for more space, power, and payload which can hardly be afforded by small aircrafts.

Redundant strip pressure sensors mounted on airfoil skins may be used to obtain aerodynamic parameters (Callegari et al. 2004, 2006); however, both sensing technology and data processing methods seem too complicated. A method described in Whitmore (1991) required a hemispherical nose installed on aircraft, which limited its application on MAVs.

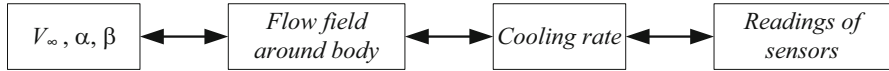
Aerodynamic parameters can be also inferred indirectly from surface flow field around aircraft wing using microsensors (Callegari et al. 2004; Callegari et al. 2006). Adoption of microsensors has shown many technical and economic advantages compared with other conventional technologies. In the past two decades, several research groups have demonstrated micromachined flow sensors based on various principles, including thermal transfer (Xu et al. 2003; Lin et al. 2004), torque transfer, and pressure distribution.

Requirements for microsensors utilizing on a small aircraft contain miniaturization, easily mounting arrays on nonplanar surface, low consumption, etc. Micro hot-film flow sensors are excellent candidates for implementing the applications.

For a certain aircraft, flow field around a wing surface depends on the free-stream air speed  $V_\infty$ , the angle of attack  $\alpha$ , and the angle of sideslip  $\beta$  (Xu et al. 2003; Lin et al. 2004; Fei et al. 2007). Therefore, it is reasonable to deduce these flight parameters by measuring the flow field around the wing surface. The flow field information around the wing contains two main components: flow velocity and wall shear stress. Considering a low Prandtl number of about  $0.7 \sim 0.8$  for air, both flow velocity and shear stress can be detected by hot-film flow sensors.

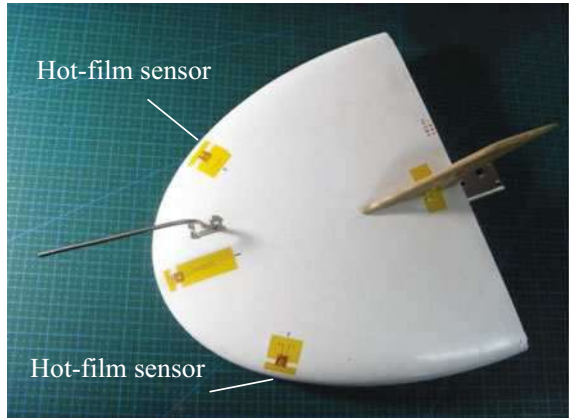
Evidently, it is almost impossible, and also unnecessary, to measure the whole flow field on an entire wing surface; flow information at several specific points may be sufficient for inferring the flight parameters. If there are a sufficiently large number of sensors, it is more probable that a good determination of the flight parameters can be obtained from the sensor readings. However, the number of sensors affects the weight, the power consumption, and the complexity of the data acquisition system – all of which should be strictly limited for small UAVs and MAVs. The number of sensors needs to be determined by taking account of a trade-off between simplicity and measurement accuracy of the sensing system. In general, four sensors are considered to be the minimum number required for inferring above three flight parameters. Figure 13 depicts the relationship between the readings of the thermal flow sensor attached on the wing surface and the three flight parameters.

Besides the sensor number, the sites where the sensors are mounted are also important. The work reported by Lian et al. (2003) indicated that on the front edge of



**Fig. 13** Relationship between the sensors’ readings and flight parameters

**Fig. 14** Collocation of hot-film flow sensor array on a MAV (Que and Zhu 2012)



the wing, flow separation and vortices seldom occurred; in addition, the flow field there changed dramatically with the change of air speed and flight aerodynamic angles; all of these implied that it was a good choice to mount the flow sensors on the front edge of the wing.

Based on the factors discussed above, a plurality of hot-film flow sensors is collocated on specific positions of the wing surface of a micro air vehicle. As illustrated in Fig. 14 (Que and Zhu 2012), hot-film sensors 1, 2, 3, and 4 are mounted symmetrically on the top left, top right, bottom left, and bottom right of the front edge of the wing, respectively.

Suppose  $n$  flow sensors are used, and denote their outputs in a vector form:

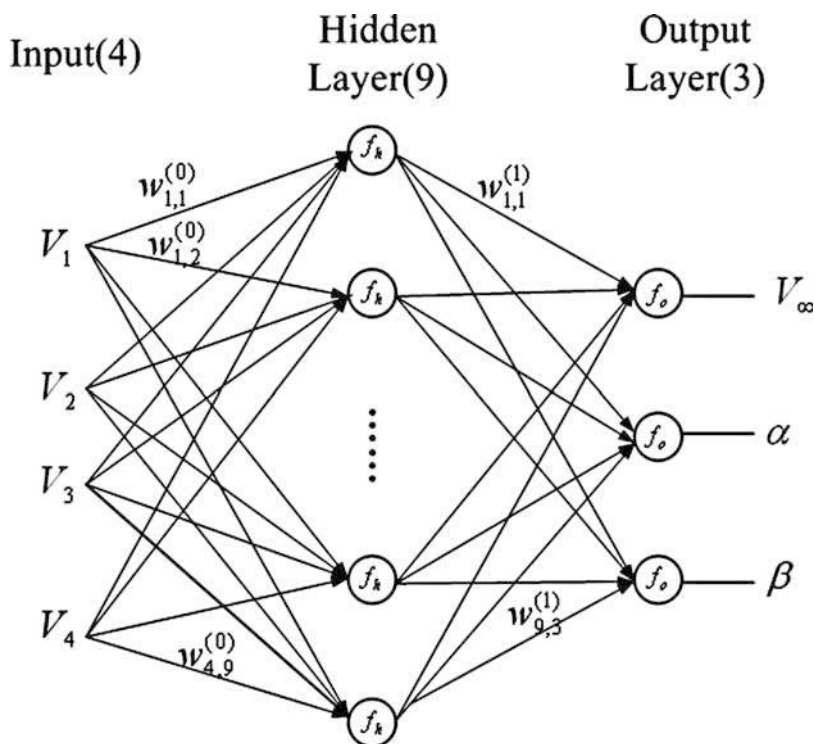
$$U = [U_1 \ U_2 \ \dots \ U_n]^T \tag{20}$$

A mathematic relation between the vector  $U$  and the three flight parameters is expressed by (to simplify the analysis, the second-order dependence of  $U$  on other parameters such as the air density is not considered):

$$U = f(V_\infty, \alpha, \beta) \tag{21}$$

where  $f$  represents the relationship model depending on aerodynamic characteristics of the air vehicle. As a matter of fact,  $f$  is generally a complex multivariate and nonlinear function, which plays a crucial role in the detection approach. An approximation of  $f$  can be obtained by various means, such as an analytical method using intrinsic fluid dynamic models, simulation, experiment-based identification,

or a mixture of the above. Theoretically, the analytical method is too complicated and inefficient (different airfoil results in different  $f$ ); simulation is feasible in some sense but is generally less valid; in contrast, the experimental identification is more reliable and effective. An experiment-based model identification method is adopted to develop  $f$ , by using a three-layer BP neural network shown in Fig. 15 as the model structure considering that it had been theoretically proved three layers of neural network could solve arbitrarily complicated nonlinear mapping problems (Clifford and Lau 1992). For applications on small UAVs and MAVs, which require both accuracy and simplicity, a three-layer BP neural network with nine neurons in the hidden layer is used, where the number of hidden neurons is determined through experimental trials. The parameters in the neural network are determined through wind tunnel calibrations. And the constructed network is used to infer the flight parameters in real time from the readings of the flow sensors. The literature (Que and Zhu 2012) demonstrates one example of testing, which indicates measurement errors of air speed, angle of attack, and angle of sideslip using the hot-film flow sensors that are approximately 0.27 m/s, 0.87°, and 0.27°, respectively, which are applicable for MAVs and UAVs.



**Fig. 15** Structure of the neural network for acquiring the flight parameters (Que and Zhu 2012)

## Micro anemometer for 2D Wind Detection

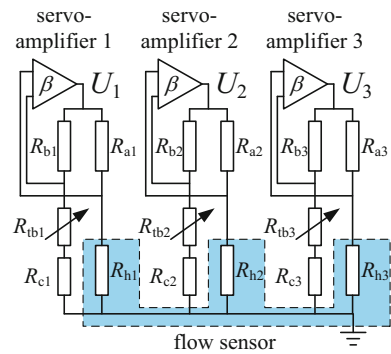
Two-dimensional (2D) fluid flow measurements are becoming more and more important in many applications, e.g., aircrafts (Fei et al. 2007; Que and Zhu 2012; Riedl et al. 2013; Callegari et al. 2008), underwater vehicles, meteorology, and agriculture. 2D ultrasonic anemometer using the directivity angle of an ultrasonic sensor was a valid way (Han et al. 2008), but the miniaturization was challenged. Using a micromachining technique to develop microsensors and array for measuring 2D flow has attracted most of attentions. Bruschi et al. (2009) proposed a micro-structure using a cylinder with a channel network connecting internal flow sensors with the lateral surface. The wind-induced pressure distribution on the cylinder's surface was detected, from which wind speed and direction were unequivocally determined. Another commonly used approach is to use multiple micro-heaters and temperature detectors patterned on a silicon substrate to detect flow-induced temperature differences, by which the flow velocity and direction angle were deduced (Kim et al. 2004; Van Oudheusden 1990; de Bree et al. 1999; Van Oudheusden and Van Herwaarden 1990; Shen et al. 2009). These sensors usually consisted of multiple isolated heaters and thermometers connecting with separate electrical pads; some of the sensors used thermal isolation structures (e.g., diaphragm) (Kim et al. 2004; Van Oudheusden and Van Herwaarden 1990) to reduce thermal capacity and thus made the sensor structure relatively complicated and large, the spatial resolution low, and the electrical wire bonding difficult as well.

An alternative approach proposed by Que and Zhu (2014; 2015) used a compact design of flow vector sensor with three or four thermal elements, each of which was composed of a roundabout wire distributed in a quadrant of a circle fabricated on a flexible substrate as described in section “[Geometric Design of Hot-Film Elements.](#)” The thermal elements were served as both joule heaters and temperature detectors. The flow velocity and direction could be deduced using a data fusion algorithm. This micro flow vector sensor had small size, simple structure, and easy fabrication. An application case by using the micro flow vector sensor to detect 2D wind is introduced here.

A scheme of the CTD mode conditioning circuits for operating the sensor with three thermal elements is shown in Fig. 16, where  $R_{h1\sim3}$  are the resistors of the sensor elements,  $R_{c1\sim3}$  are the temperature resistors used for temperature compensation,  $R_{tb1\sim3}$  are the resistors used to adjust joule heating of  $R_{h1\sim3}$ , and  $R_{a1\sim3}$  and  $R_{b1\sim3}$  are the resistors used for balancing the bridges. The Wheatstone bridge voltages  $U_1$ ,  $U_2$ , and  $U_3$  (i.e., heating voltages applied to the bridges) are served as the outputs of the sensors. Under CTD mode, the temperatures of the thermal elements keep constant by adjusting the heat powers (represented by heating voltage  $U$ ), which are outputs of the sensors. Therefore, it is reasonable to detect the 2D flow parameters (velocity and direction denoted as  $[V, \theta]$ ) by measuring the heating voltages of the thermal elements denoted as  $U_1$ ,  $U_2$ , and  $U_3$  for the three-element sensor and  $U_1$ ,  $U_2$ ,  $U_3$ , and  $U_4$  for the four-element sensor, where  $V$  is flow velocity and  $\theta$  is flow angle.

The data fusion algorithm for deducing the flow velocity and direction angle from the sensor's multiple outputs is very important for the flow vector sensor. An

**Fig. 16** Scheme of the conditioning circuit for micro flow vector sensor with three elements (Que and Zhu 2015)

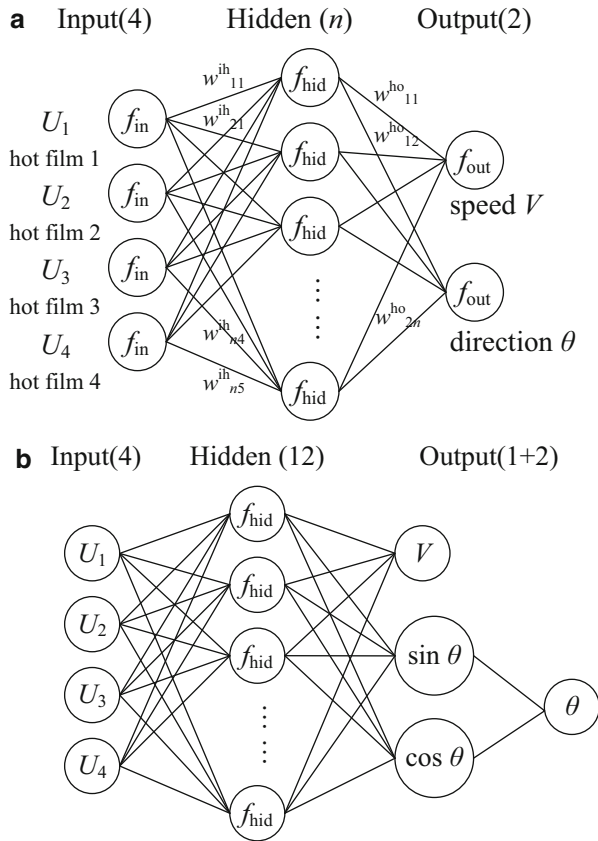


experiment-based model identification method can be adopted to develop the relationship model between multiple sensor outputs and multiple flow parameters. A three-layer back propagation (BP) neural network is a commonly used model structure to solve MIMO problem, as shown in Fig. 17a. When establishing the BP neural network, the direction angle of the flow vector is not a good output variable for the network because when the vector angle varies from  $0^\circ$  to  $360^\circ$ , the output of the sensing element follows approximately sine or cosine law that is not consistent with the monotonicity of angle variation. Therefore, using sine and cosine of the flow angle instead of the angle itself as the outputs of the network is more appropriate for deducing the flow vector as shown in Fig. 17b. The flow angle can be figured out by synthesizing  $\sin\theta$  and  $\cos\theta$ .

One experimental example is shown in Fig. 18 (Que and Zhu 2015), which demonstrates the outputs of the thermal elements  $U_1$ ,  $U_2$ , and  $U_3$  for three-element sensor and  $U_1$ ,  $U_2$ ,  $U_3$ , and  $U_4$  for four-element sensor against the flow velocity of  $3 \sim 30$  m/s and the flow angle of  $0^\circ \sim 360^\circ$ . For three-element sensor, the mean square error of the flow velocity reaches 0.65 m/s in the range of  $3 \sim 30$  m/s, and the error of the flow angle reaches  $2.69^\circ$  in the range of  $0^\circ \sim 360^\circ$ . For four-element sensor, the error of the flow velocity reaches 0.59 m/s, and the error of direction angle reaches  $0.96^\circ$ . It is seen that four-element sensor possesses better accuracy than three-element sensor.

Due to the compact design of the flow vector sensor, it enables the feasible implementation of a wearable anemometer by integrating a low-power Bluetooth and a personal mobile device with the sensors. Figure 19 shows the wearable anemometer (Zhao et al. 2016), comprised of a micro flow vector sensor, a signal conditioning circuit, a Bluetooth LE module, and a rechargeable battery, all of which are packaged in a watch-like case. The sensor data are transmitted wirelessly to a smartphone, where the sensor signals are collected and further used to execute the extraction of flow velocity and angle. The anemometer is small, wireless, and low power configured for a long-term usage, which can be applied for personal electronics, and distributed flow measurements in agricultural and meteorological monitoring.

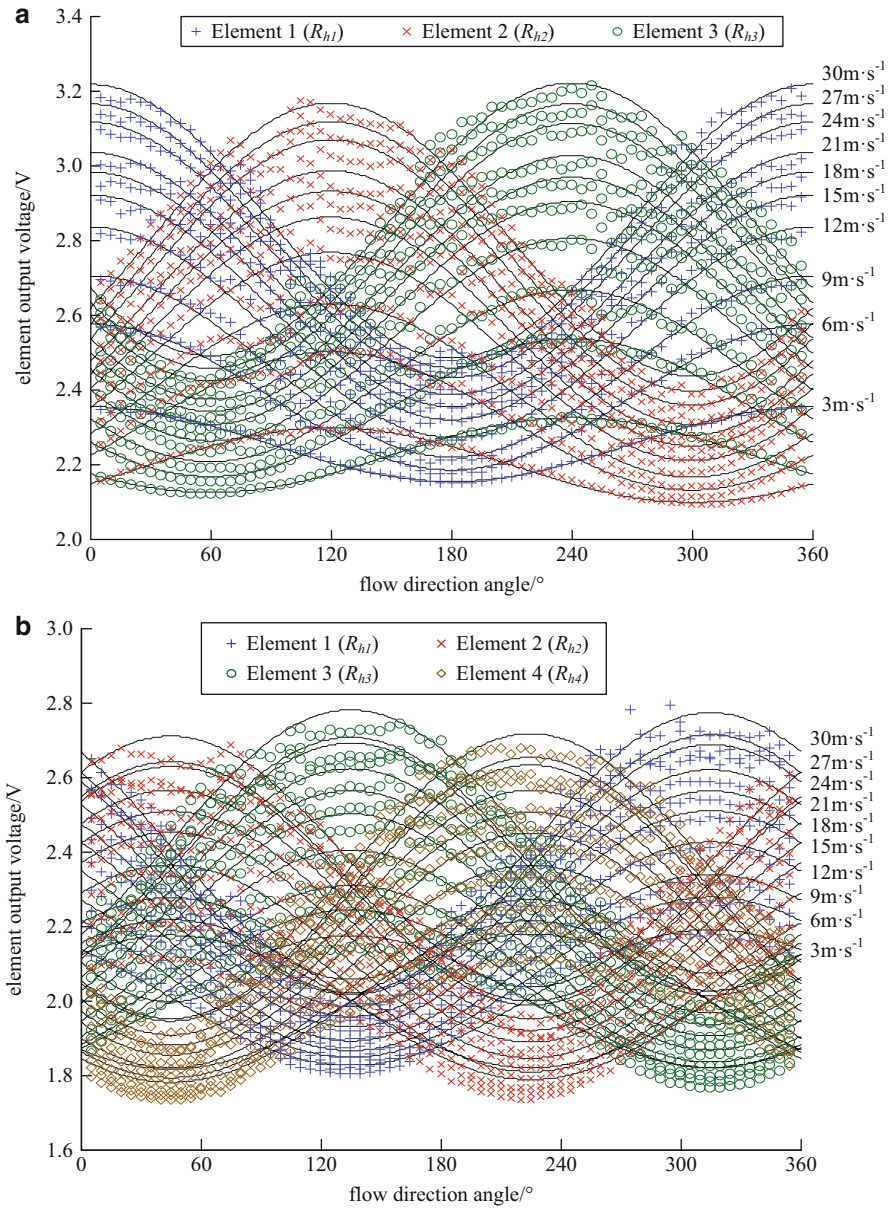
**Fig. 17** (a) Structure of BP neural network. (b) Improved network structure with sine and cosine as outputs (Que and Zhu 2014)



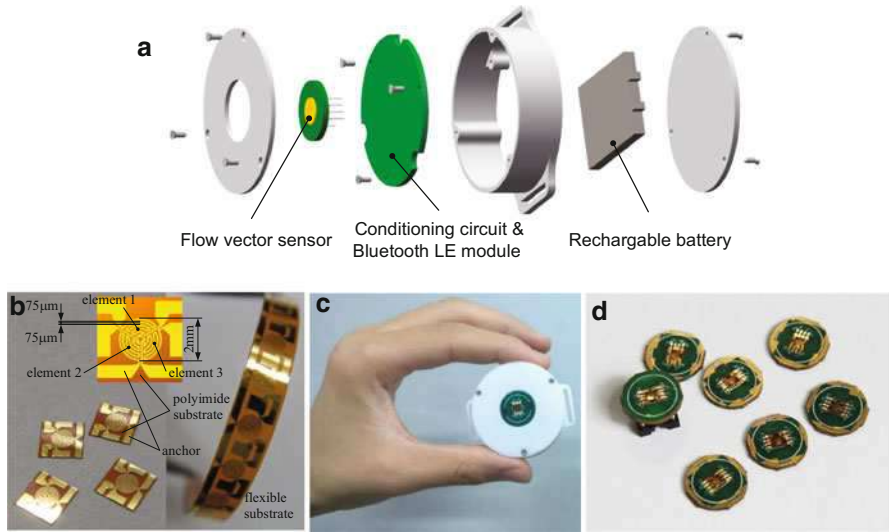
### Respiratory Airflow Monitor

The micro thermal flow sensor can be also used to detect human respiratory airflow for monitoring and diagnosing respiratory diseases, e.g., obstructive sleep apnea syndrome (OSAS), which is the most common chronic respiratory disease that bring serious risks to human health. To diagnose OSAS, traditional equipment called polysomnography (PSG) (Kushida et al. 2005) is used in hospitals and operated by physicians, which is relatively inconvenient and expensive. Consequently, portable sleep monitor (PSM) has been proposed as a substitute for PSG in pre-diagnostic assessment on patients with suspected OSAS by using micro-flow sensors (Tardi et al. 2015; Cao et al. 2012). Figure 20 illustrates a portable sleep monitor using a micro flow sensor to monitor respiratory airflow, a tri-axis micro-accelerometer to monitor body posture, and a micro-photoelectric sensor to monitor blood oxygen saturation (Cao et al. 2012). The sensor signals are wirelessly transmitted to a mobile device to execute preliminary assessment and further transmitted to a remote monitoring terminal for professional diagnosis.

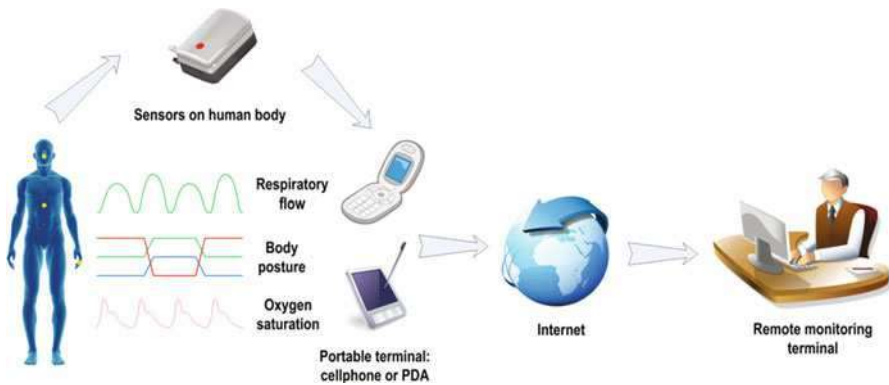




**Fig. 18** Output voltages of sensor versus the flow velocity and flow direction (flow velocities are labeled for element 1 only) (Que and Zhu 2015). **(a)** Three-element sensor. **(b)** Four-element sensor

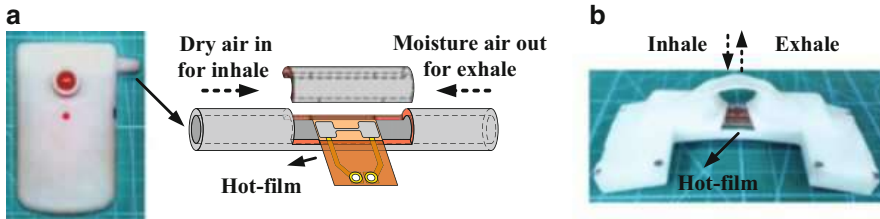


**Fig. 19** Configuration and prototype of a wearable anemometer (Zhao et al. 2016). (a) Configuration of anemometer. (b) Micro flow vector sensor. (c) Prototype of anemometer. (d) Sensor with I/O interface circuit



**Fig. 20** Configuration of wireless portable sleep monitor

There are several kinds of respiratory sensors that have been reported. One is indirectly monitoring respiration from variation in thorax volume; the other is using airflow sensor to directly detect respiratory airflow (Folke et al. 2003). The former method includes transthoracic impedance sensors, micro- and macro-bending fiber-optic sensors (Dziuda 2015), 3D image sensors, and ultrasonic sensors, which could detect the variation of chest or abdominal circumference. However, the indirect methods are easily disturbed by body movements, which result in invalid measurements. Respiratory airflow is a key vital signal for monitoring respiratory disorder.



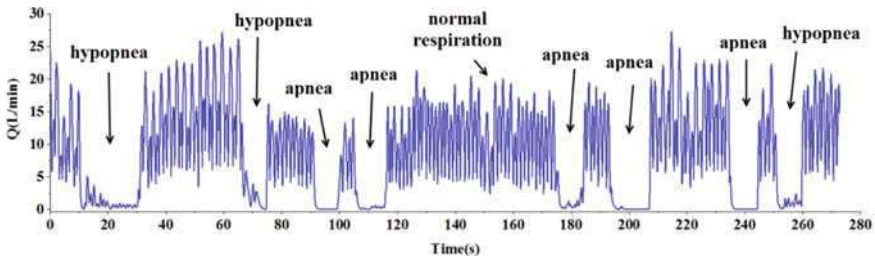
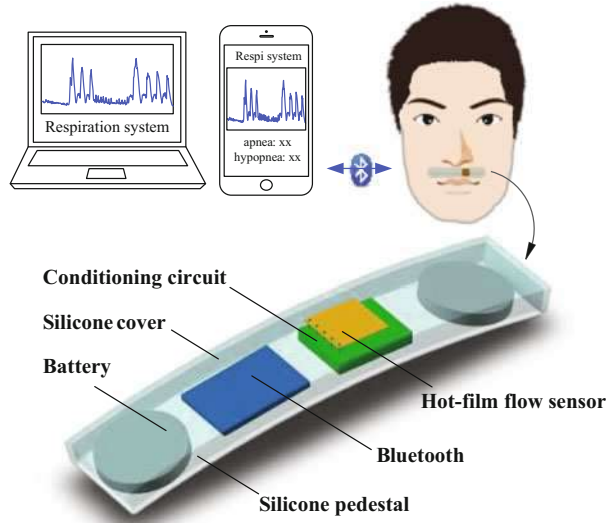
**Fig. 21** Usage modes of micro hot-film flow sensors for respiration monitoring. (a) Hot-film installation in a tube. (b) Hot-film installation on surface of sensing strip adhered underneath nostril

The apnea and hypopnea indexes are important in diagnosis of OSAS. In addition, acquiring comprehensive respiratory airflow parameters, such as minute ventilation (MV), peak inspiratory flow (PIF), respiratory rate (RR), and tidal volume (TV), is also necessary in diagnoses of OSAS, chronic obstructive pulmonary disease, and asthma. Kinds of airflow sensors are used to monitor respiratory airflow, such as pressure sensor, hot-wire/hot-film sensor, and ultrasonic flowmeter (Schena et al. 2015). The differential pressure flow sensor and the hot-wire/hot-film flow sensor are usually used in the portable medical instruments and wearable devices. Differential pressure flow sensors need a pipeline and diaphragm structure. Comparatively, micro hot-wire/hot-film sensors have small sizes, simple structures, high sensitivities and are low cost.

To be mentioned, two types of sensing modes are proposed for micro thermal flow sensor used in respiration monitor as shown in Fig. 21. One typical use is installing a microthermal flow sensor into a tube as shown in Fig. 21a. A nasal cannula that has two prominent small pipes inserted into nostrils collects nasal airflow and delivers the airflow to the tube where the flow sensor is installed (Cao et al. 2012). However, the narrow tube generally hinders nasal airflow and increases respiratory resistance, thereby causes unsmooth and uncomfortable breathing for the patient. In addition, the long nasal cannula is easily squeezed by the human body and twined with limbs. Another sensing mode is called noninvasive monitor shown in Fig. 21b, where a flexible thermal flow sensor is adhered on a sensing strip that can be attached underneath the patient's nostrils to detect inhaling and exhaling airflows in real time without the use of nasal cannula and mask (Jiang et al. 2015).

Figure 22 shows a configuration of a wearable noninvasive respiration monitor, which is comprised of a micro hot-film flow sensor, a signal conditioning circuit, a Bluetooth 4.0 LE module, and a rechargeable battery (Jiang et al. 2015). The monitor is wirelessly connected with a mobile device (e.g., smartphone) via Bluetooth. The monitor is attached on the underneath of the nostrils. The respiratory flow breathed out and into the nostrils blows over the strip and is detected by the thermal flow sensor adhered on the strip. Tube-free configuration makes the carriers comfortable. Figure 23 demonstrates an experimental result for monitoring apnea and hypopnea events using the respiration monitor (Jiang et al. 2015).

**Fig. 22** Noninvasive respiratory monitoring (Jiang et al. 2015)



**Fig. 23** Apnea and hypopnea signals detected by using micro hot-film flow sensor (Jiang et al. 2015)

### Concluding Remarks

Micro thermal flow sensors have wide applications, which have been actualized with simple structures and low-cost fabrication. Compared with other flow sensors, the hot-film flow sensor is compatible with flexible substrate and thus can be structured as a flexible sensor for various wearable applications. A layout design for the hot-film flow sensor is very important to expand its usages; besides, sensitivity and response time as well as power consumption are significantly concerned issues. Large-area fabrication is now carried through by using such as roll-to-roll technology. It is obvious that future prospects are promising for micro flexible hot-film flow sensors and system.

**Acknowledgments** The works were supported by National Natural Science Foundation of China (Grant No. 51735007) and National High-tech Program “863” of China (Grant No. 2012AA02A604 and 2006AA04Z257).

---

## References

- Bruschi P, Dei M, Piotta M (2009) A low-power 2-D wind sensor based on integrated flow meters. *IEEE Sensors J* 9(12):1688–1696
- Bruun HH (1995) Hot-wire anemometry-principles and signal analysis. Oxford University Press, Tokyo/Oxford/New York
- Callegari S, Talamelli A, Zagnoni M, Golfarelli A, Rossi V, Tartagni M, Sangiorgi E (2004) Aircraft angle of attack and air speed detection by redundant strip pressure sensors. In: Proceedings of IEEE sensors. IEEE, pp 1526–1529
- Callegari S, Zagnoni M, Golfarelli A, Tartagni M, Talamelli A, Proli P, Rossetti A (2006) Aircraft angle of attack and air speed detection by redundant strip pressure sensors. *Sensors Actuators A* 130/131:155–165
- Callegari S, Zagnoni M, Golfarelli A, Tartagni M, Talamelli A, Proli P, Rossetti A (2008) Experiments on aircraft flight parameter detection by on-skin sensors. *Sensors Actuators A Phys* 130-131:155–165
- Cao Z, Zhu R, Que R-Y (2012) A wireless portable system with microsensors for monitoring respiratory diseases. *IEEE Trans Biomed Eng* 59(11):3110–3116
- Chen J, Liu C (2003) Development and characterization of surface micromachined, out-of-plane hot-wire anemometer. *J Microelectromech Syst* 12(6):979–988
- Chen J, Fan Z, Zou J, Engel J, Liu C (2003) Two-dimensional micromachined flow sensor array for fluid mechanics studies. *J Aerosp Eng* 16(2):85–97
- Clifford L, Lau Y (1992) Neural networks: theoretical foundations and analysis. IEEE, New York
- de Bree H-E, Jansen HV, Lammerink TS, Krijnen GJ, Elwenspoek M (1999) Bi-directional fast flow sensor with a large dynamic range. *J Micromech Microeng* 9(2):186–189
- Dziuda Ł (2015) Fiber-optic sensors for monitoring patient physiological parameters: a review of applicable technologies and relevance to use during magnetic resonance imaging procedures. *J Biomed Opt* 20(1):010901–010901
- Fang Z (1999) Aircraft flight dynamics and automatic flight control. National Defense Industry Press, Beijing
- Fei H, Zhu R, Zhou Z, Wang J (2007) Aircraft flight parameter detection based on a neural network using multiple hot-film flow speed sensors. *Smart Mater Struct* 16(4):1239–1245
- Folke M, Cernerud L, Ekström M, Hök B (2003) Critical review of non-invasive respiratory monitoring in medical care. *Med Biol Eng Comput* 41(4):377–383
- Hagen FW, Seidel H (1994) Deutsche airbus flight test of rosemount smart probe for distributed air data systems. *IEEE Aerosp Electron Syst Mag* 9(4):7–14
- Han D, Kim S, Park S (2008) Two-dimensional ultrasonic anemometer using the directivity angle of an ultrasonic sensor. *Microelectron J* 39(10):1195–1199
- Hultmark M, Smits AJ (2010) Temperature corrections for constant temperature and constant current hot-wire anemometers. *Meas Sci Technol* 21(10):105404
- Jiang P, Zhao S, Zhu R (2015) Smart sensing strip using monolithically integrated flexible flow sensor for noninvasively monitoring respiratory flow. *Sensors* 15(12):31738–31750
- Kim S, Nam T, Park S (2004) Measurement of flow direction and velocity using a micromachined flow sensor. *Sensors Actuators A Phys* 114(2):312–318
- King LV (1914) On the convection of heat from small cylinders in a stream of fluid: determination of the convection constants of small platinum wires with applications to hot-wire anemometry. *Philos Trans R Soc Lond Ser A* 214:373–432. Containing Papers of a Mathematical or Physical Character

- Kushida CA, Littner MR, Morgenthaler T, Alessi CA, Bailey D, Coleman J Jr, Friedman L, Hirshkowitz M, Kapen S, Kramer M (2005) Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep* 28(4):499–521
- Lekakis I (1996) Calibration and signal interpretation for single and multiple hot-wire/hot-film probes. *Meas Sci Technol* 7(10):1313
- Lian Y, Shyy W, Viieru D, Zhang B (2003) Membrane wing aerodynamics for micro air vehicles. *Prog Aerosp Sci* 39(6):425–465
- Lin Q, Jiang F, Wang X-Q, Xu Y, Han Z, Tai Y-C, Lew J, Ho C-M (2004) Experiments and simulations of MEMS thermal sensors for wall shear-stress measurements in aerodynamic control applications. *J Micromech Microeng* 14(12):1640–1649
- Liu P, Zhu R, Liu X, Zhang F, Zhou Z (2009a) A low-cost integrated micro system for flow velocity and direction measurement. *Transducers 2009 - 15th International Conference on Solid-State Sensors, Actuators and Microsystems*, p 276–279
- Liu P, Zhu R, Que R (2009b) A flexible flow sensor system and its characteristics for fluid mechanics measurements. *Sensors* 9(12):9533–9543
- Nguyen N (1997) Micromachined flow sensors—a review. *Flow Meas Instrum* 8(1):7–16
- Nguyen N-T (2005) A novel thermal sensor concept for flow direction and flow velocity. *IEEE Sensors J* 5(6):1224–1234
- Que R, Zhu R (2012) Aircraft aerodynamic parameter detection using micro hot-film flow sensor array and BP neural network identification. *Sensors* 12(8):10920–10929
- Que R, Zhu R (2014) A two-dimensional flow sensor with integrated micro thermal sensing elements and a back propagation neural network. *Sensors* 14(1):564–574
- Que R-Y, Zhu R (2015) A compact flexible thermal flow sensor for detecting two-dimensional flow vector. *IEEE Sensors J* 15(3):1931–1936
- Que RY, Zhu R, Wei QZ, Cao Z (2011) Temperature compensation for thermal anemometers using temperature sensors independent of flow sensors. *Meas Sci Technol* 22(8):085404
- Riedl X, Leuckert J, Engert M, Kupke W, Wagner R, Nitsche W, Abbas A, Bauer K (2013) Transition measurement with microstructured hot film sensor arrays on a laminar flow airfoil model. In: *New results in numerical and experimental fluid mechanics VIII*. Springer, Berlin, pp 641–648
- Schena E, Massaroni C, Saccomandi P, Cecchini S (2015) Flow measurement in mechanical ventilation: a review. *Med Eng Phys* 37(3):257–264
- Shen G-P, Qin M, Huang Q-A (2009) A system-level model for a silicon thermal flow sensor. *Microsyst Technol* 15(2):279–285
- Shyy W, Berg M, Ljungqvist D (1999) Flapping and flexible wings for biological and micro air vehicles. *Prog Aerosp Sci* 35(5):455–506
- Tardi G, Massaroni C, Saccomandi P, Schena E (2015) Experimental assessment of a variable orifice flowmeter for respiratory monitoring. *J Sensors* 2015: 752540
- Van Baar J, Wiegerink R, Lammerink T, Krijnen G, Elwenspoek M (2001) Micromachined structures for thermal measurements of fluid and flow parameters. *J Micromech Microeng* 11(4):311
- Van Oudheusden B (1990) Silicon thermal flow sensor with a two-dimensional direction sensitivity. *Meas Sci Technol* 1(7):565–575
- Van Oudheusden B, Van Herwaarden A (1990) High-sensitivity 2-D flow sensor with an etched thermal isolation structure. *Sensors Actuators A Phys* 22(1–3):425–430
- Van Putten A, Middelhoek S (1974) Integrated silicon anemometer. *Electron Lett* 21(10):425–426
- Whitmore SA (1991) Development of a pneumatic high-angle-of-attack flush airdata sensing (HI-FADS) system; National Aeronautics and Space Administration, Ames Research Center, Dryden Flight Research Facility
- Xu Y, Tai Y-C, Huang A, Ho C-M (2003) IC-integrated flexible shear-stress sensor skin. *J Microelectromech Syst* 12(5):740–747
- Zhao S, Jiang P, Zhu R, Que R (2016) Wearable anemometer for 2D wind detection. *15th IEEE Sensors Conference, Sensors 2016*, 7808599
- Zhu R, Que R, Cao Z (2013) Micro thermal flow sensors/systems on flexible PCB and extensive applications. In: *2013 8th IEEE international conference on nano/micro engineered and molecular systems (NEMS)*. IEEE, pp 179–182



# Passive MEMS DC Electric Current Sensors

Dong F. Wang, Huan Liu, Xuesong Shang, Weikang Xian,  
Yipeng Hou, Xu Yang, Toshihiro Itoh, and Ryutaro Maeda

## Contents

Development Background .....	626
DC/AC Current in Daily Life and Electric Appliances with “Rise-Fall”	
Current Switching .....	626
Cantilever-Based Magnetic/Current Sensor .....	629
Theoretical Considerations of Cantilever-Based DC Current Sensor .....	631
Passive MEMS DC Current Sensor and Its Working Principle .....	631
Actuating MEMS DC Current Sensor and Its Working Principle .....	632
Theory of Maxwell Equation .....	633
Theory of Cantilever Bending Vibration .....	635
Theory of PZT Plates Connect in Series/Parallel .....	637
Structural Design and Prototype Fabrication .....	639
Micro Fabrication of the Prototype MEMS-Scale Piezoelectric Single-Side DC Current	
Sensor Device .....	641
Demonstration of Measurements with the Piezoelectric DC Sensor Device .....	644
Passive DC Sensor Device Experiment .....	644
Examination by Using the Fabricated MEMS-Scale Piezoelectric Single-Side DC	
Current Sensor Device .....	647
Actuating DC Sensor Device Experiment .....	649

---

D. F. Wang (✉)

Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering,  
Jilin University, Changchun, Jilin, China

Research Center for Ubiquitous MEMS and Micro Engineering, AIST, Tsukuba, Japan

e-mail: [dongfwang@jlu.edu.cn](mailto:dongfwang@jlu.edu.cn)

H. Liu · X. Shang · W. Xian · Y. Hou · X. Yang

Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering,  
Jilin University, Changchun, Jilin, China

T. Itoh

Department of Human and Engineered Environmental Studies, University of Tokyo, Chiba, Japan

R. Maeda

Ubiquitous MEMS and Micro Engineering Research Center, AIST Tsukuba, Ibaraki, Japan

Next-Generation Current Sensor Design and Simulation in Structural .....	650
Summary and Perspective .....	652
References .....	654

---

### Abstract

In this chapter, DC current sensors applicable to two-wire appliances are presented. Firstly, background information of current sensors is presented showing the properties of different current measurement applications. Working principles of passive and actuating cantilever-based current sensors are then demonstrated theoretically, and the expressions of magnetic force and piezoelectric outputs are derived. Specifically, the design of partitioning piezoelectric thin film from one plate to  $n$  plates in series is discussed. Following are structural design and prototype fabrication process. With the achieved prototypes, two types of current measurements are carried respectively for verification and demonstration of the working principles. Finally, a conclusion is drawn and perspective of cantilever-based current sensors is given.

---

### Keywords

DC current sensors · Two-wire appliances · DC/AC inducted magnetic fields/forces · Passive/actuating sensing · Amplitude changes · Resonant frequency shifts · Piezoelectric cantilevers · Series/parallel connections · On-off switching · Current monitoring

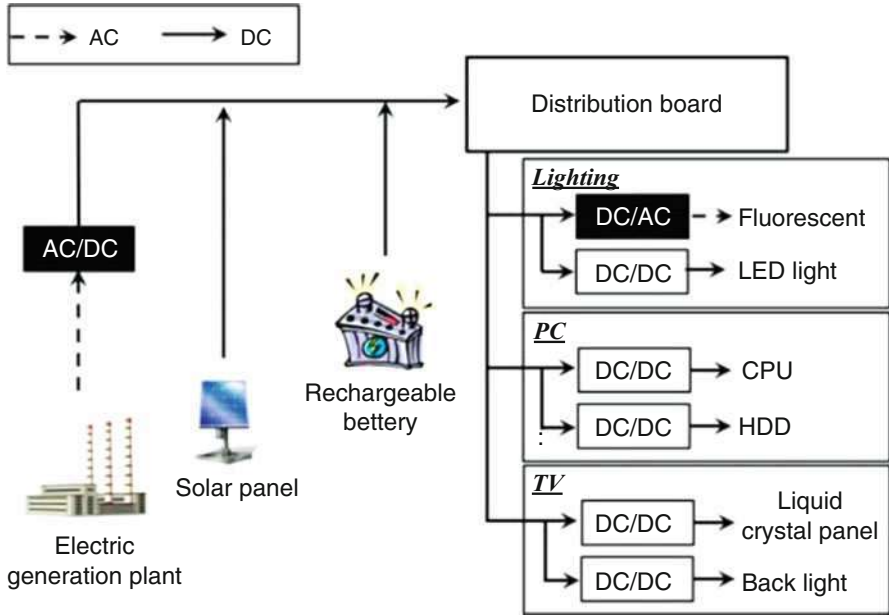
---

## Development Background

### DC/AC Current in Daily Life and Electric Appliances with “Rise-Fall” Current Switching

The energy consumption by factories, automobiles, and even human’s daily life increases  $\text{CO}_2$  exhaust, which subsequently aggravates the greenhouse effect. The total amount of  $\text{CO}_2$  exhaust in the world in 2014 was about 32.3 billion tons, and about one-third was caused by residential section and commercial section. About 40% of the amount of  $\text{CO}_2$  is caused by electrical consumption of household equipment and Information and Communication Technology (ICT) devices. Moreover, the electricity consumption of ICT devices will increase by about 4.2 times by the year of 2025. The electricity consumption of internet data center (IDC) is also rapidly increasing with the augment amount of the data traffic on the internet. It is estimated to grow by two order of its present value by the year of 2025. In addition, IDC has been anticipated to achieve a decrease in AC to DC conversion loss. Something similar is being conducted at DC houses consisting of a solar battery or a storage cell. As shown in Fig. 1, DC/DC appliances are used widely in daily life except for fluorescent lighting, which means it is essential to monitoring the DC electricity consumption so as to establish an effective electricity management system. Therefore, it is essential to monitoring the DC electricity consumption so as to establish an effective electricity management system.





**Fig. 1** Schematic of direct current (DC/DC) appliances widely used in daily life except for fluorescent lighting, where DC/AC is required

Various kinds of current sensor devices (Chen et al. 2012; Lim et al. 2011; Yang et al. 2014, 2015; Leland et al. 2010; Phan et al. 2015; Zheng et al. 2015; Franciscangelis et al. 2013) have been developed for different measurement applications.

**AC Electric Appliances Without “Rise-Fall” Current Switching**

Figure 2 shows the AC current behavior in an electric kettle (Model TFK-17001, TOMIC). The pulse waveform of the electric kettle varies alternatively from 4 A to  $-4$  A at 50 Hz. The waveforms have the constant maximum values and do not change during operation. That is what we mean here without “rise-fall” current switching.

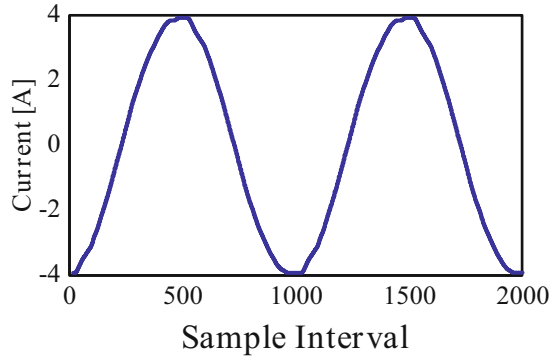
**AC Electric Appliances with “Rise-Fall” Current Switching**

Figure 3 shows the AC current behaviors in an ultrasonic humidifier (Model SC-A28, Deli). The increase and decrease behaviors (the change varies from 0.5 A to 1 A) are mechanically switched for “maximum humidity” of +1 A, “minimal humidity” of +0.5 A in the case of an ultrasonic humidifier, DC electric appliances with “rise-fall” current switching.

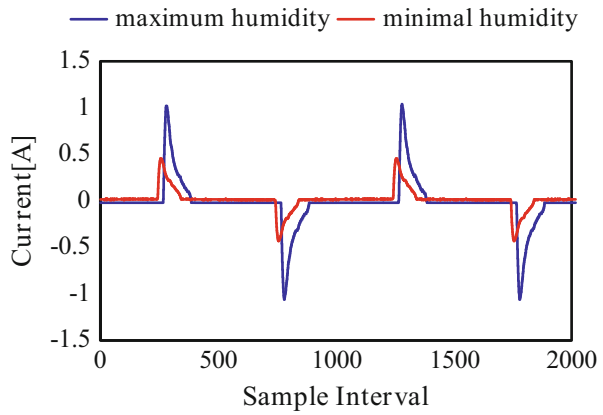
**DC Electric Appliances with “Rise-Fall” Current Switching**

Figure 4 shows the rise and fall behaviors of a DC current flowed in a cell phone with charging “On-Off” switching. In this case, a one Ladder shape change about 0.5 A can be observed from APPLE iPhone 5s in charging, and two Ladder shape change from 0.45 A to 0.8 A can be observed from MEIZU in charging.

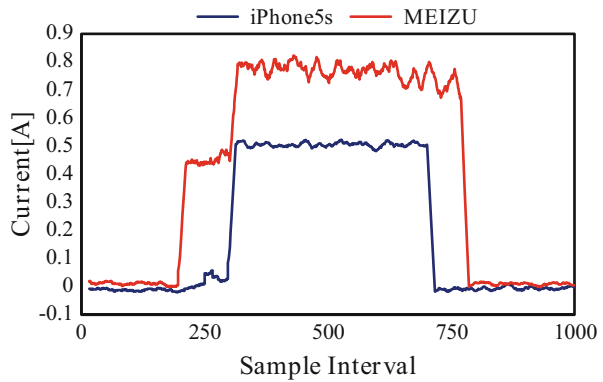
**Fig. 2** The increase and decrease behaviors in consumer electric appliances: electric kettle

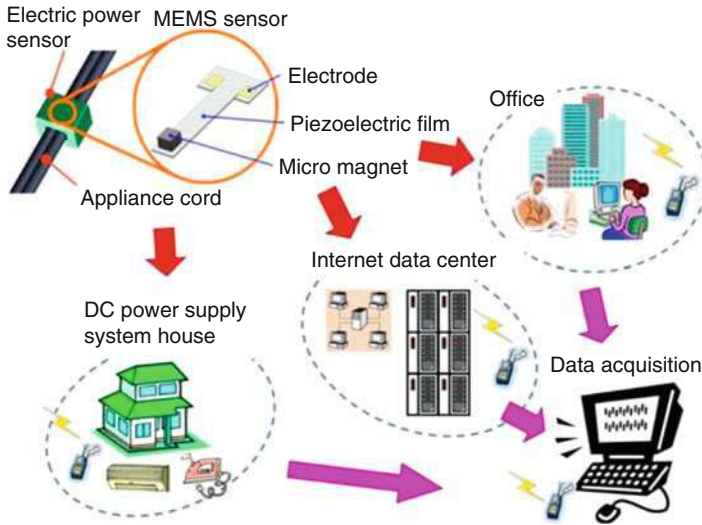


**Fig. 3** Increase and decrease behaviors in consumer electric appliances: an ultrasonic humidifier



**Fig. 4** Increase and decrease behaviors in consumer electric appliances: a cell phone





**Fig. 5** Schematic of the proposed passive and noncontact MEMS DC sensor directly applicable to one-wire or two-wire electric appliances without using cord separator for end-use monitoring of DC power supply

### Cantilever-Based Magnetic/Current Sensor

Over three kinds of popularly used consumer electric appliances were examined for both AC and DC power supply, putting stress on the changing behaviors of the current, like increase and decrease behaviors, or rise and fall behaviors. In this book, a novel MEMS-scale prototype DC sensor, which is self-driven and applicable to both one-wire and two-wire appliance cord, has been microfabricated and preliminarily investigated. The proposed MEMS DC sensor is believed to be very useful to various kinds of DC systems in the near future, as schematically shown in Fig. 5.

Although Hall element based direct current sensor (Popovic 2000; Sophian et al. 2002; Ripka 2004) is the mainstream at the present day, it can be used only for one-wire cord in principle. However, since the household equipment and ICT device use two-wire appliance cord, the Hall element based direct current sensors cannot be applied directly to two-wire cord without a cord separator or a complicated structure with the algorithm (Sun et al. 2016). In addition, a drive voltage is physically necessary for the Hall element based sensors, which is inconvenient to monitor electrical consumption at anytime and anywhere without a power supply.

So the other sensing mechanism needs to be found to measure current directly without a cord separator. Microcantilever is widely applied in measurements of weak physical fields for its high-sensitivity and high-resolution. Especially in the past 20 years, much extraordinary research has been carried out. In 1995, IBM Zurich Research Laboratory studied the correlation between intensity of weak magnetic field and resistance variation of a piezo-resistance cantilever under circumstances of static bending status and vibrating status, and managed to measure intensity of

magnetic field at the magnitude of mT (Rossel et al. 1996); in 1996, group of Johns Hopkins University researched microcantilever with magnetostrictive material on its surface, and realized measuring the intensity of magnetic field at magnitude of  $\mu\text{T}$  by the deflection of cantilever using laser (Osiander et al. 1996); in 1997, the Nanoscale Science Group of University of Cambridge made use of microcantilever with magnetic oxide particles glued on its surface and succeed to measure magnetic field at magnitude of nT by optical deflection of cantilever (Cowburn et al. 1997); in 2002, the Oak Ridge National Laboratory added a cobalt coating on cantilever and utilized fiber-optical to measure the deflection of cantilever, which caused by the current induced magnetic field. By comparing different sizes of cantilever and sensitivities, they could measure AC current with minimum of 0.1 A, at a distance of 5 mm from the sensor to the cords (Goedeke et al. 2004), and in 2009, a group in University of California, Berkeley, proposed the AC current measuring mechanism for two-cords based on microcantilever, and analyzed the characteristics of magnetic field induced by AC current. All the articles on magnetic field and AC current measurements provide wide-research ideas and have great significance to the research about microcantilever-based DC sensors.

In the related works (Isagawa et al. 2011a, b; Wang et al. 2012a, b, 2013a, b, c; Yang et al. 2017), a cantilever-based DC current sensor with or without oscillating pumping was proposed, fabricated, and verified for direct measurements of both the “steady-state” current and the “rise-fall” current switching in one-wire or two-wire electrical appliance cord. In the case with oscillating pumping, it was found that the relative change in the maximum value of the output voltage is approximately proportional to the applied DC current. While in the case without oscillating pumping, it was found that the detected peak value of the impulse output voltage varied linearly with the applied DC electric current ranging from 0.5 to 3 A. The above preliminary studies reveal that the proposed sensing principles could be applied to DC current electricity monitoring and management of in-building and outdoor facilities for intelligent urban infrastructures.

The experimental methodology (Wang et al. 2017a, b) includes the positioning of the sensor, the measurement range for sensitivity characterization, and different connecting modes of PZT plates. While the position of the sensor has been figured out, experiments are conducted in two-dimension of different testing ranges to explore the positional characteristics. Sensitivity characteristics have been studied in DC current from 8 to 400 mA and AC current from 0.5 to 3 A respectively. Besides, different connection modes have also been investigated in the devices of 10 PZT plates. And by measuring the remnant polarization of the devices, the characteristics were studied.

Theoretical considerations for further proposing a self-sensing, noncontact passive piezoelectric MEMS scale DC current sensor, applicable to one-wire or two-wire DC electric appliances, are mainly carried on from the following three aspects, the magnetic force acted on the Neodymium micro-magnet in a magnetic field generated by the DC electric current, the effect of connection mode and thickness of partitioning piezoelectric thin film on the total output voltage, and the cantilever stiffness subject to the fundamental sensing principle in DC current measurements.

These aspects are believed to be essential for both structural design and practical approach of the proposed sensor referring to the preliminary studies briefed above (Wang et al. 2013a, b).

## Theoretical Considerations of Cantilever-Based DC Current Sensor

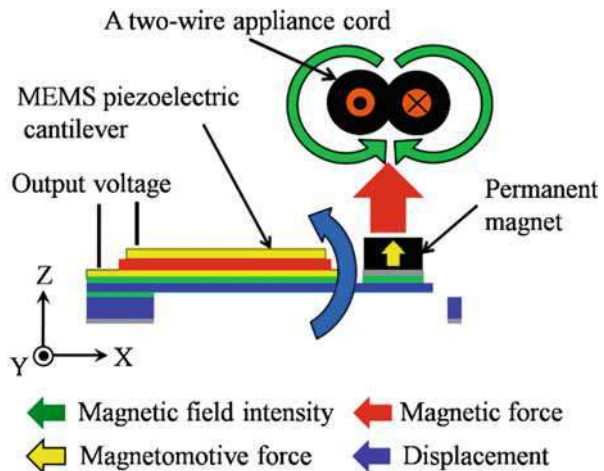
### Passive MEMS DC Current Sensor and Its Working Principle

The measured waveforms of the applied currents generally depend on the impedance of the electric appliances, so one can monitor the electricity consumption by measuring the current. In the case of AC power supply, an alternative current generates an alternative magnitude field, while in the case of DC power supply, a steady current cannot generate a variable magnetic field but a static one, which stays unchanged except for “on-off” switching. However, if the DC electric appliances are operated with the “rise-fall” current switching, the generated magnetic field can be changed correspondingly. This implies that an innovative DC current sensor might be proposed if this changed magnetic field can be picked up (Isagawa et al. 2011a, b).

The sensing principle (Wang et al. 2012a, b) of the proposed sensor is thus shown in Fig. 6. The critical component is a cantilever made up of piezoelectric film and a permanent micromagnet fixed on the end. When a direct current from DC power supply flows via a two-wire appliance cord, the piezoelectric film coated cantilever is bent by the created magnetic force acted on the micromagnet, and the output voltage is generated by the piezoelectric film and the applied DC current is therefore detected.

Based on the sensing principle described above, the output voltage of the proposed piezoelectric MEMS DC current sensor depends on the sudden change or changing behaviors of the DC current, like increase and decrease behaviors,

**Fig. 6** Sensing principle of the proposed passive piezoelectric MEMS DC current sensor



or rise and fall behaviors, including “on-off” switching. Compared with commercially available Hall element based current sensors, the proposed sensor has two significant advantages. One is that the driving voltage is not necessary because of the bending mechanism of the piezoelectric film. The other is that the current flowed in a one-wire or two-wire appliance cord can be measured directly without using cord separator. The proposed passive DC current sensor is therefore expected to be utilized for monitoring the electricity consumption by one-wire or two-wire appliance cord.

## Actuating MEMS DC Current Sensor and Its Working Principle

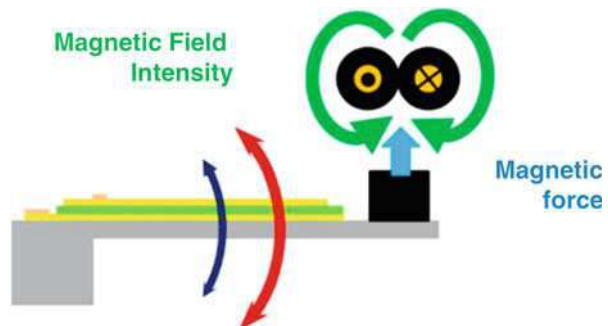
Another method to achieve measurement of current amplitude by piezoelectric-cantilever structure is using the resonant frequency shift of cantilever structure as Fig. 7 shown (Wang et al. 2013a, b). A sine voltage is first applied to the piezoelectric plate for actuating. The cantilever consisting of piezoelectric plates for both actuating and sensing is oscillated due to the actuation. Then, the piezoelectric plate for sensing generates electric charge (output voltage) due to the oscillation in the cantilever. The measured output voltage from the sensing element is influenced by the magnetic field generated from the applied dc current flow via the permanent magnet. As a result, the relative change of resonant frequency due to the applied dc current is determined.

Theoretically, the resonant frequency of a cantilever with a mass (a micromagnet in the present work) is expressed by the Eq. 1:

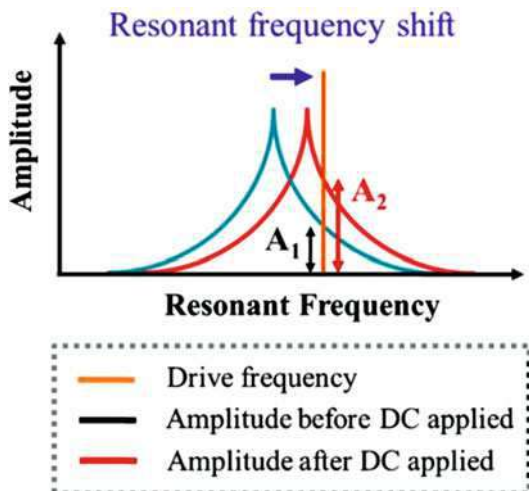
$$f_0 = -\frac{1}{2\pi} \sqrt{\frac{3EI}{mL^3}} \quad (1)$$

In the above equation,  $E$  is Young’s modulus,  $I$  is the moment of inertia,  $m$  is the micro magnet’s mass, and  $L$  is the length of the cantilever. If a magnetic force  $F$  is applied to the micromagnet,  $m$  changes to  $(m - F/g)$ , where  $g$  is the acceleration of gravity. As a result, the apparent value of the resonant frequency of the cantilever is therefore shifted as expressed in the Eq. 2:

**Fig. 7** Sensing principle of the proposed MEMS dc current sensor integrated with piezoelectric actuating and sensing plates. The blue directional marker represents the vibration before applying the dc current, while the red directional marker stands for the vibration after applying the dc current



**Fig. 8** Amplitude change due to the resonant frequency shift over the applied dc current. The black directional marker represents the amplitude before applying the dc current, while the red directional marker stands for the amplitude after applying the dc current



$$f_0 = -\frac{1}{2\pi} \sqrt{\frac{3EI}{(m - F/g)L^3}} \quad (2)$$

The apparent hardness of the cantilever increased when the cantilever was pulled by the magnetic force generated by the applied dc current. As schematically described in Fig. 8, the resonance frequency of the cantilever is measured to slightly shift to close to the driving (excitation) frequency, which is higher than the as-fabricated resonant frequency. As a result, the vibration amplitude, as well as the output voltage, as well as the output voltage was expected to be increased as shown in Fig. 8. Therefore, both the slight resonant frequency shift and the applied higher excitation frequency should account for the vibration amplitude change.

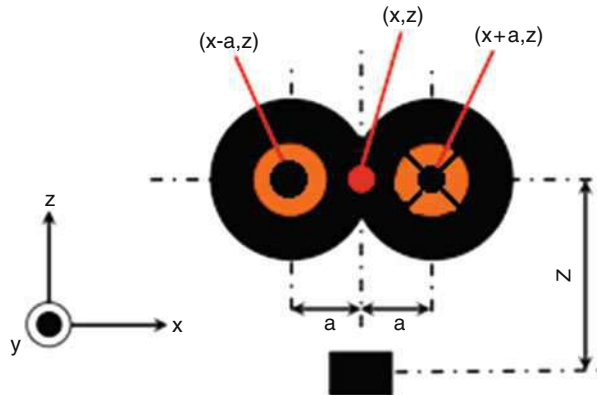
## Theory of Maxwell Equation

When the wires are energized, a magnetic field is generated. And the magnet on the cantilever will be subject to magnetic force. The magnetic force on the permanent magnet is proportional to the integral of the field gradient over the magnet's volume (Yang et al. 2001; Shutov et al. 2005; Leland et al. 2009). When the magnet is placed close to a long current-carrying cord, the magnetic forces in the plane normal to the wire can be described by the following Eqs. 3 and 4:

$$F_y = B_r \iiint \frac{d}{dy} (H_y) dV \quad (3)$$

$$F_z = B_r \iiint \frac{d}{dz} (H_z) dV \quad (4)$$

**Fig. 9** Schematic figure showing the relative position between the two-wire appliance cord and micromagnet in coordinate axis



In Eqs. 3 and 4,  $y$  and  $z$  denote the horizontal and vertical directions, respectively,  $F$  is the magnetic force on the magnet,  $H_y$  and  $H_z$  are horizontal and vertical components of magnetic field in amperes parameter,  $B_r$  is the remanence of permanent magnet in Tesla, and  $V$  is the magnet's volume. The remanence of the permanent magnet is assumed to be uniform and aligned in the positive  $z$ -direction.

In case of a two-wire appliance cord, we define  $a$  as the distance between the center of the appliance cord and the center of right or left one. The  $z$ -direction gradient of the  $z$ -component of the magnetic field surrounding a two-wire appliance cord is given by Eqs. 5 and 6 reported in our studies (Wang et al. 2012a, b, 2013a, b, c), and the magnetic force of the  $z$ -direction in central two-cord can be expressed by the Eq. 7. The two cord dimensional relations as shown Fig. 9. Figure 10 plots the  $y$ -direction gradient of the  $y$ -component of the magnetic field surrounding a two-wire appliance cord with a distance of 1.8 mm.

$$\frac{dH_z}{dz} = \frac{dH_{zl}}{dz} + \frac{dH_{zr}}{dz} \quad (5)$$

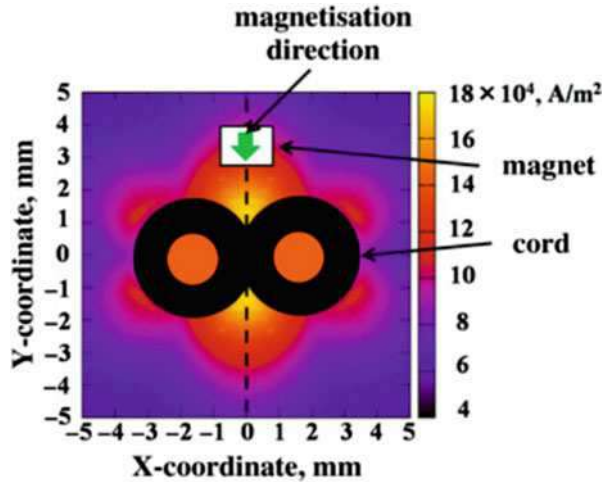
$$\frac{dH_z}{dz} = -\frac{iz}{\pi} \left( \frac{(x-a)}{(x-a)^2 + z^2} - \frac{(x+a)}{((x+a)^2 + z^2)} \right) \quad (6)$$

$$F_z = -\frac{2izaB_r}{\pi(z^2 + a^2)^2} V \quad (7)$$

In Eq. 7,  $F_z$  is the magnetic force of  $z$ -direction on the magnet;  $i$  is the current of wires;  $a$  is the distance between wire center and  $z$ -axis, which project at  $x$ -direction;  $z$  is the distance between wire center and magnet center, which project at  $z$ -axis; and  $V$  is the magnet's volume.



**Fig. 10** Plot of the magnitude of the vertical component of y-direction magnetic field gradient around a two-wire appliance cord, 10 A current assumed



### Theory of Cantilever Bending Vibration

The z-axis, y-axis, and x-axis were defined and the origin of the z-axis as the axis vertical to the surface (thickness direction), the width direction, the length direction, and the bottom of the cantilever in Fig. 11, respectively. Also, the width of individual PZT plate  $w_E$  is expressed by the following Eq. 8:

$$w_E = \frac{w - (n - 1)\varphi}{n} \tag{8}$$

where  $n$  is the number of PZT plate,  $w$  is the whole width of PZT, and  $\varphi$  is the gaps between the adjacent PZT plate after partition. The accumulated charge in individual PZT plate is expressed as follows from Gauss’s law:

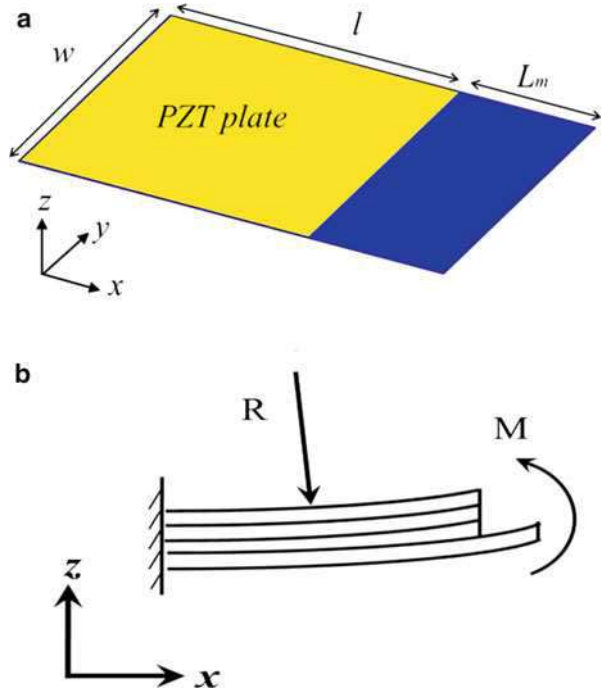
$$Q_{ind} = \iint_{l, w_E} D dw dl \tag{9}$$

where  $D$  is the electrical-field displacement in z-direction and  $dw dl$  is the area of individual PZT. For piezoelectric sensing, the electrical-field displacement  $D$  without applying any external electrical field is described by Eq. 10:

$$D = d_{31}\sigma_p \tag{10}$$

where  $d_{31}$  and  $\sigma_p$  are transverse piezoelectric constant ( $-50 \text{ pmV}^{-1}$ ) and x-direction stress when applied bending moment on the tip of cantilever, respectively. Also, x-direction stress  $\sigma$  is expressed as follows Eq. 11 (Zou et al. 2008),

**Fig. 11 (a)** Schematic figure showing a cantilever-based piezoelectric sensor, where the yellow region is single PZT plate and the blue region is Si substrate,  $w$  is the width of PZT plate,  $l$  is the length of PZT plate, and  $l + L_m$  is the length of cantilever. **(b)** Schematic figure showing the  $z$ - $x$  cross-section of the cantilever-based piezoelectric sensor, the each layer along the  $z$ -axis direction expressed by top electrode layer, piezoelectric layer, bottom electrode layer, and Si substrate layer. Where  $M$  is the torques created by the created magnetic force acted on the micromagnet and  $R$  is the curvature radius of the bending cantilever



$$\sigma = \frac{E_p Z_p \left( \frac{1}{2} L_m + l \right) F_z}{\sum_i E_i (I_i + A_i Z_i^2)} \tag{11}$$

where  $F_z$  is the magnetic force when counterpoises restoring force of cantilever.  $E_i$  and  $E_p$  are Young's moduli of each layer and PZT thin film, and  $I_i$  is the moment of inertia of each layer, respectively.  $A_i$  is the  $z$ - $y$  cross-section area which described by Eq. 12:

$$A_i = \{w - (n - 1)\varphi\}h_i \tag{12}$$

where  $h_i$  is the thickness of each layer, and  $\varphi = 0$  in case of  $z$ - $y$  cross-section of the substrate. In Eq. 11,  $Z_p$  is the distance between the position  $z_p$  of center of the PZT thin films and the position  $z_N$  of neutral axis parallel to the length direction of cantilever, and can be expressed as Eq. 13.  $Z_i$  is the distance between the position  $z_i$  of center of the each layer and the neutral axis parallel to the length direction of cantilever and expressed as Eq. 14.

$$Z_p = z_p - z_N \tag{13}$$

$$Z_i = z_i - z_N \tag{14}$$

Where the neutral axis  $Z_N$  is expressed as follows Eq. 15 (Kobayashi et al. 2010):

$$z_N = \frac{\sum_i z_i E_i A_i}{\sum_i E_i A_i} \quad (15)$$

The accumulated charge of individual PZT plate can then be expressed as follows by Eqs. 9, 10, and 11.

$$Q_{\text{ind}} = \frac{d_{31} E_p Z_p \left( \frac{1}{2} L_m + l \right) l w_E}{\sum_i E_i (I_i + A_i Z_i^2)} F_z \quad (16)$$

### Theory of PZT Plates Connect in Series/Parallel

As shown in Fig. 12a, when n pieces of PZT plates are electrically connected in series, the total accumulated charge  $Q_{\text{total}}$  is equivalent to the accumulated charge  $Q_{\text{ind}}$  in individual PZT plate.

$$Q_{\text{total}} = Q_{\text{ind}} \quad (17)$$

Therefore, the wider the width of the PZT plate, the more increment the charge of the device.

The output voltage connected in series can be expressed by Eq. 18:

$$V_{\text{total}} = \frac{d_{31} E_p Z_p \left( \frac{1}{2} L_m + l \right) l w_E}{\sum_i E_i (I_i + A_i Z_i^2)} \frac{1}{C_{\text{total}} + C_{\text{others}}} F_z \quad (18)$$

where  $C_{\text{others}}$  is the capacitance of the measurement system.  $C_{\text{total}}$  is the total capacitance of the device and is expressed by Eq. 19:

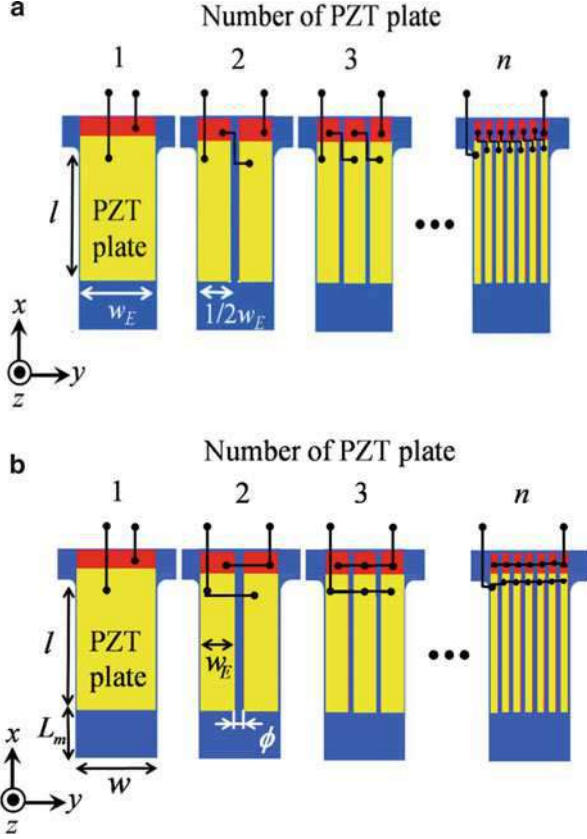
$$C_{\text{total}} = \frac{C_{\text{ind}}}{n} \quad (19)$$

The capacitance of individual PZT plate  $C_{\text{ind}}$  is given by

$$C_{\text{ind}} = \varepsilon \varepsilon_0 \frac{S_E}{h_p} \quad (20)$$

where  $\varepsilon$  and  $\varepsilon_0$  are the vacuum dielectric constant ( $8.85 \times 10^{-12} \text{ Fm}^{-1}$ ) and relative dielectric constant (1000) of PZT thin films, respectively.  $S_E$  is the area of the

**Fig. 12** (a) Schematic figure for partitioning piezoelectric thin film from 1 plate to n plates in series. (b) Schematic figure for partitioning piezoelectric thin film from 1 plate to n plates in parallel



electrode and  $h_p$  is the thickness of the PZT thin films. Therefore, in order to obtain a high  $V_{total}$  in Eq. 18, it will be effective to decrease the capacitance  $C_{total}$  and increase the number of PZT plates at the same time.

As shown in Fig. 12b, when n pieces of PZT plates are electrically connected in parallel, the total output charge  $Q_{total}$  is n times more than the output charge of individual PZT plate and expressed as Eq. 21.

$$Q_{total} = nQ_{ind} \tag{21}$$

The output voltage connected in parallel can be expressed by Eq. 22.

$$V_{total} = \frac{nd_{31}E_pZ_p\left(\frac{1}{2}L_m + l\right)lw_E}{\sum_i E_i(I_i + A_iZ_i^2)} \frac{1}{C_{total} + C_{others}} F_z \tag{22}$$

The total electric capacity  $C_{total}$  can be expressed as following Eq. 23.

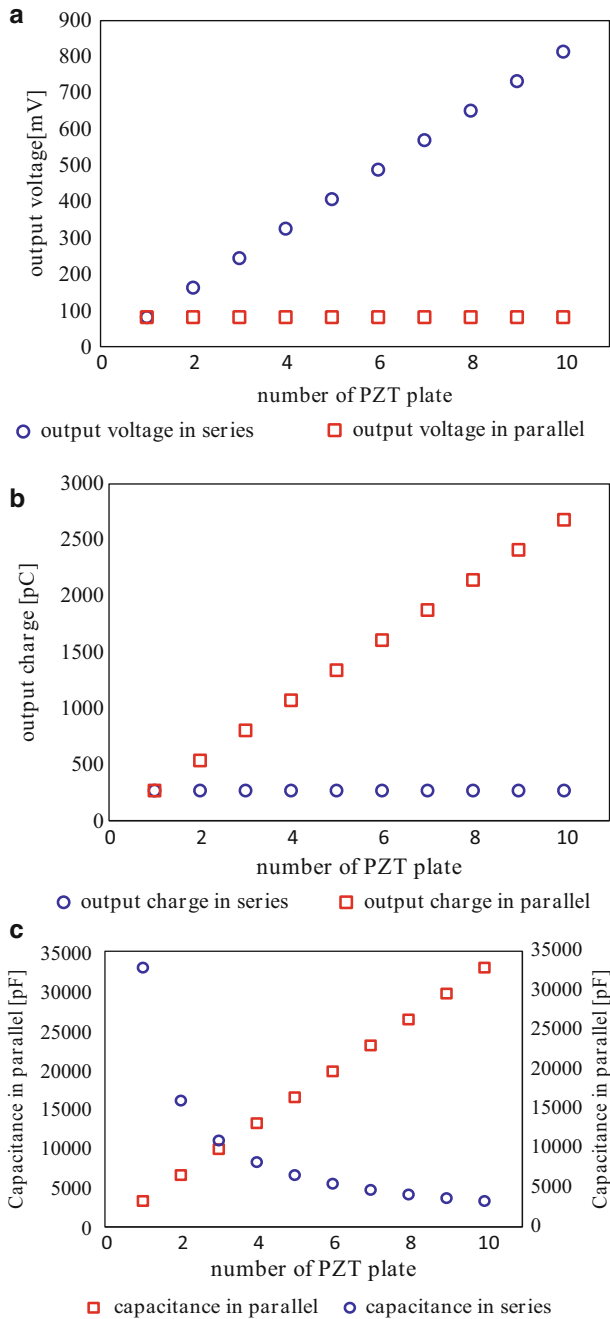
$$C_{\text{total}} = nC_{\text{ind}} \quad (23)$$

While the capacitances induced by others are neglected. Then the output voltage, charge, and capacitance of series and parallel can be compared with the connect number of PZT plate, as shown in Fig. 13a, b, c (Wang et al. 2017).

## Structural Design and Prototype Fabrication

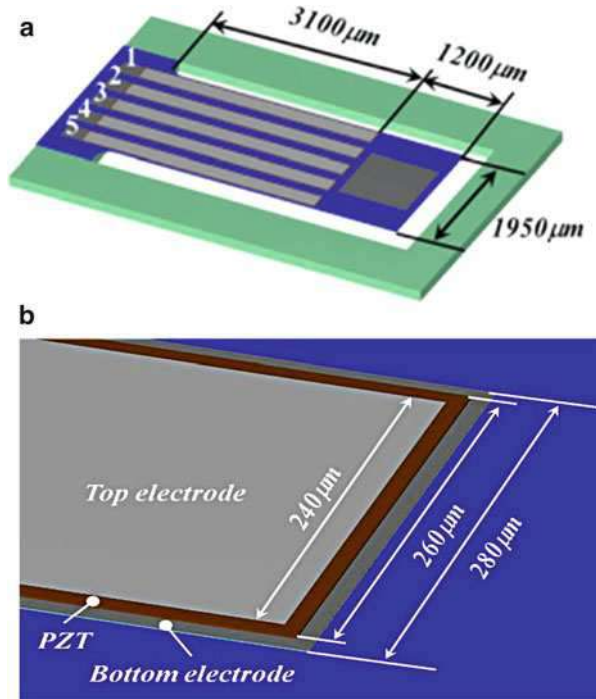
The substrate and connect them in series with each other. Figure 14a (Wang et al. 2013a) shows a typical design of our MEMS DC current sensor consisting of 5-PZT plates indicating such kind of design approach, in light of the theoretical considerations discussed in the previous section. A Neodymium magnet with a size of 1 mm × 1 mm × 1 mm is designed to be located at the cantilever tip. The distance between the center of the Neodymium magnet and the center of the two-wire appliance cord is 3.8 mm. The remanence magnetization  $B_r$  is assumed as 1.2 T. The size of cantilever is chosen as 4300 μm in length and 1950 μm in width to withstand the magnetic force when a current of 10 A is applied, and the center of magnet is located at 3.8 mm from the center of the appliance cord. The square space for the lower electrode of Pt/Ti is the alignment mark for positioning the permanent magnet. Taking account of an undercut about 10 μm, the width of the upper electrode, PZT thin film, and the lower electrode are designed to be 240 μm, 260 μm, and 280 μm respectively, as described in Fig. 14b. It is noted from Fig. 15a, b (Wang et al. 2017) that, the thicker the PZT plate, the much decrease the values of  $Q_{\text{ind}}$  and  $C_{\text{ind}}$ . However, since the decrease of  $C_{\text{ind}}$  is greater than that of  $Q_{\text{ind}}$ , an increasing behavior and a maximum point at thickness of 4 μm can be observed for  $V_{\text{ind}}$  from Fig. 15c (Wang et al. 2017). Considering the difficulty in device processing, however, a thickness of around 2 μm is thus selected for the present study. Electrical connects in series by wire bonding for signal amplification will be operated. The square space for the lower electrode of Pt/Ti is the alignment mark for positioning the permanent magnet. Applying a resultant force rising from the magnet mass and the maximum magnetic force when 10 A DC current is flowed, the internal stress of each layer when varying the silicon thickness ranging from 1 μm ~ 10 μm can thus be calculated using Eq. 11 and summarized in Fig. 16 for comparison. It is noted that higher internal stress can be found in top electrode Pt/Ti layer if fixing the thickness values for top electrode, PZT, bottom electrode, thermal SiO<sub>2</sub>, and substrate silicon at 200 nm, 2 nm, 200 nm, and 1 nm respectively. Since the yield stress of Pt is about 70 MPa, a reasonable thickness range for substrate silicon should be around 4 nm. The material parameters as shown in Table 1.

In general, the amount of electric charge generated from PZT thin film increases as its thickness decreases. When the thickness decreases to a certain value, dielectric breakdown will occur before a required displacement is achieved. On the other hand, if the thickness goes to be thicker than the elastic layer, the neutral axis will exist in PZT thin film and cause to cancel the polarization matching. It is thus necessary to



**Fig. 13** (a) The output voltage in series and parallel respectively with a number of PZT plate. With the number of PZT plate increasing, the output voltage has a linear increase in series, but in parallel the output voltage keeps a constant as same as the one plate condition. (b) The output charge in

**Fig. 14** (a) Structurally designed self-sensing, noncontact MEMS DC current sensor device with 5-PZT partition plates with geometrical dimensions. (b) Structurally designed single PZT plate with geometrical dimensions



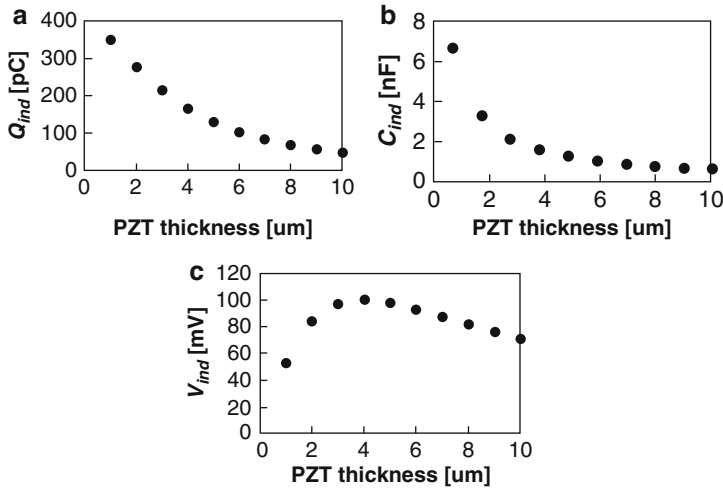
keep a thickness balance, usually from 1 μm ~ 10 μm, between PZT thin film and the elastic layer of silicon (Wang et al. 2017).

### Micro Fabrication of the Prototype MEMS-Scale Piezoelectric Single-Side DC Current Sensor Device

As shown in Fig. 17 (Isagawa et al. 2011a, b; Wang et al. 2013c), the prototype devices are fabricated from multilayer of Pt/Ti/PZT/Pt/Ti/SiO<sub>2</sub> deposited on silicon-on-insulator (SOI) wafers using a 5-mask micromachining process.

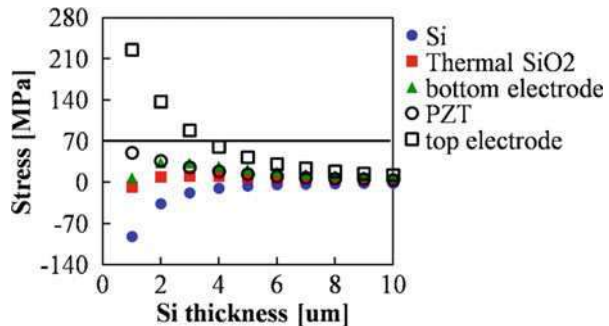
Deposition of the multilayer is started from thermal oxidation of SOI wafers followed by Pt/Ti bottom electrodes deposition. After (100)-oriented PZT thin film is deposited by a sol-gel process, Pt/Ti top electrodes are then sputtered.

←  
**Fig. 13** (continued) series and parallel with a number of PZT plate. With the number of PZT plate increasing, the output charge keeps a constant as same as the one plate condition in series, but in parallel the output charge has a linear increase. (c) The capacitance of series and parallel with a number of PZT plate. With the number of PZT plate increasing, in parallel, the output charge has a linear increase, contrary, the output charge as inverse proportional function of the number of PZT plate



**Fig. 15** (a) Output charge as a function of the thickness of piezoelectric thin film (single plate). (b) Electrostatic capacity as a function of the thickness of piezoelectric thin film (single plate). (c) Output voltage as a function of the thickness of PZT film

**Fig. 16** Internal stress on different layers as a function of Si thickness



The etching process is as follows. Pt/Ti top electrodes are first etched by an Ar ion through mask No. 1. PZT thin films are then wet-etched with an aqueous solution of HF, HNO<sub>3</sub>, and HCl through mask No. 2. After Pt/Ti bottom electrodes are etched by an Ar ion through mask No. 3, both the thermal SiO<sub>2</sub>, structural Si and buried oxide (BOX) are etched by RIE with CHF<sub>3</sub> gas (SiO<sub>2</sub>) and SF<sub>6</sub> gas (Si) through (mask No. 4). Finally, the substrate Si is etched from the backside to release the cantilever through mask No. 5.

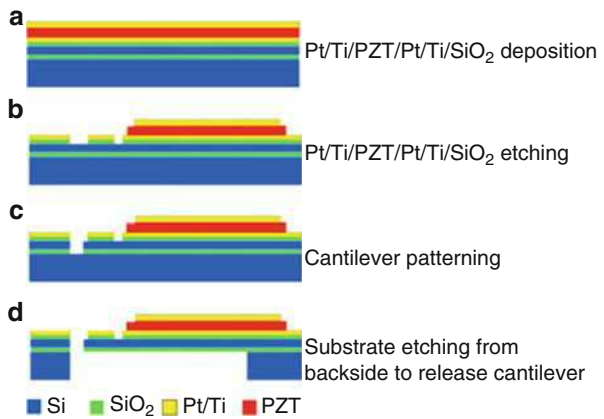
The P-E curves in Fig. 18 are found to be approximately overlapped after each etching step, which means that the etching process has a little effect on the polarization performance of PZT plates. We thus randomly select three PZT plates to test their performance in the same environment.



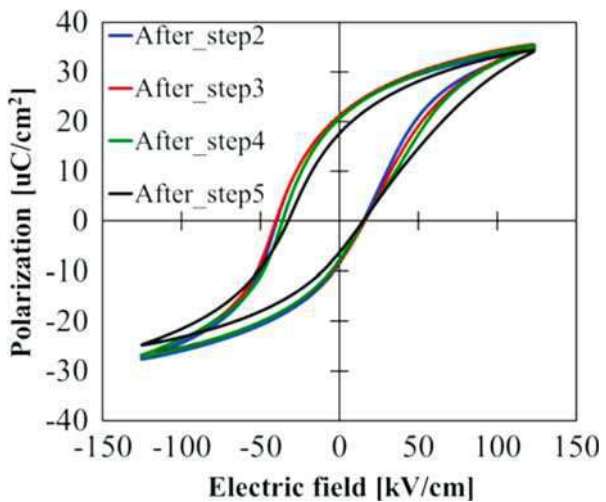
**Table 1** Comparison of the thickness, Young’s modulus, and the distance from the bottom surface of the cantilever to the center of different layers

Material	$i$	$h_i(\mu\text{m})$	$E_i(\text{N/m})$	$Z_i(\mu\text{m})$
Top electrode Pt/Ti	5	0.2	168	7.3
PZT	4	2	72.5	6.2
Bottom electrode Pt/Ti	3	0.2	168	5.1
Thermal SiO <sub>2</sub>	2	1	73.1	4.5
Si	1	4	190	2

**Fig. 17** Schematic diagram of the fabrication process chart: (a) Pt/Ti/PZT/Pt/Ti/SiO<sub>2</sub> deposition, (b) Pt/Ti/PZT/Pt/Ti/SiO<sub>2</sub> etching, (c) cantilever patterning, and (d) substrate etching from the backside to release cantilever



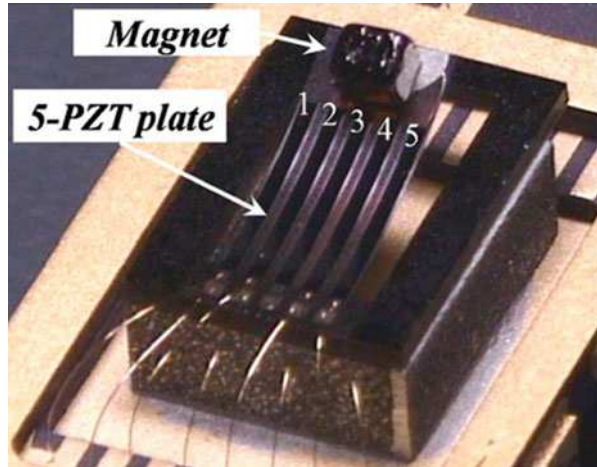
**Fig. 18** P-E curves of the PZT plate after each etching step in the electric field varying from -150 kV/cm to +150 kV/cm



**Table 2** Comparison of polarization values of three PZT plates

Plate numbers	+Pr [ $\mu\text{C}/\text{cm}^2$ ]	-Pr [ $\mu\text{C}/\text{cm}^2$ ]	$\{\text{Pr} - (-\text{Pr})\}/2$ [ $\mu\text{C}/\text{cm}^2$ ]
PZT plate no. 1	8.39	-1.15	4.77
PZT plate no. 1	9.47	-1.19	5.33
PZT plate no. 1	9.36	-1.28	5.32

**Fig. 19** A magnified optical image of the micro-fabricated MEMS piezoelectric single-side DC current sensor with 5-PZT plates and a permanent micro-magnet fixed on the end of the cantilever



As summarized in Table 2, the comparison of polarization values of three PZT plates shows that a little difference is observed after the same etching process. The maximum difference in Table 2 is  $0.56 \mu\text{C}/\text{cm}^2$ , which is very small.

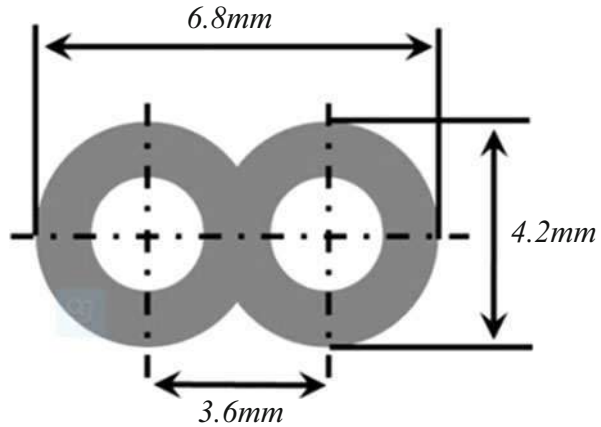
Figure 19 shows the optical picture of the fabricated MEMS piezoelectric single-side DC current sensor device. It can be noted that the cantilever is warped due to the BOX layer remained on the backside.

## Demonstration of Measurements with the Piezoelectric DC Sensor Device

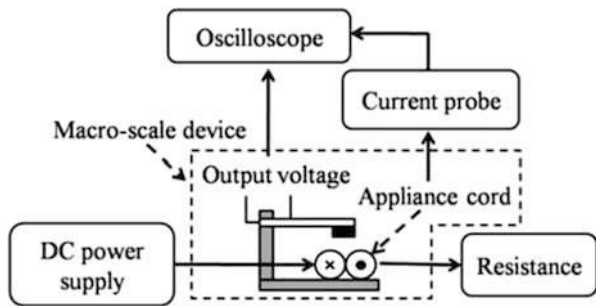
### Passive DC Sensor Device Experiment

As explained in the previous section, when a direct current is applied to the appliance cord, the output signal from the proposed cantilever-based piezoelectric DC sensor is supposed to be very impulsive. Therefore, a macroscale device has been fabricated to demonstrate whether the output signal arising from direct current can be measured or not. The macroscale device consisted of a bimorph cantilever with a permanent magnet located at the cantilever tip. The size of the bimorph and the permanent

**Fig. 20** Cross-section of a typical two-wire appliance cord, which is examined in present study



**Fig. 21** Measurement setup for demonstration by using the fabricated device and the typical two-wire appliance cord

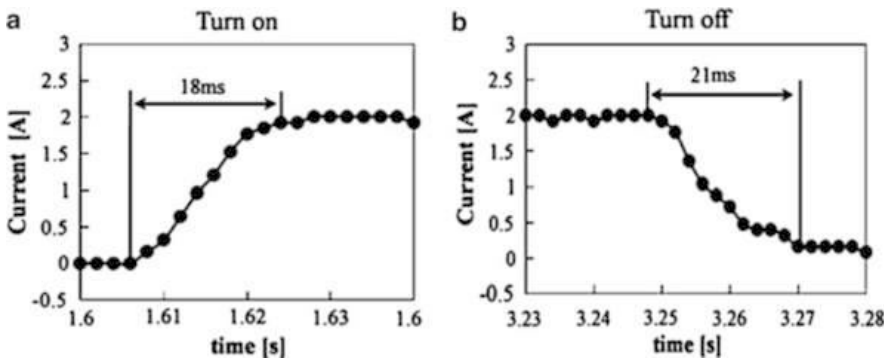
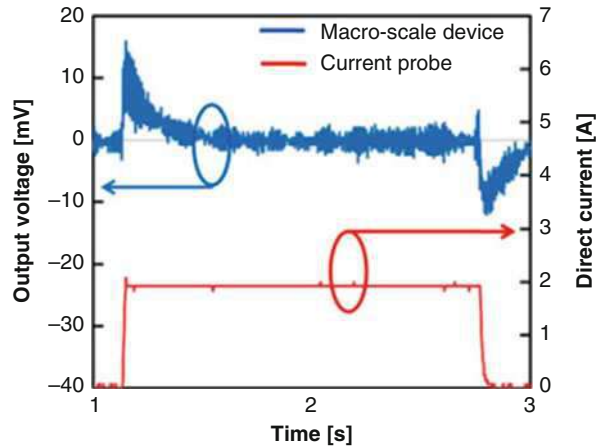


magnet is  $28 \times 13.4 \times 9$  (mm<sup>3</sup>) and  $5 \times 5 \times 2$  (mm<sup>3</sup>), respectively. The remanence magnetization  $B_r$  is assumed as 1.2 T (Wang et al. 2012b).

As the intended application of this research is to monitor residential and commercial electricity use, a two-wire “zip-cord,” common to many appliances as shown in Fig. 20, is selected to be examined in the present study. Figure 21 shows a simple measurement set-up for preliminary demonstration by the macroscale device. The output voltage, arising from the direct current supplied by a DC power supply (AND Co. AD-8735A, JPN), is measured by an oscilloscope (Tektronix Co. TDS2014, USA). The applied direct current is measured by a current probe (Tektronix Co. TCP312, USA). The output voltage and the applied direct current were measured with a 2 ms sampling interval. However, the center of the permanent magnet is located at 25.5 mm from the base of cantilever and 3.6 mm from the center of the appliance cord.

As a result, the output voltage impulsive signals measured by the macroscale device. Figure 22 typically shows that the output voltage impulse signal was clearly detected when 2 A was applied to a two-wire appliance cord. Figure 23a, b further show the detailed of the “rise time” and “fall time” of the applied DC current

**Fig. 22** A typical measurement showing the output voltage impulses when a direct current of 2 A was applied to a two-wire appliance cord (on), and when the applied current was discontinued (off), respectively. The output voltage from the current probe was also drawn for comparison

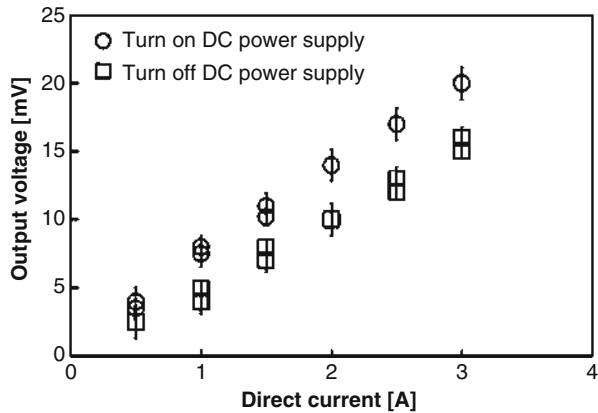


**Fig. 23** DC current in a two-wire appliance cord: (a) rise time after “on” switching; (b) fall time after “off” switching

as 18 and 21 ms, respectively. The peak value was measured as  $-10$  mV when turning on the DC power supply, while that was measured as 17 mV when turning off the DC power supply. The output voltage impulse signal converged to zero within 0.07 s. Such measurement can be conducted from a lower applied current of 0.5 A to a higher one of 3 A. However, the reason for the difference in the amplitude of the measured output values depending on on-off switching is not clear, since the symmetry architecture was used in present study, rather than a construction with a piezoelectric layer only on the top or on the back. In fact, the mass effect of the magnet should be further considered, and this might account for the reason that the output voltages were not equal.

Figure 24 shows a linear relation between the absolute peak values (in error bar) of output voltage impulse signal and the applied DC current. The sensitivities of the macroscale device were derived as around 7 and 5 mV/A for the cases of turning on and turning off the DC power supply, respectively.

**Fig. 24** Empirical peak values (in error bar) of the output voltage impulse as a function of the applied direct current, as measured in Fig. 10



### Examination by Using the Fabricated MEMS-Scale Piezoelectric Single-Side DC Current Sensor Device

A simple experimental set-up was constructed for the examination by using both macroscale device and MEMS-scale device. Two-wire appliance cord was placed under the micromagnet, which was fixed by manipulator onto the PZT cantilever tip to the theoretically pinpointed position of PZT plate No. 3, as shown in Fig. 19 (Wang et al. 2013b). The impulse output voltage of the piezoelectric cantilever was then measured when a direct current was flowed from the DC power supply. The distance from the center of two-wire cord to the center of micromagnet was adjusted from 3.8 mm to 2.8 mm. The output voltage, arising from the direct current supplied by a DC power supply (AND Co. AD-8735A, JPN), is measured by an oscilloscope (Tektronix Co. TDS2014, USA). The applied direct current is measured by a current probe (Tektronix Co. TCP312, USA).

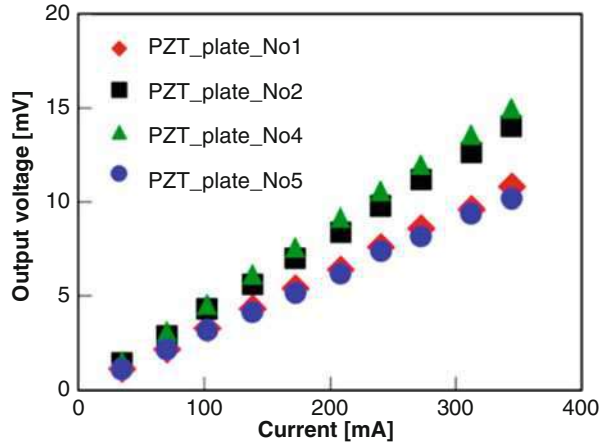
The output voltage and the applied direct current were measured with a 2 ms sampling interval.

The output voltage as a function of the alternating current was plotted for each single PZT plate. As shown in Fig. 25, slightly higher output voltage was obtained from single PZT plate No. 2 or No. 4, when an AC current of 30 Hz was applied. However, no measurement was carried out for the center PZT plate No. 3, since the tip of which might be affected by the micromagnet during the integrating process.

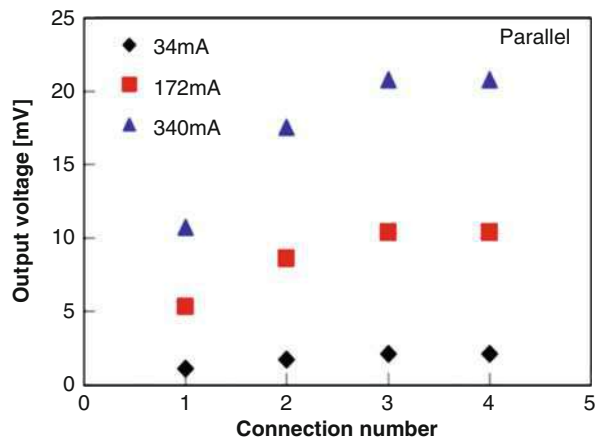
The change of the total output voltage as a function of the connection number of single PZT plates in parallel, varying from one plate to four plates, was plotted with the applied AC current of 34, 172, and 340 mA, respectively. As shown in Fig. 26, the total output voltage was first increased with the increasing number of parallel plates to some degree, and then stayed almost unchanged from three plates in the study.

Although an increase in the voltage sensitivity was experimentally observed, the mechanism of the nonlinear responses for the proposed structure due to the applied

**Fig. 25** Measurement results showing the output voltage as a function of AC current of 30 Hz with a respect to different single PZT plate corresponding to Fig. 6



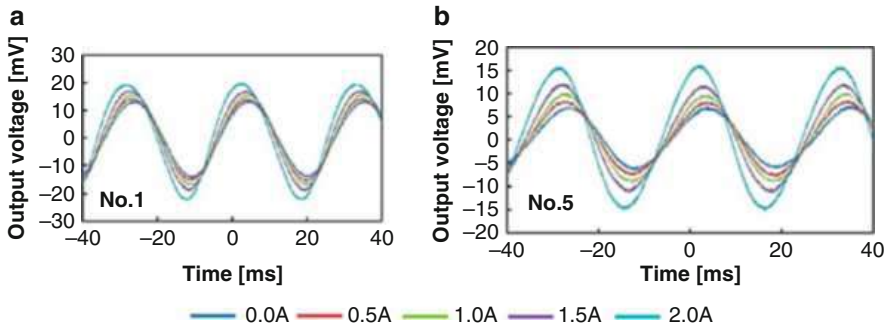
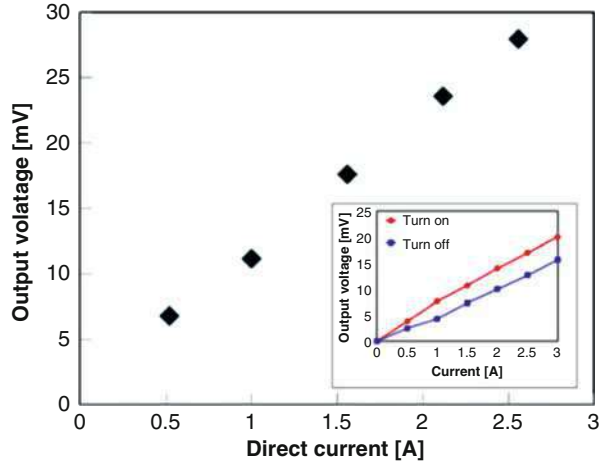
**Fig. 26** Measurement results showing the output voltage as a function of the parallel connection number of PZT plates with respect to different applied currents of 34, 172, and 340 mA, respectively



current shown in Fig. 26 should be further clarified with more experimental and theoretical evidence.

As a result, the output voltage the impulsive values measured by using the fabricated MEMS-scale piezoelectric single-side DC current sensor device. Different from the results of the macroscale piezoelectric bimorph DC current sensor, impulsive piezoelectric output voltages accompanying with a gradual decrease in decay were detected when the applied DC electric current was varied from 0.5 to 2.5 A. As shown in Fig. 27, a linear relationship between the detected peak value of the impulse output voltage and the applied DC electric current was also obtained but with a higher slope compared to the result of the macroscale piezoelectric bimorph DC current sensor.

**Fig. 27** Linear relationship between the peak value of impulse voltage and the applied DC electric current, compared with the data obtained from the macro-scale piezoelectric bimorph DC current sensor (inset figure)

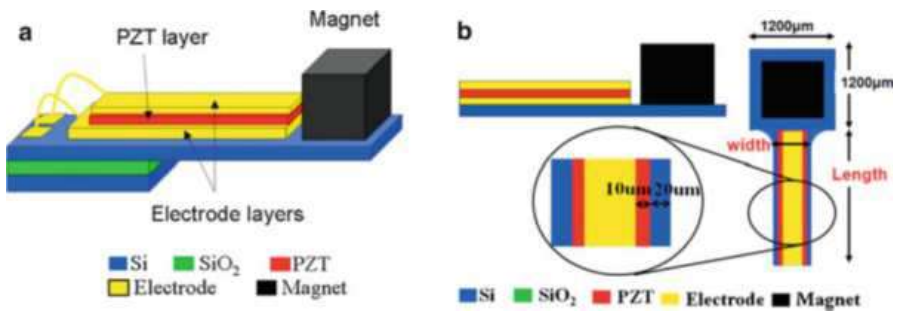
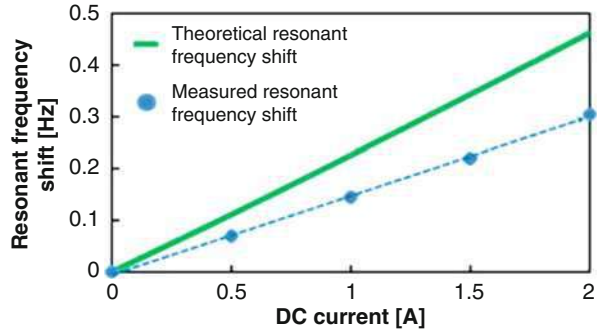


**Fig. 28** Measured output waveforms when an excitation voltage of 6 V<sub>pp</sub> with a frequency of 32.5 Hz was applied to the actuating piezoelectric plates

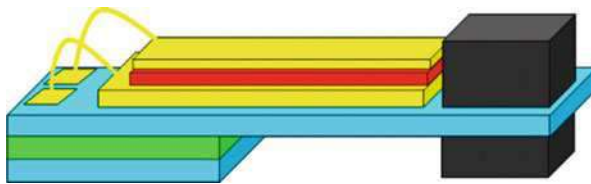
### Actuating DC Sensor Device Experiment

An experiment (Wang et al. 2013a, b) is conducted for verification of the resonant frequency shift. Figure 28 shows the measured output waveforms when an excitation voltage of 6 V<sub>pp</sub> with an excitation frequency of 32.5 Hz was applied to the actuating plates no. 2 and no. 4. Figure 28a was measured from sensing plate no. 1, while Fig. 28b was from plate no. 5. The relative change of the maximum value of the signal output voltage was obtained before and after applying DC current. DC current measurements were conducted continuously from an applied current of 0 to 2 A. In Fig. 29, the resonant frequency shift is demonstrated and increased linearly over the applied dc current.

**Fig. 29** Resonant frequency shift of the integrated MEMS dc current sensor was observed to be increased over the applied dc current both theoretically and experimentally. The green line is the theoretical resonant frequency shift, while the blue spots are for the experimental one



**Fig. 30** Scheme of the proposed cantilever resonance type magnet-integrated sensor device, named as design 1, where (a) permanent micromagnet is set up on the top point of the Si substrate; (b) geometrical details

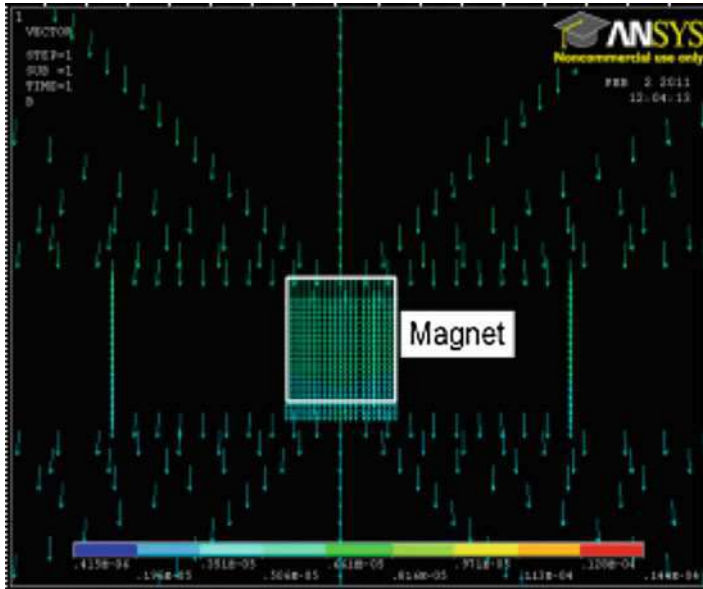


**Fig. 31** Scheme of the proposed cantilever resonance type magnet-integrated sensor device, named as design 2, where a permanent micromagnet is separated into two parts

## Next-Generation Current Sensor Design and Simulation in Structural

Two structural designs, named as design 1 and design 2, have been comparatively analyzed using ANSYS in order to obtain larger frequency shifts (higher magnetism sensitivity) due to the applied exterior magnetic field (Wang and Maeda 2015). The analytical results show that, in the range of 0–10 mT, the frequency shifts are small,





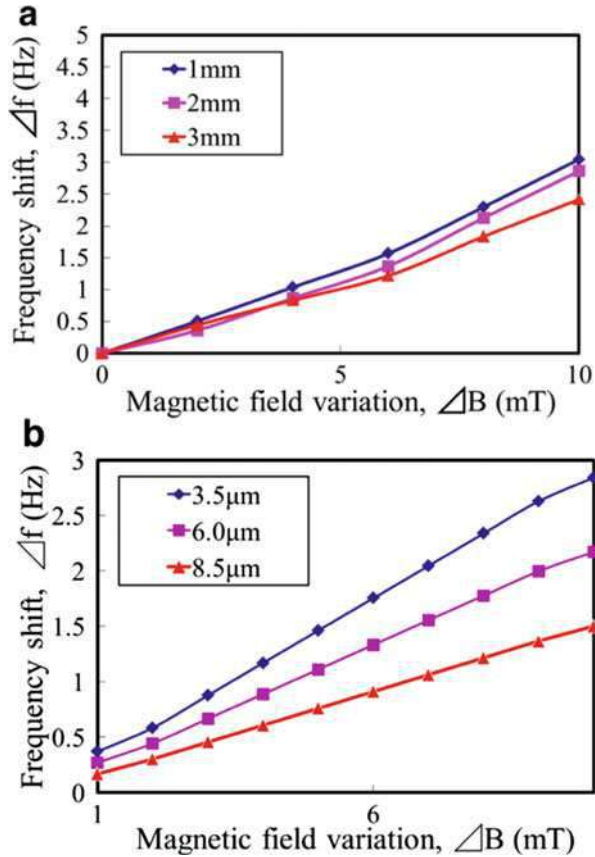
**Fig. 32** Analytical result of flux density vector for flux density at 0.005 mT, with an introduced large magnet in the center. Similar behaviors can also be confirmed for the other flux densities at 0.05, 0.5, 10, and 100 mT

while under 100 mT, a relatively larger frequency shift of about 30 Hz can be theoretically obtained.

The proposed cantilever resonance type magnet-integrated sensor devices are depicted in Figs. 30 and 31 separately. Typically, there are three layers on the top of Si substrate, and the piezoelectric thin film is placed between two electrode layers (Pt thin films). In order to detect or reflect the external magnetizing field, a permanent micromagnet is set up on the top point of the Si substrate (design 1) as shown in Fig. 30, and is separated into two parts (design 2) as shown in Fig. 31. The thickness of one part in design 2 is half of that in design 1. In analysis, the size of the above permanent micromagnet is designed as  $1 \text{ mm}^3$  ( $1 \times 1 \times 1 \text{ mm}$ ), and the material is neodymium. A large magnet is introduced and is supposed to generate an external homogeneous magnetizing field, strong enough than that of the permanent micromagnet. Figure 32 shows an analytical result of the flux density vector with the center large magnet. As depicted in the figure, it is clear that a flux density of 0.005 mT from the external magnetizing field has expanded straight from up to down with respect to the center large magnet. Similar behaviors can also be confirmed for the other flux densities at 0.05, 0.5, 10, and 100 mT. The analysis of the suction force from the introduced large magnet is therefore believed to be reasonable.

As shown in Fig. 33a, while the cantilever displacement increased with the increase of the cantilever length. The effect of the cantilever thickness on the

**Fig. 33** Frequency shift as a function of magnetic field variation, (a) with respect to different cantilever length; (b) with respect to different cantilever thickness

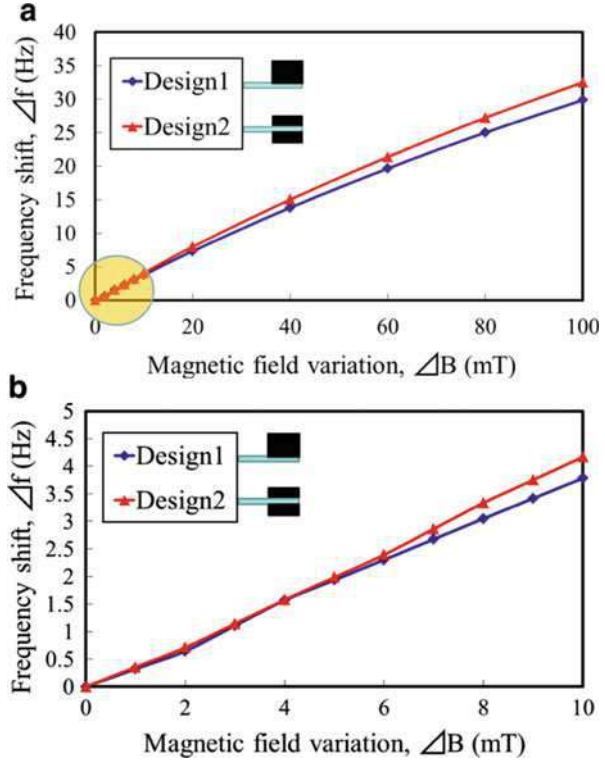


resonance frequency shift was further analyzed. It is noted from Fig. 33b that the frequency shift decreased with the increase of the cantilever thickness. As shown in Fig. 34, a frequency shift of about 29 Hz can be obtained for design 1 under 100 mT, while about 31 Hz can be obtained for design 2 under the same external field, slightly (about 10%) higher than that in design 1.

## Summary and Perspective

In this chapter, current sensor applicable to the two-wire appliance is presented. There is urgent need of development for current sensing because of rapidly increasing electricity consumption to which monitoring is of great importance. Different current measurement applications show characteristic properties which are shown in detailed. Working principles of passive and actuating cantilever-based current sensor are then demonstrated respectively. The former use amplitude changes of piezoelectric

**Fig. 34** Frequency shift as a function of the magnetic field variation, where (b) is an enlarged one for the range of 0–10 mT in (a)



output to derived current amplitude, while the latter use resonate frequency shifts. Expressions of magnetic force and piezoelectric outputs are developed. Specifically, the design of partitioning piezoelectric thin film is discussed for amplification of piezoelectric outputs. Consideration of structural design and prototype fabrication process is presented. A prototype is developed for verification and demonstration of the working principles.

The presented work, however, is a preliminary research about the cantilever-based current sensor. Further research is required for two reasons. Firstly, an explicit model of the dynamic response of the cantilever-based structure is needed due to complex electricity variation in different applications. Secondly, designs with sophistication are expected in practical application in which perturbations like ambient vibration and electromagnetic noise are presented effecting current measurements.

**Acknowledgments** This research was granted by the Japan Society for the Promotion of Science (JSPS) through the “Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program),” initiated by the Council for Science and Technology Policy (CSTP).

This work is partially supported by the National Natural Science Foundation of China (Grant No. 51675229). Part of this work is also financially supported by Scientific Research Foundation for Leading Professor Program of Jilin University (Grant No. 419080500171 & No. 419080500264).

## References

- Chen YC, Yu SC, Cheng SH, Cheng YT (2012) A flexible inductive coil tag for household two-wire current sensing applications. *Sensors J IEEE* 12(6):2129–2134
- Cowburn RP, Moulin AM, Welland ME (1997) High sensitivity measurement of magnetic fields using microcantilevers. *Appl Phys Lett* 71(15):2202–2204
- Franciscangelis C, Fruett F, Florida C, Bezerra EW, Leonardi AA, Cunha AS et al (2013) Theoretical and experimental comparison between power and degree of polarization based optical fiber current sensors. *J Microw Optoelectron Electromagn Appl* 12(2):617–627
- Goedeke SM, Allison SW, Datskos PG (2004) Non-contact current measurement with cobalt-coated microcantilevers. *Sens Actuators A Phys* 112(1):32–35
- Isagawa K, Wang DF, Kobayashi T, Itoh T, Maeda R (2011a) Development of a MEMS DC electric current sensor applicable to two-wire electrical appliance cord. In: Proceedings of the 2011 6th IEEE international conference on nano/micro engineered and molecular systems, Kaohsiung, pp 932–935
- Isagawa K, Wang DF, Kobayashi T, Itoh T, Maeda R (2011b) Developing MEMS DC electric current sensor for end-use monitoring of DC power supply. In: Proceedings of the 13rd international conference on design, integration & packaging of MEMS/MOEMS (DTIP 2011), Aix-en-Provence, pp 231–236
- Kobayashi T, Okada H, Masuda T, Itoh T (2010) A digital output piezoelectric accelerometer using patterned Pb(Zr, Ti)O<sub>3</sub>, thin films electrically connected in series. *Smart Mater Struct* 19(10):533–536
- Leland ES, Wright PK, White RM (2009) A MEMS AC current sensor for residential and commercial electricity end-use monitoring. *J Micromech Microeng* 19(9):094018
- Leland ES, Sherman CT, Minor P, White RM (2010) A new MEMS sensor for AC electric current. *Sensors* 143:1177–1182. IEEE
- Lim KS, Harun SW, Damanhuri SSA, Jasim AA, Tio CK, Ahmad H (2011) Current sensor based on microfiber knot resonator. *Sens Actuators A Phys* 167(1):60–62
- Osiander R, Ecelberger SA, Givens RB, Wickenden DK (1996) A microelectromechanical-based magnetostrictive magnetometer. *Appl Phys Lett* 69(69):2930–2931
- Phan TA, Hara M, Oguchi H, Kuwano H (2015) Current sensors using fe–b–nd–nb magnetic metallic glass micro-cantilevers. *Microelectron Eng* 135:28–31
- Popovic RS (2000) Not-plate-like hall magnetic sensors and their applications. *Sens Actuators A Phys* 85(1–3):9–17
- Ripka P (2004) Current sensors using magnetic materials. *J Optoelectron Adv Mater* 6(2):587–592
- Rossel C, Bauer P, Zech D, Hofer J, Willemin M, Keller H (1996) Active microlevers as miniature torque magnetometers. *J Appl Phys* 79(11):8166–8173
- Shutov MV, Sandoz EE, Howard DL, Hsia TC, Smith RL, Collins SD (2005) A micro-fabricated electromagnetic linear synchronous motor. *Sens Actuators A Phys* 121(2):566–575
- Sophian A, Gui YT, Taylor D, Rudlin J (2002) Design of a pulsed eddy current sensor for detection of defects in aircraft lap-joints. *Sens Actuators A Phys* 101(1–2):92–98
- Sun C, Wen Y, Li P, Ye W (2016) Self-contained wireless hall current sensor applied for two-wire zip-cords. *IEEE Trans Magn* 99(7):1–1
- Wang DF, Maeda R (2015) Analytical study on cantilever resonance type magnet-integrated sensor device for micro-magnetic field detection. *Microsyst Technol* 21(6):1167–1172
- Wang DF, Isagawa K, Kobayashi T, Itoh T (2012a) Developing passive piezoelectric MEMS sensor applicable to two-wire DC appliances with current switching. *Micro Nano Lett* 7(1):68–71
- Wang DF, Isagawa K, Kobayashi T, Itoh T (2012b) Passive piezoelectric DC sensor applicable to one-wire or two-wire DC electric appliances for end-use monitoring of DC power supply. *Microsyst Technol* 18(11):1897–1902
- Wang DF, Suzuki Y, Kobayashi T, Itoh T (2013a) Developing integrated piezoelectric direct current sensor with actuating and sensing elements. *Micro Nano Lett* 8(12):858–860

- Wang DF, Suzuki Y, Suwa Y, Kobayashi T, Itoh T, Maeda R (2013b) Integrated piezoelectric direct current sensor with actuating and sensing elements applicable to two-wire DC appliances. *Meas Sci Technol* 24(12):537–540
- Wang DF, Isagawa K, Kobayashi T, Itoh T, Maeda R (2013c) Passive piezoelectric single-side MEMS DC current sensor with five parallel PZT plates applicable to two-wire DC electric appliances without using cord separator. *Microsyst Technol* 19(6):923–927
- Wang DF, Member IEEE, Liu H, Li X, Li Y, Xian W, Kobayashi T, Itoh T, Maeda R (2017a) Passive MEMS DC electric current sensor: part I – theoretical considerations. *IEEE Sensor J* 17:1230
- Wang DF, Member IEEE, Liu H, Li X, Li Y, Xian W, Kobayashi T, Itoh T, Maeda R (2017b) Passive MEMS DC electric current sensor: Part II—experimental verifications. *IEEE Sensor J* 17(5):1238–1245
- Yang HH, Myung NV, Yee J, Park DY, Yoo BY, Schwartz M et al (2001) Ferromagnetic micro-mechanical magnetometers. *Transducers '01 eurosensors XV*. Springer, Berlin/Heidelberg
- Yang X, Liu H, Wang Y, Wang Y (2014) A giant magneto resistive (GMR) effect based current sensor with a toroidal magnetic core as flux concentrator and closed-loop configuration. *IEEE Trans Appl Supercond* 24(3):1–5
- Yang X, Li Y, Zheng W, Guo W (2015) Design and realization of a novel compact fluxgate current sensor. *IEEE Trans Magn* 51(3):1–4
- Zheng B-C, Yan S-C, Chen J-H, Cui G-X, Xu F, Lu Y-Q (2015) Miniature optical fiber current sensor based on a graphene membrane. *Laser Photon Rev* 9(5):517–522
- Zou Q, Tan W, Kim ES, Loeb GE (2008) Single- and triaxis piezoelectric-bimorph accelerometers. *J Microelectromech Syst* 17(1):45–57
- Xu Yang, Yupeng Fu, Dong F Wang (2017) Electrical and mechanical performance difference on piezoelectric segmentation in a passive MEMS DC current sensor applicable to two-wire DC appliances. *IOP Measurement Science and Technology* 28 (1):015101

---

## **Part IV**

# **Chemical Microsensors**



# Microcalorimeters for Detection of Trace Energetic Chemicals

Zheyao Wang and Wenzhou Ruan

## Contents

Introduction .....	660
Operation Principle of Microcalorimeters .....	662
Qualitative Model .....	662
Transduction Principles .....	665
Numerical Simulation .....	668
Static State .....	668
Transient State .....	671
Heated Reaction .....	673
Microcalorimeters for Trace Energetic Chemical Detection .....	674
Device Fabrication .....	676
In Situ <i>Integrated Carbon Nanotubes for Preconcentration</i> .....	677
Detection of TEC Using Microcalorimeters .....	681
References .....	685

## Abstract

Detection of trace energetic chemical (TEC) vapors is a challenging task because of the extremely low vapor concentrations of most TECs. Microcalorimeters, which consist of a suspended microbridge with integrated heaters and thermistors, are emerging as a powerful tool for fast detection of TECs. By heating the TEC molecules adsorbed onto the microcalorimeters to deflagration using the heaters and measuring the induced thermal responses and the total heat using the thermistors, microcalorimeters can detect TEC vapors through differential scanning calorimetry mode or differential thermal analysis mode. Due to the large surface areas, the small heat mass, and the rapid heating rates, the microcalorimeters are able to detect TEC vapors with low detection limits and fast detection rates.

Z. Wang (✉) · W. Ruan  
Tsinghua University, Beijing, China  
e-mail: [z.wang@tsinghua.edu.cn](mailto:z.wang@tsinghua.edu.cn)

**Keywords**

Calorimetry · Microcalorimeter · Microfabrication · MEMS · Microsensor · Trace energetic chemical

**Introduction**

Detection of trace energetic chemical (TEC) vapors, such as 2,4,6-trinitrotoluene (TNT) and 1,3,5-trinitro-1,3,5-triazacyclohexane (RDX), is a rapidly increasing task in forensics, anti-terrorist activities, global demining projects, and environment monitoring. Most of these applications, if not all, need portable detection technologies with low limit of detection (LOD), high selectivity, and fast detection time. This is a result of the extremely low partial pressures of the chemicals in the air, complex interferences, and various kinds of TECs (Nambayah and Quickenden 2004; Singh 2007). For example, the partial pressure of TNT is a few parts per billion (ppb) at room temperature and in the sub-part per trillion (ppt) range in the air above a buried mine (Bogue 2011).

The extreme vapor pressures of TECs make the detection of trace chemicals in a complex environment a still significant technological challenge. While some commercial products for detecting TECs are now available, the need of highly sensitive and fast detection methods is continuously increasing and the research into alternative technologies continues at a pace (Moore 2004). Various detection methods with potential to achieve low LOD are being investigated for TEC detection, such as electrochemical sensors (Wang 2007; Patel et al. 2003), biosensors (Shankaran et al. 2005), fluorescence- and Raman-based optical methods (Naddo et al. 2008; Sun et al. 2010), mass spectrometry (MS) (Mullen et al. 2006), ion mobility spectrometry (IMS) (Ewing et al. 2001), and sensors based on nano- and microfabrication technologies (Aguilar et al. 2010; Lee et al. 2002; Pinnaduwege et al. 2003, 2004a, 2005; Pinnaduwege 2004b; Yi et al. 2008).

Electrochemical sensors provide some selectivity (Wang and Thongngamdee 2003) but suffer from limited sensitivity and require mobile electrolytes, which may cause stability issues and delayed response time. Also, electrodes can be easily fouled, and interfering problems may occur as some interferences are electrochemically active. Nanomaterials, e.g., carbon nanotubes (CNTs) and nanowires, have been employed to construct electrochemical sensors for detection of TEC vapors (Sergio and Arben 2012). Typical CNT applications include CNT field-effect transistors (FET) or resistors (Chen et al. 2010) and CNT-modified glassy carbon electrodes (Hrapovic et al. 2006). More recently, CNTs have been employed in TEC sensors by exploiting passive functions rather than transduction, such as enhanced Raman scattering (Sun et al. 2010) and preconcentration of target samples (Camara et al. 2011). LODs down to ppb level have been achieved for TNT detection by exploiting the extremely large surface-to-volume ratio (Roberts et al. 2009). Despite high sensitivity, significant challenges, such as low selectivity, low signal-to-noise ratio (SNR), long recovery times, interference, and device fabrication difficulties, still remain to be addressed (Sinha et al. 2006; Zribi et al. 2005).



IMS/MS exhibits inherent advantages like low LODs in TEC vapor detection, but they need sophisticated instruments either bulky or expensive. To improve selectivity, separation techniques such as gas chromatography (GC) (Agah et al. 2005; Lu et al. 2003) are needed to separate target chemicals before detection, which involves time-consuming on-site sample collection and pretreatment or transportation to laboratories. Handheld IMS with detection time less than 10 s are now commercially available. However, the radioactivity of the ion sources and the expensive total cost of ownership limit mass-deployable applications. Recently, microfluidic devices, which can integrate multiple sample handling processes with on-chip sensors, have been investigated for TEC detection with the potential of providing on-site separation of explosives (Wang 2004).

Microsensors also attract considerable research attention for trace chemical detection. Two typical representatives are the microcantilevers, which can detect TEC vapors by measuring chemical adsorption-induced resonant frequency shifts (Xu et al. 2010) or microcantilever deflection (Pinnaduwege et al. 2007), and the microcalorimeters, which measure thermal reaction of TEC vapors (Carreto-Vazquez et al. 2010; Cavicchi et al. 2004; Liu et al. 2007; Piazzon et al. 2010; Youssef et al. 2009; Zuck et al. 2008). Microcantilevers usually use a self-assembled monolayer (SAM) immobilized on the surfaces to selectively adsorb target chemicals through covalent binding. Although the minute dimensions enable low LODs, several limitations exist. For example, water vapors that have concentrations orders of magnitude higher than TEC vapors severely restrict the selectivity when certain SAMs are used (Senesac and Thundat 2008; Southworth et al. 2010). Some interferences adsorbed or landing on the sensor surfaces occupy the effective adsorbing areas and thus negatively influence the performance. In addition, long adsorption and recovery times (if applicable), normally hundreds to thousands of seconds, are needed for microcantilevers to reach stable states (Goeders et al. 2008; Xu et al. 2010).

Microcalorimeters, which have been widely applied for measuring thermal properties of chemical and biological reactions by analyzing the exothermic and endothermic characteristics of thermal reaction (Carreto-Vazquez et al. 2010; Piekiet et al. 2011; Wang et al. 2008), were first employed for TEC vapor detection in 1999 (Chen et al. 1999). Due to the extremely low thermal mass, microcalorimeters with LOD down to 6 pg for TNT detection (Pinnaduwege et al. 2003; 2004; 2005; Pinnaduwege LA; Yi et al. 2008) and 3 ppm for combustible gases (Cavicchi et al. 2004) have been achieved. The capability in measuring the thermal fingerprints of target chemicals also allows microcalorimeters to be able to differentiate TECs. However, the low vapor pressures of most TECs and the small surfaces of the microcalorimeters need long times to adsorb enough TEC molecules for detection, normally around several tens to hundreds seconds (Pinnaduwege et al. 2003).

Although considerable research effort has been allocated and different technologies have been developed to tackle this issue, many obstacles still remain for TEC detection in portable applications, especially for detection of TEC vapors (Fainberg 1992; Steinfeld and Wormhoudt 1998). Dogs remain the best detectors for TEC vapors to

date (Furton and Myers 2001); however, widespread deployment of canine teams is neither practical nor cost effective.

## Operation Principle of Microcalorimeters

A microcalorimeter normally consists of a microstructure for adsorption of TEC molecules, a heating device for triggering thermal processes of the TEC, and a temperature sensor for measuring the thermal properties of the thermal processes. The microstructure is commonly a microcantilever, a microbridge, or a microplate. When the microcalorimeter is exposed to TEC vapors, TEC molecules are adsorbed onto the surfaces of the microstructure. By heating the microstructure to specific temperatures, thermal processes such as melting, decomposition, or deflagration of the TEC molecules occur on the surfaces of the microstructure. The thermal processes are inherently endothermic or exothermic, which cause temperature changes of the microstructure by heat transfer. The temperature changes or the corresponding heat changes of the microstructure are measured by the temperature sensor to quantitatively analyze the adsorbed TEC. It is apparent that the microcalorimeters should have low heat capacity, such that the temperature changes induced by the thermal processes can be maximized.

## Qualitative Model

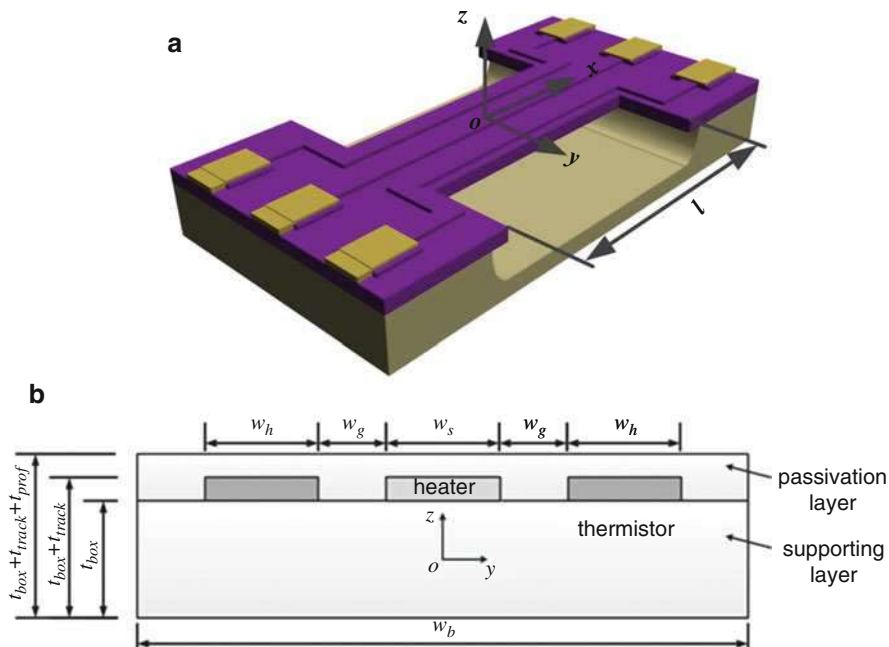
A theoretical model takes a microbridge as an example to demonstrate the temperature changes in a thermal process. A simplified microbridge model is illustrated in Fig. 1. The microbridge is made of silicon dioxide and is suspended to minimize the thermal dissipation through the substrate. The heating device and the temperature sensor are made of silicon and are embedded in the microbridge. For simplification, the temperature gradient across the thickness of the microbridge is ignored. It is reasonable to deduce this simplification in a relatively long dynamic process because the tiny thickness of the microbridge (normally less than 1  $\mu\text{m}$ ) allows a temperature difference of 100  $^{\circ}\text{C}$  across the thickness to vanish in 0.4  $\mu\text{s}$ . The temperature is also assumed to be uniform across the width of the microbridge. Therefore, the thermal transfer problem of the microbridge is simplified as a one-dimensional problem.

The total thermal conductivity,  $K$ , along the length of the microbridge is given by

$$K = \sum_{i=1}^n k_i w_i t_i \quad (1)$$

where  $k_i$  is the thermal conductivity of the  $i^{\text{th}}$ -layer material of the microbridge.

By denoting the coefficient of heat convection for the microbridge to the air with  $h$ , one can obtain the thermal conduction of a unit length by convection,  $H$ :



**Fig. 1** Schematic illustration of a microbridge-based microcalorimeter. (a) Whole configuration. (b) Cross-sectional view of the microbridge

$$H = 2h \left( w_b + \sum_{i=1}^n t_i \right). \tag{2}$$

Although constant thermal conductivities are used for these materials, it should be noted that if the mean free path of phonons exceeds the thickness of a thin structure, the interface phonon scattering may cause distinct reduction in thermal conductivity (Ju and Goodson 1999). The thermal conductivity of a thin structure with a thickness of several hundred nanometers is about 10%–20% lower in magnitude than the corresponding bulk material (Bruzo et al. 2003; Kleiner et al. 1996) and may reduce as much as 50% if the thickness is less than 100 nm (Ju and Goodson 1999; Yamane et al. 2002).

In static state, the equation of thermal conduction can be expressed as

$$\nabla \cdot k \nabla T(x, y, z) + q(x, y, z) = 0 \tag{3}$$

where  $k$ ,  $T$ , and  $q$  are, respectively, the thermal conductivity, the temperature, and the power density at a place with coordinate  $(x, y, z)$ .

As aforementioned, the temperature distribution as well as the power density changes only along the length, i.e.,  $x$ -axis; Eq. (3) can be rewritten as following by assuming  $P(x)$  as a power density per unit length:

$$\frac{\partial^2 T}{\partial x^2} - \frac{H(T - T_0)}{K} + \frac{P(x)}{K} = 0 \quad (4)$$

where  $T_0$  denotes the room temperature and is treated as a constant, and  $K$  is given in Eq. 1.

Equation (4) takes both heating power and thermal dissipation through convection into consideration. For a simple case that convection is neglected and the thermal parameters of the materials are constant during heating, one finds an approximate quadratic solution for Eq. (4):

$$T(x) = -\frac{P_0}{2K}x^2 + T_1 + \frac{P_0}{8K}l^2 \quad (5)$$

where  $P_0$  is the constant heating power and  $P(x) = P_0$ , and the temperature at the end of the microbridge is  $T_1$ .

The temperature distribution solved from Eq. (5) indicates that the highest temperature occurs at the center of the microbridge, and the lowest temperature occurs at either end of the microbridge. In addition, the temperature is proportional to the heating power, i.e., the larger the heating power, the higher the temperature. Equation (5) implies that large temperature changes can be obtained by increasing the heating power, increasing the length of the microbridge, and decreasing the thermal conductivity.

As a microcalorimeter for TEC detection, increasing the heating power can be achieved by absorbing more TEC molecules. Because long absorption time is not always allowed in applications, one alternative way is to enhancing the absorption ability by using large surface areas. Increasing the length of the microbridge benefits from the resultant reduced thermal conductivity rather than the increased area because the thermal dissipation of the microbridge also increases with the length.

Detection of TEC by heating the TEC to a specific temperature to trigger chemical reactions is a dynamic thermal process. The chemical reactions of TEC are inherent endothermic or exothermic processes and cause temperature changes versus time that can be measured. For a dynamic process, the heat transfer can be described by the following equation (Lienhard and Lienhard 2005):

$$\frac{\partial^2 T}{\partial x^2} - \frac{H(T - T_0)}{K} + \frac{P}{K} = \frac{\rho C}{K} \frac{\partial T}{\partial t} \quad (6)$$

where  $\rho$  and  $C$  are, respectively, the density and the heat capacity of the microbridge.

By neglecting thermal dissipation from the microbridge, i.e., the heat generated by the heating power is completely adsorbed by the microbridge, one has

$$Q = P\Delta t = \rho V C \Delta T \quad (7)$$

where  $Q$  is the total heat,  $\Delta t$  is the duration of the heating power,  $V$  is the volume of the microbridge, and  $\Delta T$  is the temperature changes.

Equation (7) shows that in a dynamic process, decreasing the total heat capacity  $\rho VC$  is beneficial to increasing the temperature change  $\Delta T$ . If the material and thus the heat capacity  $C$  are predetermined, small volume allows large temperature changes. However, it should be noted that reducing the volume of the microcalorimeter is only effective to reducing the heat capacity when it is achieved by decreasing the thickness of the microbridge while maintaining the surface; otherwise, the adsorption ability and thus the heating power reduce proportionally. That is, decreasing the volume and increasing the surface areas are helpful to improve the performance.

The aforementioned models are simplified models by assuming a uniform temperature distribution, considering material properties as constant, and neglecting thermal dissipation. Taking these factors into consideration makes the models significantly complex. For example, if large temperature gradients exist along the length of the microbridge, only part of the adsorbed TEC close to the center of the microbridge deflagrates because of high temperatures. Although these factors make the models more complex, the conclusions are consistent with those deduced from simplified models.

## Transduction Principles

Temperature-sensitive devices that can be easily microfabricated include silicon diodes, thermal couples, and metal or silicon thermistors. Silicon diodes are only applicable to low temperature because the high temperatures to deflate TEC (normally higher than 200 °C) lead to large amount of thermal carriers in silicon that may cause failure of semiconductor devices. Moreover, the temperature coefficients of silicon diodes decrease remarkably with the increase in temperatures. Besides these direct temperature-sensing methods, the secondary effects of temperature, such as temperature gradient-induced bending of bimorphs that can be measured by integrated piezoresistors, are also applicable to temperature sensing.

### Resistance Change Through Piezoresistivity

Bimorph structures bend when loaded with temperature changes because of the difference in the coefficients of thermal expansion (CTE). The bending can be measured by implanted silicon piezoresistors, of which the resistance varies with changes in carrier mobility as a result of strain (lattice deformation) induced by bending.

For a commonly used silicon piezoresistor on a long and thin microcantilever (one end fixed and the other free), the out-of-plane stresses caused by bending are relatively small compared with in-plane stresses and thus are ignored. Then the resistance changes of the piezoresistor can be expressed as

$$\frac{dR}{R} = \frac{d\rho}{\rho} = \pi_l \sigma_l + \pi_t \sigma_t \quad (8)$$

where  $dR/R$  is the relative changes of the resistance  $R$  of a silicon piezoresistor,  $\rho$  is the conductivity,  $\pi_l$  and  $\pi_t$  are the piezoresistivity coefficients along the longitudinal and transverse directions of the piezoresistor, respectively, and  $\sigma_l$  and  $\sigma_t$  are the stresses of the corresponding directions.

Using the assumptions of uniform temperature distribution and uniform doping concentration across the thickness of the piezoresistor, one can obtain the sensitivity of temperature-induced resistance changes,  $S_T$ , which is (Wang et al. 2005)

$$S_T \triangleq \frac{\partial(\Delta R/R)}{\partial(\Delta T)} = \frac{(\pi_l + \pi_t)E_p(z_p - z_N) \sum_{i=1}^n \left[ E_i w_i t_i \left( \frac{\sum_{i=1}^n E_i w_i t_i \alpha_{T,i}}{\sum_{i=1}^n E_i w_i t_i} - \alpha_{T,i} \right) \left( \frac{t_i}{2} + \sum_{j=1}^{i-1} t_j \right) \right]}{\sum_{i=1}^n \left\{ \frac{E_i w_i t_i^3}{12} + E_i w_i t_i \left[ \left( \frac{t_i}{2} + \sum_{j=1}^{i-1} t_j \right) - z_N \right]^2 \right\}} \quad (9)$$

where  $\Delta T$  is the temperature change;  $E_p$  is Young's modulus of the piezoresistor;  $E_i = E/(1 - \nu)$  is biaxial modulus of the  $i$ -th layer material;  $E$  and  $\nu$  are, respectively, its Young's modulus and Poisson's ratio;  $w_i$  and  $t_i$  are its width and thickness, respectively; and  $\alpha_{T,i}$  is its coefficient of thermal expansion.  $z_p$  and  $z_N$  denote the coordinates in thickness of the neutral plane of the piezoresistor and the multilayer structure, and  $z_p - z_N$  represents the distance between the neutral plane of the piezoresistor and the multilayer structure.

Equation (9) explicitly describes the resistance change induced by thermal expansion is a function of material properties and dimensions. By using Eq. (9) and silicon and silicon dioxide as the microcantilever materials, the temperature-induced sensitivity of resistance is on the order of 1 ppm/K. Although placing the resistors close to the fixed end of the microcantilever exploits the stress concentration and enhances the sensitivity (Yu et al. 2007), the estimated result is also roughly on the order of 1~10 ppm/K.

### Resistance Change Through Temperature Coefficient of Resistance

The resistivity of semiconductor material such as silicon is determined by the carrier concentration and mobility, which are temperature dependent. Therefore, a silicon resistor is directly sensitive to temperature changes and can be used for temperature sensing.

The temperature coefficient of resistance (TCR) can be expressed as

$$TCR = \frac{d\rho/\rho}{dT} \quad (10)$$

where the resistivity  $\rho$  is given by

$$\rho = \frac{1}{nq\mu_n + pq\mu_p} \quad (11)$$

where  $q$  is the amount of unit electrical charge;  $\mu_n$  and  $\mu_p$  are, respectively, the mobilities of n-type and p-type carriers; and  $n$  and  $p$  are, respectively, the concentrations of n-type and p-type carriers. For silicon, the mobility is determined by lattice scattering and ionized impurity scattering and can be given by

$$\mu = \frac{q}{m^*} \frac{1}{AT^{3/2} + BN_i/T^{3/2}} \quad (12)$$

where  $m^*$  denotes the effective mass,  $T$  is temperature,  $N_i$  is the impurity concentration,  $A$  and  $B$  are constants,  $AT^{3/2}$  represents lattice scattering, and  $BN_i/T^{3/2}$  represents ionized impurity scattering.

The intrinsic carrier mobility  $n_i$  is given by

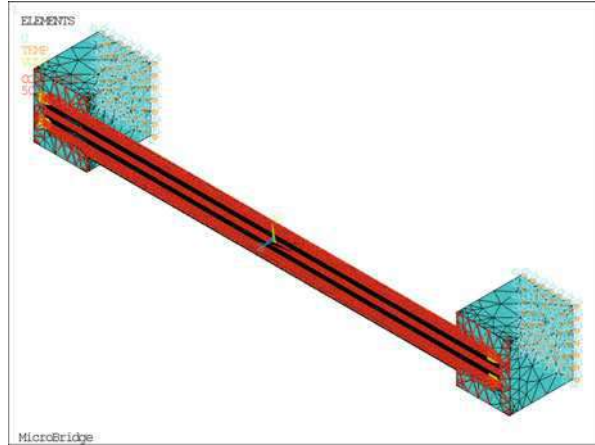
$$n_i = N_S \exp\left(-\frac{E_g}{2k_B T}\right) \quad (13)$$

where  $N_S$  is the effective state concentration which is on the order of  $10^{19} \text{ m}^{-3}$  at room temperature and increase with temperature,  $E_g$  is the energy gap, and  $k_B$  is Boltzmann's constant.

Equations (12) and (13) show that the carrier concentration and the mobility increase with temperature, and thus the resistivity decreases with temperature. In fact, the dependence of the resistivity on temperature is more complex. In low temperatures, the intrinsic excitation is so low that it can be neglected, and the carriers are mainly originated from doped impurities. The mobility of impurities increases with temperature because of the enhanced ionized impurity scattering, and hence the resistivity decreases with increase in temperature according to (12). For relatively high temperatures, the doped impurities are completely ionized, whereas ionization of intrinsic carriers is still insignificant, and thus the carrier concentration does not vary significantly with temperature change. However, at high temperatures lattice vibration dominates carrier scattering, reducing the mobility and the resistivity with temperature increase. This temperature range covers room temperature and is the common case for most sensors. At higher temperatures, intrinsic excitation causes significant increase in carrier concentration and the conduction mechanism falls into intrinsic conduction. For example, intrinsic carrier concentration doubles every temperature increase of  $8^\circ \text{C}$ . Therefore, at high temperatures changes of the intrinsic carriers with temperature dominate the resistivity, and the resistivity decreases with temperature increase. This inflection temperature is determined by the bandgap and the doping concentration.

As a temperature sensor, a large temperature range with a monotonic relationship between resistance and temperature is needed. To make a trade-off between

**Fig. 2** Mesh illustration for simulation



monotonic ranges and temperature coefficients, a doping concentration in the range of  $10^{19}\sim 10^{20}/\text{cm}^3$  is proper for TEC detections, in which the temperature coefficient is on the order of  $10^3$  ppm/K (Boukabache and Pons 2002).

## Numerical Simulation

### Static State

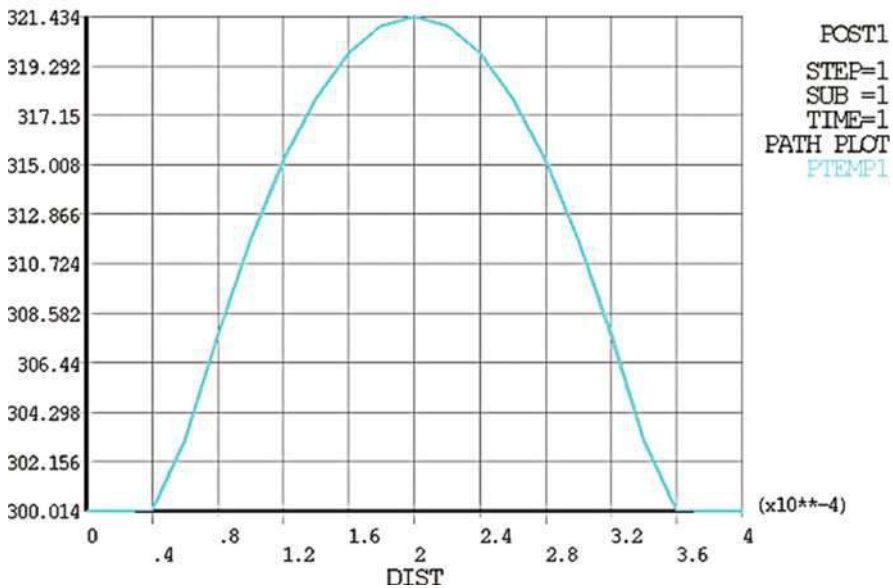
Figure 2 shows the mesh of the microcalorimeter for static simulation using ANSYS™ with element number of Solid 227. The microbridge is heated by two resistive heaters embedded in the microbridge with a voltage of 5 V. The temperature at the center of the microbridge is obtained and used to calculate the stresses induced by thermal expansion. The thermal stresses are then used to calculate the resistance changes through piezoresistivity. The substrate is treated as deformation-free during heating and its temperature is set to 300 K as a constant. Thermal dissipation includes heat transfer to the substrate and convection to the ambient environment with a convection factor of  $10 \text{ W}/(\text{m}^2\cdot\text{K})$ . All the materials are assumed as isotropic, and the material properties are listed in Table 1. Note that because of the variation in the stresses through the length of the piezoresistor, the relative changes of the piezoresistor per temperature change (K) are obtained by integration using Eq. (9).

Figure 3 shows the simulated static-state temperature distribution across the length of the microbridge. The temperature follows roughly a parabolic relation, with the highest temperature appearing at the center of the microbridge and the lowest temperature at the two ends of the microbridge. The temperatures at the two ends are close to 300 K. Although the magnitude of the highest temperature varies with the piezoresistance, the dimensions and the materials of the microbridge, and the externally applied heating power, the temperature curves follow the same distribution for various microcalorimeters.



**Table 1** Material parameters for simulation

Parameter	Definition	Value
$h_2$	Thickness of the heating resistor	200 nm
$h_1$	Thickness of the protection layer	400 nm
$\rho_2$	Density of the heating resistor	2328 kg/m <sup>3</sup>
$\rho_1$	Density of the protection layer	2200 kg/m <sup>3</sup>
$\alpha_2$	Thermal expansion coefficient of the heating resistor	$2.6 \times 10^{-6}/K$
$\alpha_1$	Thermal expansion coefficient of the protection layer	$0.5 \times 10^{-6}/K$
$K_2$	Thermal conductivity of the heating resistor	157 W/(m K)
$K_1$	Thermal conductivity of the protection layer	1.4 W/(m K)
$E_2$	Young's modulus of the heating resistor	$130 \times 10^3$ MPa
$E_1$	Young's modulus of the protection layer	$70 \times 10^3$ MPa
$C_2$	Specific heat of the heating resistor	0.702 J/(g K)
$C_1$	Specific heat of the protection layer	1.0 J/(g K)



**Fig. 3** Simulated static temperature distribution along the length of the microbridge

Figure 4 shows the simulated stress distribution across the length of the microbridge caused by thermal expansion. As expected, the out-of-plane stress,  $\sigma_z$ , is very small and thus can be ignored. Because the two ends of microbridge are clamped at the substrate, stress concentration appears symmetrically at the two places close to the two ends ( $\pm 150 \mu\text{m}$ ). The stress in the longitudinal direction is roughly twice in magnitude of that in the transverse direction at the same location.

Figure 5 shows the piezoresistance-temperature sensitivity versus the length and the width of the microbridge, i.e., the changes in the resistance of the piezoresistor

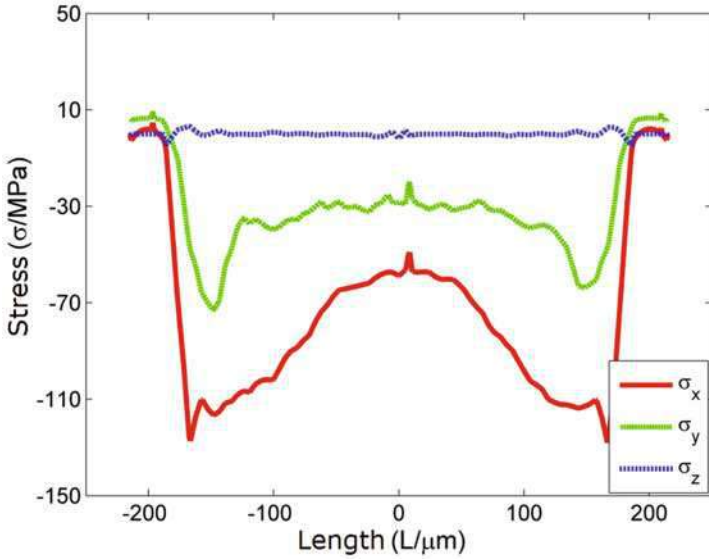


Fig. 4 Simulated thermal stresses along the length of the microbridge

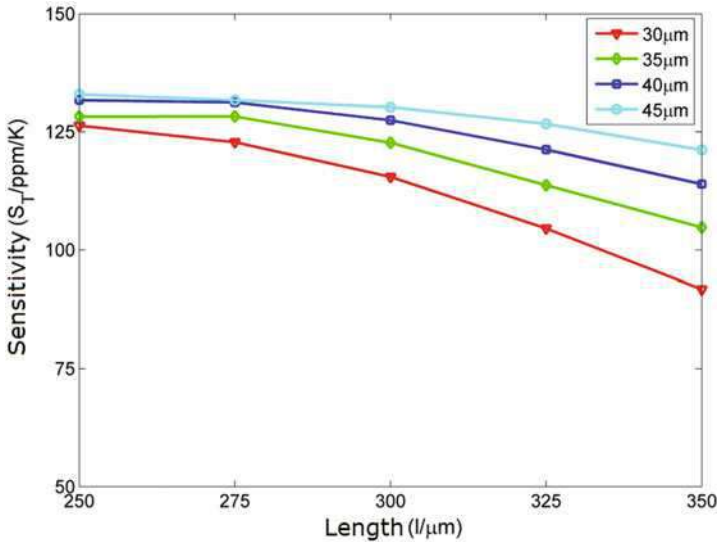


Fig. 5 Simulated changes of resistance per K versus the length and width of the microbridge

per unit temperature change (K). The longitudinal direction of the piezoresistors is in the  $[110]$  direction of a  $(100)$  silicon wafer and the transverse direction is in  $[\bar{1}\bar{1}0]$  direction. It can be found that increase in the length of the microbridge decreases the piezoresistance-temperature sensitivity because the stress concentration areas of

a short microbridge occupy a relatively large portion of the microbridge. Increase in the width of the microbridge increases the sensitivity because the influence of the end clamp of the microbridge is more significant than a narrow microbridge.

It also shows that the piezoresistance-temperature sensitivity is around  $10^2$  ppm/K, which is about two orders higher in magnitude than theoretical estimation. This significant improvement is attributed to the fixed boundary condition, which is not taken into consideration in theoretical estimation. Compared with the resistance-temperature sensitivity, which is on the order of  $10^3$  ppm/K, the temperature-induced piezoresistance change is far low. The thermal expansion induced resistance changes have the same contribution to the resistance changes, they may superposed in applications.

## Transient State

Figure 6 shows the device model for transient simulation, in which only thermal processes are considered. The device is meshed with Solid 98 elements, and the substrate of the microbridge is set to 300 K. The dimension of the microbridge is  $300 \times 35 \mu\text{m}$ , and the heating power of the heaters is  $5.21 \times 10^{12} \text{ W/m}^3$ . When the temperature reaches the deflagration temperature of TNT, an extra heating power of  $2.52 \times 10^{11} \text{ W/m}^3$  is added to mimic the heat released from deflagration of TNT with a density of  $10 \text{ fg}/\mu\text{m}^2$  in 1 ms. Table 2 shows the properties of TNT.

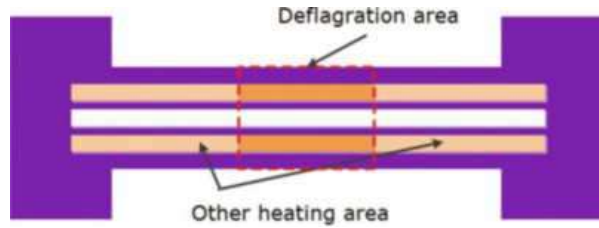
Figure 7 shows the transient temperature distribution. A temperature range between 570 and 583 K is considered as a proper range for TNT deflagration according to the TNT properties, and the areas with temperature higher than 583 K have no TNT because of deflagration. As deflagration of TNT is an inherently exothermic reaction, an extra heat is generated.

When solid-state TNT is heated in a container, TNT sublimates and its vapor pressure increases significantly with the heating temperature. It is proved that the relationship between the sublimation rates of TNT and temperature follows Arrhenius equation (Gershanik and Zeiri 2010). As the TNT sublimation rate at  $70^\circ\text{C}$  is about  $10.1 \text{ ng}/(\text{cm}^2 \cdot \text{s})$ , for a microbridge with a size of  $300 \times 35 \mu\text{m}$  and a uniform temperature distribution, the loss of TNT through sublimation in 30 ms is about 20 fg. The heat needed for sublimation of 20 fg TNT is about 9.7 pJ, which causes a 0.97 K temperature decrease of the microbridge that has a heat capacity of 10 nJ/K.

If the temperature reaches  $80^\circ\text{C}$ , TNT begins to melt and the phase changes from solid-state to liquid-state. In phase change, TNT absorbs heat which converts to melting latent heat, but the temperature remains unchanged. According to Table 2, the heat needed for melting of 100 pg TNT is about 9.8 nJ, which is large enough to generate a temperature decrease of the microbridge as large as 1 K. After melting, the thermal conductivity of liquid TNT is improved because of the existence of thermal conduction and convection, and the temperature rises more quickly than solid-state when being heated.

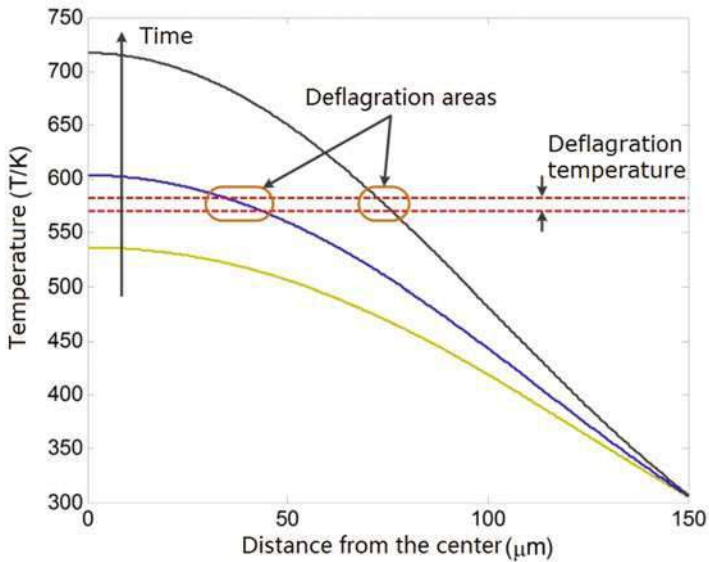
The heat transfer in TNT phase change can be given as (Sun and Garimella 2005)

**Fig. 6** Device model for transient process simulation



**Table 2** The properties of TNT (Gershanik and Zeiri 2010; Hobbs et al. 2009; Sun and Garimella 2005)

Symbol		Definition	Value
$\rho$		Density	1560 kg/m <sup>3</sup>
$c_{ps}$		Specific heat(21 °C)	1130 J/(kg K)
$k_s$		Solid-state conductivity	0.25 W/(m K)
$k_l$		Liquid-state conductivity	2 W/(m K)
$T_m$		Melting point	80.9 °C
$T_d$		Explosion point	295 °C
$\Delta H_s$		Sublimation enthalpy	484 kJ/kg
$\Delta H_m$		Melting latent heat	98.4 kJ/kg
$\Delta H_d$		Reaction enthalpy	5400 kJ/kg



**Fig. 7** The temperature distribution of the microbridge in transient heating

$$\nabla^2 T = \frac{\rho c}{k} \frac{\partial T}{\partial t} = \frac{\rho}{k} \left( c_p + \frac{\partial f_l}{\partial T} \Delta H \right) \frac{\partial T}{\partial t} \quad (14)$$

where  $\rho$  is the density of TNT,  $k$  is the thermal conductivity, and  $c$  is the effective heat capacity and includes the heat capacity of solid-state TNT,  $c_{ps}$ , and the influence of latent heat,  $\partial f_l / \partial T \cdot \Delta H$ , where (Lamberg et al. 2004; Sun and Garimella 2005)

$$f_l = \begin{cases} 1 & T > T_m + {}^m\varepsilon \\ \frac{T - T_m + {}^m2\varepsilon}{2\varepsilon} & |T - T_m| \leq \varepsilon - {}^m\varepsilon \\ 0 & T < T_m \end{cases} \quad (15)$$

where  $T_m$  is the melting point. In the range of  $T_m \pm \varepsilon$ , TNT melts and absorbs heat, but the temperature remains unchanged. Therefore, in phase change the external heat provides the energy for changing from solid-state to liquid-state, and the heat capacity jumps without temperature change. This takes enthalpy change into consideration. In numerical calculation, specific heat and temperature change can be treated by converting into enthalpy using Eqs. (14) and (15).

After melting, the liquid TNT evaporates and the evaporation rates increase with heating temperature until 295 °C at which deflagration occurs. Deflagration is an exothermic reaction, which releases a large amount of heat to be detected by temperature sensors. To avoid evaporation of liquid TNT during heating, the heating duration should be short.

## Heated Reaction

To simulate the thermal reaction process, the TNT properties listed in Table 2 and the microbridge properties listed in Table 1 are used. The heating voltage is a triangular wave with a period of 50 ms. For simplification, evaporation during heating is neglected. TNT melting at around 81 °C is considered by defining the relation between per unit enthalpy of the material and temperature (Hess and Miaoulis 1995):

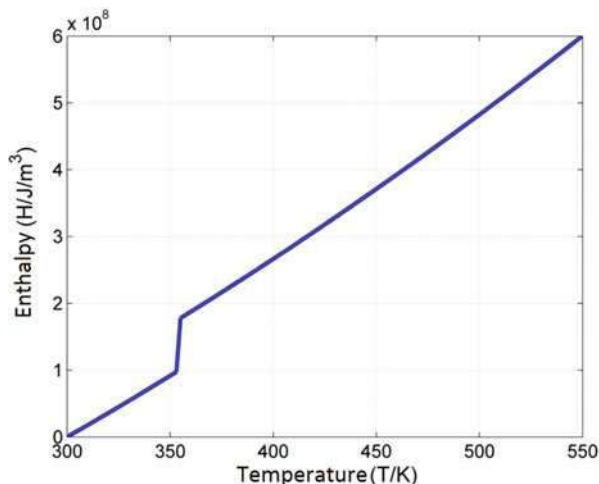
$$H = \int \rho c(T) dT \quad (16)$$

If melting occurs at 80~82 °C, by neglecting the density changes with temperature and by assuming that the heat capacity of TNT,  $c_{TNT}$ , follows a linear relation with temperature (Hobbs et al. 2009), one has

$$c_{TNT} = 1.677T + 630 \quad (17)$$

Then the changes in the enthalpy of TNT with latent heat considered vs. temperature can be obtained, and the curve is shown in Fig. 8.

**Fig. 8** The changes of enthalpy vs. temperature



It is assumed that the TNT completely deflagrate when heated to 295 °C to simplify the very complex thermal reaction. The heat released from deflagration is applied to the microbridge as an external thermal load. Figure 9 shows the temperatures of the microbridge with and without TNT versus heating time. It can be seen that without TNT, the temperature increases smoothly with increase with the heat time. With TNT, the temperature increase slows down at 25 ms, when melting occurs and the heat supplied by the heating power converts to latent heat. At around 47 ms, the temperature increases sharply because deflagration occurs which converts the chemical energy in TNT to heat.

By using an identical microbridge as a reference on which there is no TNT, the temperature differences between the sensing microbridge and the reference microbridge represent the temperature changes caused by TNT deflagration, i.e., the difference of the two curves shown in Fig. 9.

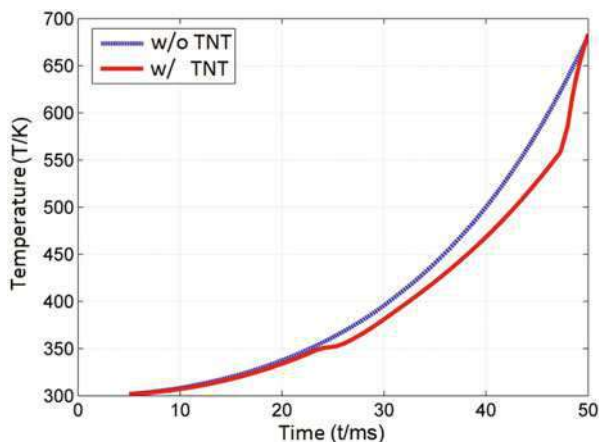
Figure 10 shows the influences of the duration of the heating pulse on the temperature changes. It can be seen that with a short heating pulse, the temperature curve shifts to the left and the temperature peak is higher. This can be expected as short heating pulse reduces the time and the amount of TNT evaporation, allowing more TNT to be heated to deflagration. The uniformity of the heating temperature across the microbridge is also important because uniform temperature distribution allows simultaneous deflagration on the whole microbridge, avoiding evaporation and gradual deflagration which reduce the total released heat of deflagration.

---

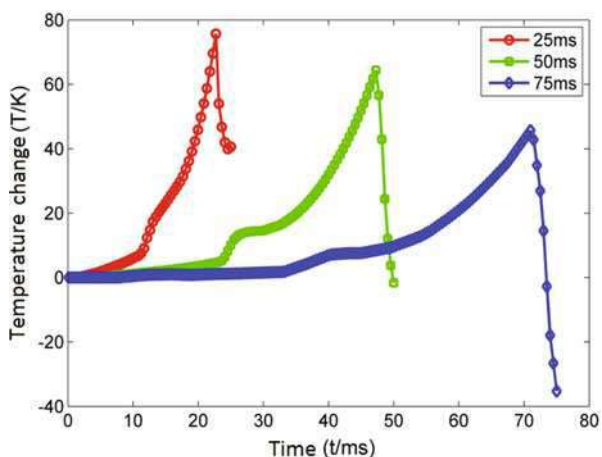
## Microcalorimeters for Trace Energetic Chemical Detection

The microcalorimeter, as shown in Fig. 11, consists of a suspended microbridge, two heating resistors, and a temperature sensor (thermistor). The microbridge should be made of materials with low thermal conductivity and low heat capacity to maximize

**Fig. 9** The temperature changes vs. time

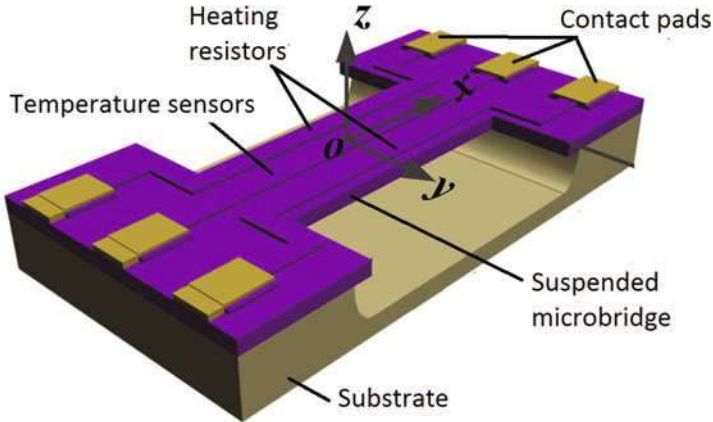


**Fig. 10** The influences of the pulse duration on the temperature change



the temperature increase. The heating resistors and the thermistor should be made of materials with high thermal conductivity to achieve quick thermal transfer. In addition, all the materials should be able to withstand temperatures as high as 400 °C.

Upon these requirements, the microbridge is made of two layers of silicon dioxide. The bottom silicon dioxide layer serves as a supporting structure and the top one acts as a passivation layer. The heating resistors and the thermistor, which are made of single crystalline silicon, are sandwiched in between the two layers of silicon dioxide. As the heat released from TEC deflagration is extremely low, the microbridge is suspended to minimize thermal dissipation to the substrate, such that the temperature increase of the microbridge by TEC deflagration is maximized and is detectable. This is achieved by removing the materials beneath the microbridge.



**Fig. 11** Schematic illustration of the microcalorimeter

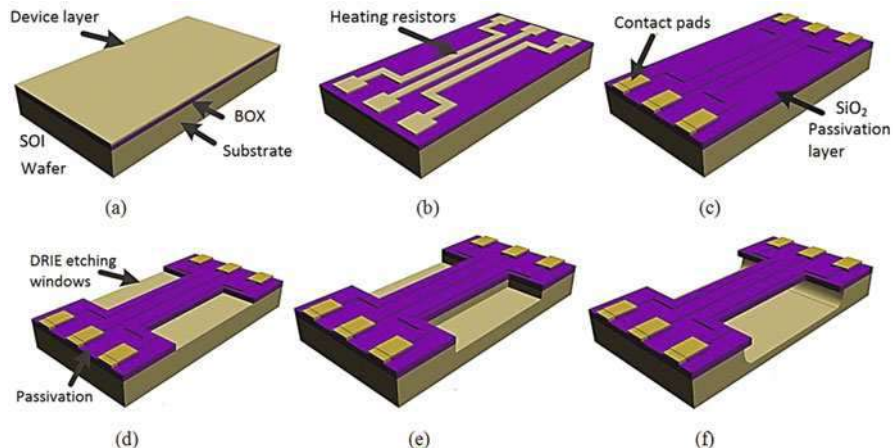
## Device Fabrication

The microcalorimeter is fabricated using microfabrication technologies. The detailed fabrication processes are illustrated in Fig. 12 (Zhou et al. 2009). As shown in Fig. 12a, the microcalorimeter is fabricated from a silicon-on-insulator (SOI) wafer. The device layer (single-crystalline silicon) of the SOI wafer is used to fabricate the heating resistors and the thermistor, and the buried oxide (BOX) layer of the SOI wafer is used to fabricate the supporting layer. Because of the low thermal conductivity of silicon dioxide, the thermal dissipation through the two ends of the microbridge to the substrate is minimized.

As shown in Fig. 12b, the device layer is then implanted with boron after patterning using lithography. Patterning allows the heating resistors and the thermistor, as well as the contact areas, to be doped with different boron concentrations, such that the resistances can be optimized. Then the heating resistors and the thermistor are defined by reactive ion etching (RIE), followed by deposition and etching of contact metal. To avoid device noise caused by damages in RIE etching, it is desired to remove the damaged layer after RIE etching using wet etching. The two heating resistors are placed on the two sides of the thermistor to achieve uniform temperature distribution across the width of the microbridge. Heating resistors with more complicated shapes can be used to achieve uniform temperature profiles across the length. For example, by using lens shape heaters, the nonuniformity of the heating temperature is reduced from 12.5% to 5.4% (Greve et al. 2009).

As shown in Fig. 12c, a  $\text{SiO}_2$  passivation layer is deposited using CVD. Then the contact windows are opened by RIE etching of the passivation layer. The passivation layer should be dense to protect the resistors from being affected by contaminations. As shown in Fig. 12d, the passivation layer together with the BOX layer is etched sequentially using RIE to define the releasing windows, through which dry etching is performed to remove the substrate beneath the microbridge.





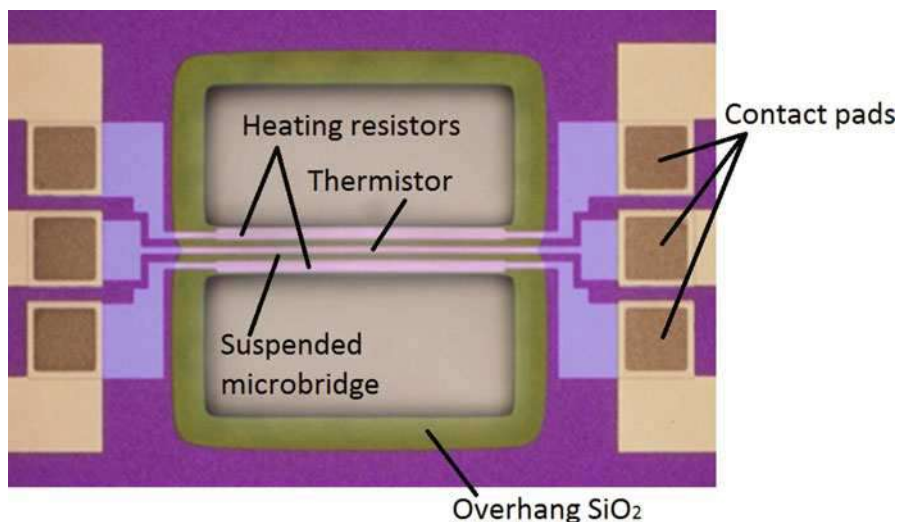
**Fig. 12** Fabrication processes of the microcalorimeter

Releasing the microbridge to suspension using a two-step dry etching. The first step in etching is anisotropic deep reactive ion etching (DRIE), which extends the windows into the substrate vertically to a depth of several tens micrometers, as shown in Fig. 12e. Then isotropic RIE is performed, which etches the substrate laterally and vertically. Lateral etching from the two sides of the microbridge removes the substrate beneath it, and thus suspends the microbridge, as shown in Fig. 12f. Since the silicon device layer is etched away and the resistors remain as silicon islands, the thermal conduction to the substrate is minimized.

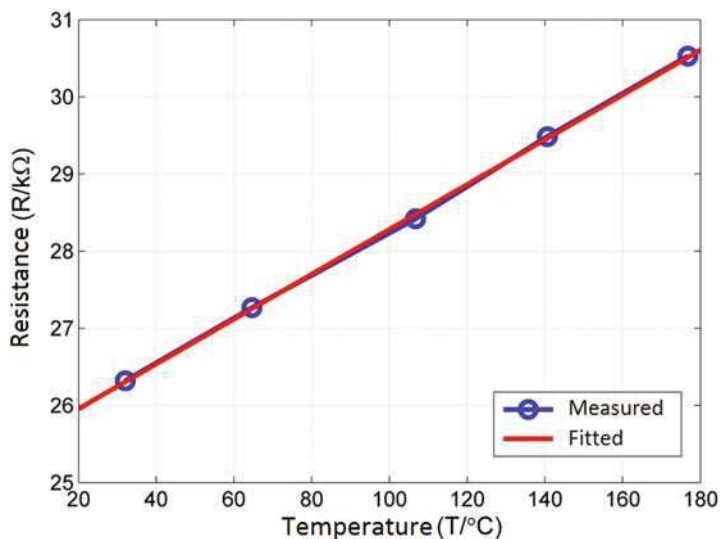
Figure 13 shows the optical photo of a microcalorimeter. The microcalorimeter is suspended over a cavity etched in the substrate. The long and narrow microbridge allows low thermal dissipation to the substrate and large temperature increase by deflagration. It should be noted that uniform temperature distribution across the length of the microbridge contradicts low thermal dissipation to the substrate through the two ends, because uniform temperature distribution also makes the two ends of the microbridge to have high temperature, enhancing the thermal dissipation to the substrate. Figure 14 shows the measured resistance changes versus the temperature. It can be seen that the resistance increases linearly with the temperature. The sensitivity is about 1100 ppm/K.

### ***In Situ Integrated Carbon Nanotubes for Preconcentration***

The operation of the microcalorimeter depends on adsorption of TEC molecules on the surface of the microbridge. For a given adsorption time, the LOD is determined by the amount of the adsorbed TEC, the heat capacity of the microcalorimeter, and the temperature resolution of the thermistor. Thus, it is highly desired to improve the adsorption ability of the microbridge with the heat capacity kept as constant or to

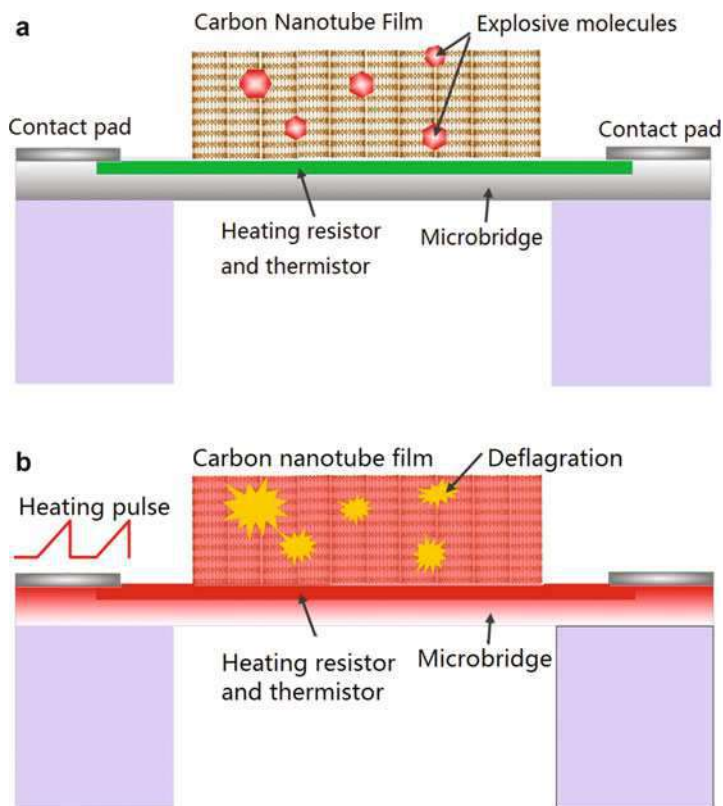


**Fig. 13** Optical photo of a microcalorimeter



**Fig. 14** Measured resistance versus temperature

reduce the heat capacity with the adsorption ability being unaltered. In addition, short adsorption time is also critical to many applications which require fast detection time. However, because of the small surfaces of the microcalorimeter and the low partial pressures of most TECs, it normally takes a rather long time to adsorb sufficient target TEC molecules for detection (Pinnaduwege et al. 2003; Sheehan and Whitman 2005).



**Fig. 15** Microcalorimeter with carbon nanotube films. (a) Adsorption of TEC molecules on the CNT films. (b) Explosion of TEC molecules

Neither the LOD nor the detection time can be improved by increasing the area of the microbridge. Although increasing the area of a microbridge increases the amount of TEC adsorbed proportionally in a given period, it also increases the heat capacity at the same level. Therefore, the extra heat provided by the additional adsorbed TEC is consumed by the additional heat capacity, with the temperature increase caused by TEC deflagration remaining unchanged. To improve the LOD or reduce the detection time, it is needed to improve the ratio of the amount of the adsorbed TEC molecules to the heat capacity of the microcalorimeter by using a material or structure.

One possible solution to improve the LOD is to use a carbon nanotube (CNT) film integrated on the surface of the microbridge. When the microcalorimeter is exposed to TEC vapors, the chemical molecules are adsorbed onto the surfaces of the CNTs, as shown in Fig. 15a. By heating the microbridge to the deflagration temperature of the TEC, the CNT film is heated to almost the same temperature simultaneously due to the extremely high thermal conductivity of CNTs. Thus, the chemical molecules adsorbed on CNT surfaces are ignited to deflagration, and extra heat is generated on

the CNT film and conducted to the microbridge, as shown in Fig. 15b. The extra heat causes an additional temperature increment superimposed on the heating temperature, and by measuring the temperature changes with the thermistor, the endothermic and exothermic characteristics of the thermal reaction are obtained as thermal fingerprints of the TEC.

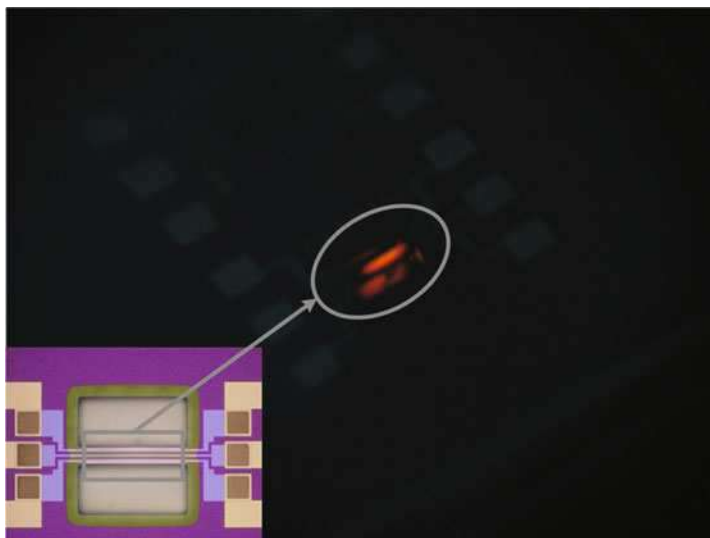
The benefits of using CNT films include two aspects. First, the large surface areas, extremely high thermal conductivity, and high temperature stability of CNTs (Berger et al. 2000; Hiraoka et al. 2010) enable CNTs to be a high-performance adsorbent for chemical vapors. Second, the aromatic structures in nitrobenzene explosives are likely to form weak intermolecular forces between CNTs and TEC molecules through  $\pi$ -stacking interaction (Tournus et al. 2005; Woods et al. 2007), which facilitates efficient and relatively stable adsorption of TEC molecules to CNT surfaces.

To achieve firm CNT integration and high thermal conductivity along the heat-passing routes, direct synthesis of CNTs on the microbridge surfaces by laser heating assisted chemical vapor deposition (CVD) has been developed (Ruan et al. 2011, 2012). Direct synthesis of CNTs avoids the problems of CNT transfer (Patil et al. 2009, Zhou et al. 2012), such as weak adhesion that may cause CNT peeling off during heating, large interface thermal resistance (ITR) between CNTs and the substrate, and possible CNT length reduction and contamination (Timmermans et al. 2011).

For CNT synthesis using laser-assisted CVD, a 3 nm thick Fe film is deposited onto the microbridge as a catalyst. Nitrogen is introduced to purge the reaction chamber thoroughly, and then gases of  $C_2H_2/H_2/Ar$  with an optimized flow rates of 25/20/250 sccm are delivered into the chamber. After flow stability, a  $CO_2$  laser beam is led to the microbridge to heat it to the temperature for CNT synthesis. The laser heats the microbridge to the desired temperature instantly due to the low thermal dissipation to the substrate and the small thermal mass. Because of the high temperature, the microbridge glows during laser heating, as shown in Fig. 16. The cavity under the microbridge is out of focus of the laser beam and the temperature still remains rather low. This allows localized CNT growth on the desired suspended structure by achieving self-confined synthesis without the need of further catalyst patterning. It also prevents the temperature-vulnerable devices on the substrate from being damaged by high temperatures.

Figure 17 shows the SEM photos of the in situ synthesized CNT films on the microbridge. For short CNTs, they resemble sparse grass with a chaos state. For the CNTs longer than 10  $\mu m$ , they tend to form vertically aligned structures due to van der Waals forces. Vertically aligned CNTs with a length of 10  $\mu m$ , which form a densely packed film, are used for adsorption promotion.

One significant benefit of CNT direct synthesis is the minimized interface thermal resistance between CNTs and the microbridge. Because the LOD of the microcalorimeter is strongly dependent on the rates of temperature increasing (Senesac et al. 2009), it is highly desired to reduce the total thermal resistance along the heat transfer route, which arises from the interface thermal resistance and the intrinsic thermal resistance of CNTs. As the thermal conductivity of CNTs is much larger than



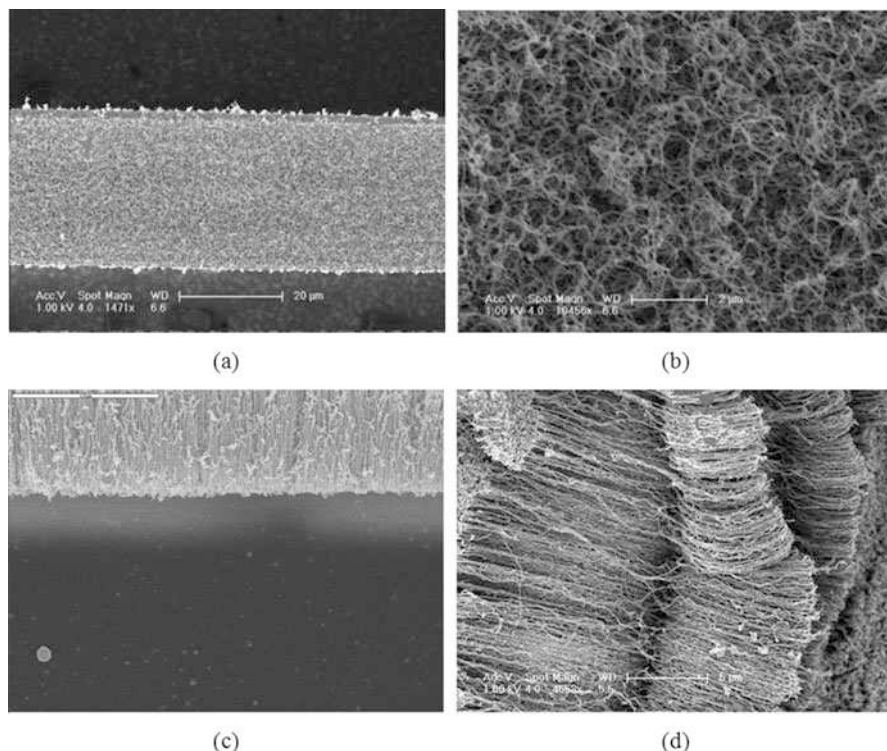
**Fig. 16** Glowing of the microbridge heated by a laser beam for CNT growth

that of the contact, the interface thermal resistance dominates the total thermal resistance. Direct synthesis enables CNTs to be tightly rooted on the substrate surface and the metal catalyst with the assistance of high synthesis temperatures, allowing good phonon interaction between CNTs and substrates to reduce interface thermal resistance. The literature reported interface thermal resistances are on the order of  $10^{-4}$  to  $10^{-5}$   $\text{m}^2\text{K}/\text{W}$  (Panzer et al. 2008; Son et al. 2008), which are much lower than that of the CNTs transferred onto substrates.

Because the heat mass of the CNT film, around 10–12 nJ/K (Ruan et al. 2012), is roughly equal to that of the microbridge, integration of CNTs on the microbridge doubles the heat mass of the whole microcalorimeter. The increase in the heat mass deteriorates the temperature increase caused by TEC deflagration. However, the surface areas are significantly increased because of the large surface to volume ratio of CNT, making the adsorption ability much larger than the heat mass. Because the improvement in adsorption provided by the huge surface areas of CNTs is much higher than the increase in the heat capacity, the overall contribution of the CNTs is significant.

## Detection of TEC Using Microcalorimeters

The large surface areas of CNTs enhance the adsorption ability to chemical molecules. Theoretical works have shown that the chemicals containing aromatic organic molecules tend to be physi-sorbed onto CNTs by non-covalent binding (Tournus et al. 2005; Woods et al. 2007). For TEC detection, two microcalorimeters in two separate chambers are operated in a differential mode. One chamber is prevented

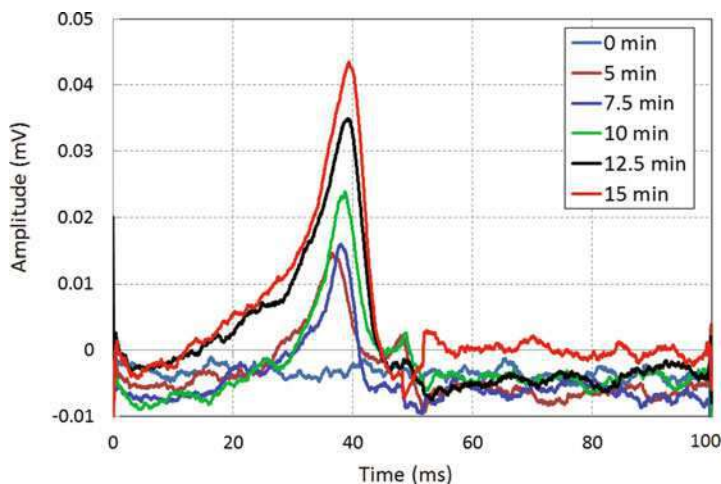


**Fig. 17** In situ synthesized CNT on the microbridge. (a) Sparsely packed CNT film. (b) Close view of sparsely packed CNT. (c) Tightly packed CNT film. (d) Close view of tightly packed CNT

from TEC vapors and the microcalorimeter inside acts as a reference; the other chamber is exposed to TEC vapors and the microcalorimeter inside serves as the sensing device. TNT vapors with saturated concentration at 60 °C, about 1ppmv (Wang 2007), are conducted to the chamber for detection.

The microcalorimeter can be operated at slow heating mode, which uses a 50-ms triangular pulse to heat the microbridge. Figure 18 shows the output voltages of the microcalorimeter as a result of deflagration of the adsorbed TNT molecules versus adsorption times. It is evident that deflagration occurs at about 30 ms, which is indicated by a sharp peak. The voltage peaks increase with adsorption time, indicating increased amount of the adsorbed TNT mass. The response curves have same shapes, which indicate that response curves of a given TEC are determined by the material properties and the heating rate. The slow heating mode is able to recognize the thermal spectra of thermal reaction. However, it significantly desorbs the TNT molecules that are already adsorbed on CNT surfaces because of long time of sublimation and evaporation and thus reduces the temperature increment induced by deflagration.





**Fig. 18** The output voltages of the microcalorimeter using 50-ms heating pulse

The microcalorimeter can be optionally operated at rapid heating mode, which uses a 5-ms pulse to heat the microbridge. This heating rate is fast enough to raise temperature with short time sublimation and evaporation and thus is able to improve the temperature increment. In addition, rapid heating mode also reduces heat dissipation by shortening the duration of high temperatures.

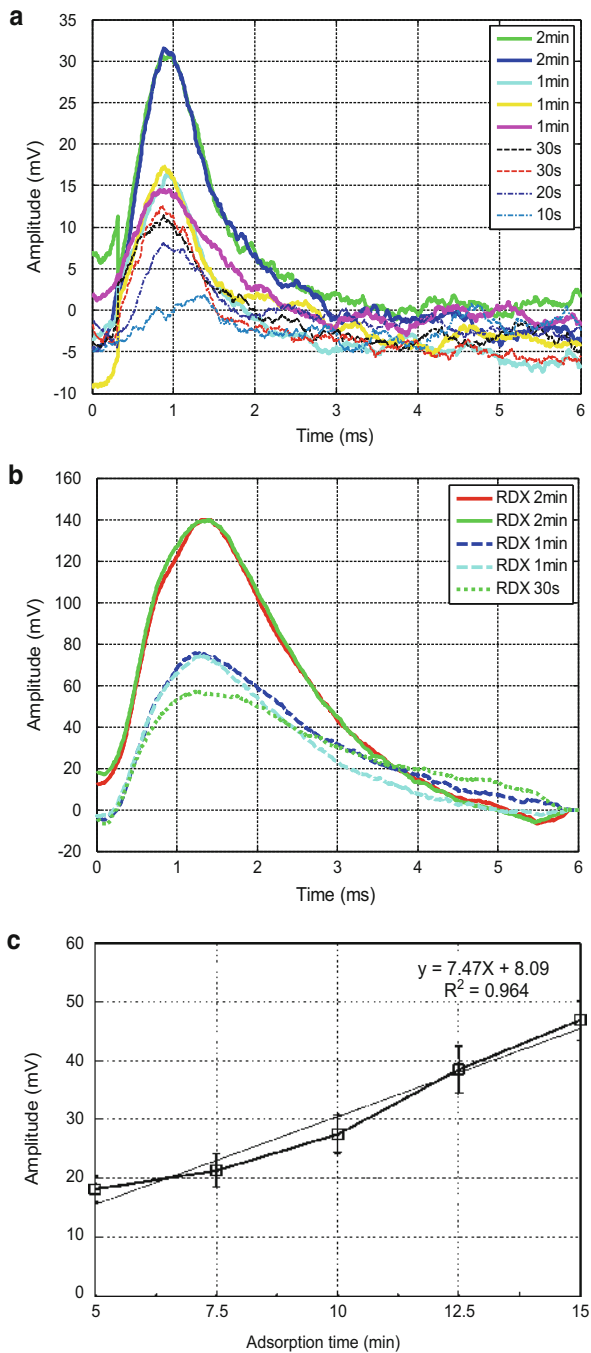
Using a rapid heating rate, the microcalorimeter is operated in a differential thermal analysis mode, which obtains the total heat changes but neglects the thermal reaction details. Figure 19 shows the response voltages of the microcalorimeter versus adsorption times using a rapid heating mode of 5 ms. The equivalent LOD of the microcalorimeter, 2.6 pg, can be theoretically estimated (Ruan et al. 2013).

The saturated vapor pressure of RDX at room temperature is much lower than that of TNT (Shankaran et al. 2005), so it is more difficult to detect trace RDX because it needs much longer adsorption time to accumulate sufficient RDX molecules for detection. In addition, without the intermolecular forces between CNT surfaces and RDX molecules, adsorption of RDX to CNTs could be slower than TNT and desorption during heating could be faster.

Although the microcalorimeter is also used to detect RDX, Fig. 19 also shows the RDX results of evaporation temperature at 84 °C, 2-min adsorption times, and 5-ms rapid heating mode. The peak voltages of 2-min adsorption are around 140 mV, which is more than four times higher than that of TNT for the same adsorption time.

Due to the inherent high-temperature feature of deflagration, the recovery time of the microcalorimeters is very short. The deflagration not only burns the TEC molecules adsorbed on the surfaces of the microcalorimeters but also accelerates evacuation and desorption of impurities and reaction products. The surfaces of the microcalorimeters are cleaned by deflagration and they are ready for next detection immediately after deflagration. Thus, deflagration provides a self-cleaning function to achieve fast recovery times, and the microcalorimeter can be used for TEC vapor

**Fig. 19** The output voltages of the microcalorimeter using 5-ms heating pulse. **(a)** TNT **(b)** RDX **(c)** The changes of the peak voltages vs. the adsorption times





detection reversibly, without the need of waiting or special treatment. The self-cleaning function also prevents the microcalorimeter from being poisoned, which is also a critical problem for chemical sensors as it could cause performance decay after several times of use. The standard deviations of the net voltage changes of the three test runs are 5.1% for fast heating and 6.4% for slow heating, which quantitatively indicate the relatively good repeatability (Ruan et al. 2013).

In practical use, the microcalorimeters may work in complex environments where interferents such as water vapor and other volatile organic chemicals (VOCs) may exist with high concentrations. Because of the deflagration mechanism, water vapor and acetone vapor with pressure of several orders of magnitude higher than that of TNT have no influences on the microcalorimeters (Ruan et al. 2013).

---

## References

- Agah M, Potkay JA, Lambertus G et al (2005) High-performance temperature-programmed micro-fabricated gas chromatography columns. *IEEE J Microelectromech Syst* 14:1039–1050
- Aguilar AD, Forzani ES, Leright M et al (2010) A hybrid nanosensor for TNT vapor detection. *Nano Lett* 10:380–384
- Berger S, Kwon YK, Tomanek D (2000) Unusually high thermal conductivity of carbon nanotubes. *Phys Rev Lett* 84:4613–4616
- Bogue R (2011) Terrorism and military actions pose the ultimate challenge to gas sensing. *Sensors Rev* 31:6–12
- Boukabache A, Pons R (2002) Doping effects on thermal behavior of silicon resistor. *Electron Lett* 38:342–343
- Bruzo MG, Komarov PL, Raad PE (2003) Thermal transport properties of gold-covered thin-film silicon dioxide. *IEEE Trans Compon Packag Technol* 26:80–88
- Camara EHM, Breuil P, Briand D et al (2011) A micro gas preconcentrator with improved performance for pollution monitoring and explosives detection. *Anal Chim Acta* 688:175–182
- Careto-Vazquez VH, Wojcik AK, Liu Y-S et al (2010) Miniaturized calorimeter for thermal screening of energetic materials. *Microelectron J* 41:874–881
- Cavicchi RE, Poirier GE, Tea NH et al (2004) Micro-differential scanning calorimeter for combustible gas sensing. *Sensors Actuators B Chem* 97:22–30
- Chen C-C, Bannister WW, Viswanathan A (1999) Micro-thermal analysis for airport detection of energetic materials. In: *Proceedings the 27th annual conference on North American thermal analysis society*, pp 67–72
- Chen PC, Sukcharoenchoke S, Ryu K et al (2010) 2,4,6-trinitrotoluene (TNT) chemical sensing based on aligned single-walled carbon nanotubes and ZnO nanowires. *Adv Mater* 22:1900–1904
- Ewing RG, Atkinson DA, Eiceman GA et al (2001) A critical review of ion mobility spectrometry for the detection of explosives and explosive related compounds. *Talanta* 54:515–529
- Fainberg A (1992) Explosives detection for aviation security. *Science* 255:1531–1537
- Furton KG, Myers LJ (2001) The scientific foundation and efficacy of the use of canines as chemical detectors for explosives. *Talanta* 54:487–500
- Gershanik AP, Zeiri Y (2010) Sublimation rate of TNT microcrystals in air. *J Phys Chem A* 114:12403–12410
- Goeders KM, Colton JS, Bottomley LA (2008) Microcantilevers: sensing chemical interactions via mechanical motion. *Chem Rev* 108:522–542

- Greve A, Olsen JK, Boisen A et al (2009) Micro-calorimetric sensor for vapour phase explosive detection with optimized heat profile. In: Proceedings IEEE sensors conference on Hawaii, pp 723–726
- Hess CK, Miaoulis IN (1995) Implicit modified enthalpy method with application to thin film melting. *Int J Numer Methods Heat Fluid Flow* 5:385–398
- Hiraoka T, Izadi-Najafabadi A, Yamada T et al (2010) Compact and light supercapacitor electrodes from a surface-only solid by opened carbon nanotubes with 2200 m<sup>2</sup>/g surface area. *Adv Funct Mater* 20:422–428
- Hobbs ML, Kaneshige MJ, Gilbert DW et al (2009) Modeling TNT ignition. *J Phys Chem A* 113:10474–10487
- Hrapovic S, Majid E, Liu Y et al (2006) Metallic nanoparticle-carbon nanotube composites for electrochemical determination of explosive nitroaromatic compounds. *Anal Chem* 78:5504–5512
- Ju YS, Goodson KE (1999) Phonon scattering in silicon films with thickness of order 100 nm. *Appl Phys Lett* 74:3005–3007
- Kleiner MB, Kuhn SA, Weber W (1996) Thermal conductivity measurements of thin silicon dioxide films in integrated circuits. *IEEE Trans Electron Devices* 43:1602–1609
- Lamberg P, Lehtiniemi R, Henell A-M (2004) Numerical and experimental investigation of melting and freezing processes in phase change material storage. *Int J Thermal Sci* 43:277–287
- Lee DS, Shim CH, Lim JW et al (2002) A microsensor array with porous tin oxide thin films and microhotplate dangled by wires in air. *Sensors Actuators B Chem* 83:250–255
- Lienhard JH IV, Lienhard JHV (2005) A heat transfer textbook, 3rd edn. Phlogiston Press, Cambridge, pp 19–26
- Liu YS, Ugaza VM, North SW et al (2007) Development of a miniature calorimeter for identification and detection of explosives and other energetic compounds. *J Hazardous Mat* 142:662–668
- Lu C-J, Whiting J, Sacks RD, Zellers ET (2003) Portable gas chromatograph with tunable retention and sensor array detection for determination of complex vapor mixtures. *Anal Chem* 75:1400–1409
- Moore DS (2004) Instrumentation for trace detection of high explosives. *Rev Sci Instrum* 75:2499–2512
- Mullen C, Irwin A, Pond BV et al (2006) Detection of explosives and explosives-related compounds by single photon laser ionization time-of-flight mass spectrometry. *Anal Chem* 78:3807–3814
- Naddo T, Yang XM, Moore JS et al (2008) Highly responsive fluorescent sensing of explosives taggant with an organic nanofibril film. *Sensors Actuators B Chem* 134:287–291
- Nambayah M, Quickenden TI (2004) A quantitative assessment of chemical techniques for detecting traces of explosives at counter-terrorist portals. *Talanta* 63:461–467
- Panzer MA, Zhang G, Mann D et al (2008) Thermal properties of metal-coated vertically aligned single-wall nanotube arrays. *ASME J Heat Transfer* 130:052401
- Patel SV, Mlsna TE, Fruhberger B et al (2003) Chemicapacitive microsensors for volatile organic compound detection. *Sensors Actuators B Chem* 96:541–553
- Patil N, Lin A, Myers ER et al (2009) Wafer-scale growth and transfer of aligned single-walled carbon nanotubes. *IEEE Trans Nanotech* 8:498–504
- Piazzon N, Rosenthal M, Bondar A et al (2010) Characterization of explosives traces by the nanocalorimetry. *J Phys Chem Solids* 71:114–118
- Piekiel NW, Cavicchib RE, Zachariaha MR (2011) Rapid-heating of energetic materials using a micro-differential, scanning calorimeter. *Thermochim Acta* 521:125–129
- Pinnaduwege LA, Gehl A, Hedden DL et al (2003) A microsensor for trinitrotoluene vapour. *Nature* 425:474
- Pinnaduwege LA, Wig A, Hedden DL et al (2004a) Detection of trinitrotoluene via deflagration on a microcantilever. *J Appl Phys* 95:5871–5875
- Pinnaduwege LA, Yi D, Tian F et al (2004b) Adsorption of trinitrotoluene on uncoated silicon microcantilever surfaces. *Langmuir* 20:2690–2694
- Pinnaduwege LA, Ji HF, Thundat T (2005) Moore's law in homeland defense: an integrated sensor platform based on silicon microcantilevers. *IEEE Sensors J* 5:775–785

- Pinnaduwege LA, Gehl AC, Allman SL et al (2007) Miniature sensor suitable for electronic nose applications. *Rev Sci Instrum* 78:055101
- Roberts ME, LeMieux MC, Bao Z (2009) Sorted and aligned single-walled carbon nanotube networks for transistor-based aqueous chemical sensors. *ACS Nano* 3:3287–3293
- Ruan W, Wang Z, Liu L et al (2011) Synthesis of carbon nanotubes on suspending microstructures by rapid local laser heating. *IEEE Sensors J* 11:3424–3425
- Ruan W, Wang Z, Li Y, Liu L (2012) In-situ heat capacity measurement of carbon nanotubes using suspending microstructure based micro-calorimetry. *IEEE Trans Nanotech* 11:367–373
- Ruan W, Wang Z, Li Y, Liu L (2013) A microcalorimeter integrated with carbon nanotube interface layers for fast detection of trace energetic chemicals. *IEEE J Microelectromechanical Syst* 22(1):152–162
- Senesac LR, Thundat TG (2008) Nanosensors for trace explosive detection. *Mater Today* 11:28–36
- Senesac LR, Yi D, Greve A et al (2009) Micro-differential thermal analysis detection of adsorbed explosive molecules using microfabricated bridges. *Rev Sci Instrum* 80:035102
- Sergio M, Arben M (2012) Nanomaterials based electrochemical sensing applications for safety and security. *Electroanalysis* 24:459–469
- Shankaran DR, Gobi KV, Sakai T et al (2005) Surface plasmon resonance immunosensor for highly sensitive detection of 2,4,6-trinitrotoluene. *Biosens Bioelectron* 20:1750–1756
- Sheehan PE, Whitman LJ (2005) Detection limits for Nanoscale biosensors. *Nano Lett* 5(4):803–807
- Singh S (2007) Sensors-an effective approach for the detection of explosives. *J Hazardous Mater* 144:15–28
- Sinha N, Ma J, Yeow JTW (2006) Carbon nanotube-based sensors. *J Nanosci Nanotechnol* 6:573–590
- Son Y, Pal SK, Borca-Tasciuc T et al (2008) Thermal resistance of the native interface between vertically aligned multiwalled carbon nanotube arrays and their SiO<sub>2</sub>/Si substrate. *J Appl Phys* 103:024911
- Southworth DR, Bellan LM, Linzon Y et al (2010) Stress-based vapor sensing using resonant microbridges. *Appl Phys Lett* 96:163503
- Steinfeld JI, Wormhoudt J (1998) Explosives detection: a challenge for physical chemistry. *Annu Rev Phys Chem* 49:203–232
- Sun D, Garimella SV (2005) Numerical and experimental investigation of the melt casting of explosives. *Propellants Explos Pyrotech* 30:369–380
- Sun Y, Liu K, Miao J et al (2010) Highly sensitive surface-enhanced raman scattering substrate made from super-aligned carbon nanotubes. *Nano Lett* 10:1747–1753
- Timmermans YM, Grigoras K, Nasibulin AG et al (2011) Lithography-free fabrication of carbon nanotube network transistors. *Nanotechnol* 22:065303
- Tournus T, Latil S, Heggie MI et al (2005) Pi-stacking interaction between carbon nanotubes and organic molecules. *Phy Rev B* 72:075431
- Wang J (2004) Microchip devices for detecting terrorist weapons. *Anal Chim Acta* 507:3–10
- Wang J (2007) Electrochemical sensing of explosives. *Electroanalysis* 19:415–423
- Wang J, Thongngamdee S (2003) On-line electrochemical monitoring of (TNT) 2,4,6-trinitrotoluene in natural waters. *Anal Chim Acta* 485:139–144
- Wang Z, Yue R, Zhang R, Liu L (2005) Design and optimization of laminated piezoresistive microcantilever sensors. *Sensors Actuators A* 120(2):325–336
- Wang L, Wang B, Lin Q (2008) Demonstration of MEMS-based differential scanning calorimetry for determining thermodynamic properties of biomolecules. *Sensors Actuators B Chem* 134:953–958
- Woods LM, Badescu SC, Reinecke TL (2007) Adsorption of simple benzene derivatives on carbon nanotubes. *Phys Rev B* 75:155415
- Xu P, Li X, Yu H et al (2010) Self-assembly and sensing-group graft of pre-modified CNTs on resonant micro-cantilevers for specific detection of volatile organic compound vapors. *J Micro-mech Microeng* 20:115003

- Yamane T, Nagai N, Katayama S, Todoki M (2002) Measurement of thermal conductivity of silicon dioxide thin films using a  $3\omega$  method. *Appl Phys Lett* 91:9772–9776
- Yi D, Greve A, Hales JH et al (2008) Detection of adsorbed explosive molecules using thermal response of suspended microfabricated bridges. *Appl Phys Lett* 93:154102
- Youssef S, Podlecki J, Al Asmar R et al (2009) MEMS scanning calorimeter with serpentine-shaped platinum resistors for characterizations of microsamples. *IEEE J Microelectromech Syst* 18:414–423
- Yu X, Tang Q, Zhang H et al (2007) Design of high sensitivity cantilever and its monolithic integration with CMOS circuits. *IEEE Sensors J* 7:489–495
- Zhou Y, Wang Z, Wang C et al (2009) Design, fabrication and characterization of a two-step released silicon dioxide piezoresistive microcantilever immunosensor. *J Micromech Microeng* 19:065026
- Zhou Y, Johnson JL, Ural A, Xie H (2012) Localized growth of carbon nanotubes on CMOS substrate at room temperature using maskless post-CMOS processing. *IEEE Trans Nanotechnol* 11:16–20
- Zribi A, Knobloch A, Rao A (2005) CO<sub>2</sub> detection using carbon nanotube networks and micro-machined resonant transducers. *Appl Phys Lett* 86:203112
- Zuck A, Greenblatt J, Zifman A et al (2008) Explosive detection by microthermal analysis. *J Energetic Mat* 26:163–180



# Biochemical Sensors Based on Piezoresistive Microcantilevers

Xiaomei Yu and Rui Zhao

## Contents

Introduction .....	690
Microcantilever .....	690
Readout Methods .....	692
Design Approach .....	693
Sensitivity .....	693
Noise .....	693
Resolution .....	694
Design Method .....	695
Fabrication .....	696
Fabrication Process with Backside Releasing Technique .....	697
Fabrication Process with Frontside Releasing Technique .....	698
Characteristics of Piezoresistive Micrcantilever .....	700
Spring Constant .....	700
Noises .....	702
Sensitivity .....	703
Monolithic Integration of Piezoresistive Microcantilever with Readout Circuts (Xiaomei Yu et al. 2008) .....	704
Chemical Sensors .....	708
Biosensors .....	712
Summary .....	712
References .....	714

---

X. Yu (✉)

National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Haidian District, Beijing, China  
e-mail: [yuxm@pku.edu.cn](mailto:yuxm@pku.edu.cn)

R. Zhao

Science and Technology on Electronic Test and Measurement Laboratory, School of Instrument and Electronics, North University of China, Taiyuan, Shanxi, China  
e-mail: [zhaorui@nuc.edu.cn](mailto:zhaorui@nuc.edu.cn)

---

**Abstract**

Microcantilevers, which are highly attractive for their small size, high sensitivity, and low cost, have been successfully used in label-free biological and chemical sensing applications during the past 20 years. In this chapter, a piezoresistive microcantilever-based biochemical sensor is introduced, in which a mechanical bending induced by a biochemical reaction or absorption on the surface of the microcantilever is changed into an electrical signal by integrated piezoresistors. Theory and design method are introduced firstly, and then fabrication technique and characteristics of the microcantilever sensors are described in details. Finally, we introduce some biochemical detection results measured with the piezoresistive microcantilever-based sensors.

---

**Keywords**

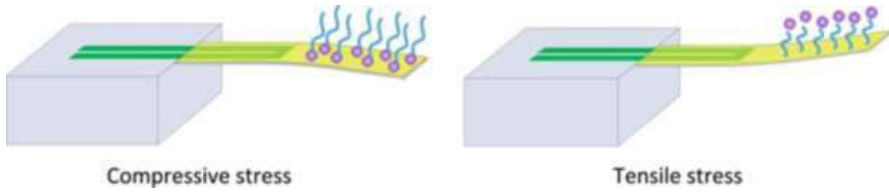
Biochemical sensors · Piezoresistive microcantilever · Deflection sensitivity · Minimum detectable deflection · Monolithic integration

---

**Introduction****Microcantilever**

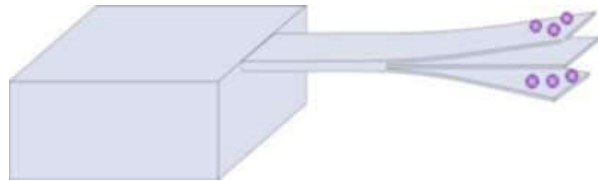
The background for using microcantilever-based sensors is found in the atomic force microscope (AFM) invented by Binnig et al. in 1986 (Binnig et al. 1986). Changes in surface stress due to adsorption (of molecules) on cantilevers was reported by Raiteri and Butt (1995), Chen et al. (1995), Butt (1996), and O'Shea et al. (1996), and this showed the potential use of microcantilevers developed for AFM as biochemical sensors. The adsorption of biochemical molecules and the molecular recognition landed on the microcantilever surface may generate changes in the bendings or resonance frequency shifts of the microcantilever, as shown in Figs. 1 and 2. The detections of these responses are referred as surface stress and effective Young's modulus of operation modes, respectively.

The surface stress-based operation mode is also referred as static modes. Working in this mode, the surface stress change comes from molecular interaction and must take place over only one of the microcantilever surface, which generally leads to the changes of the curvature. The induced curvature may give rise to positive or negative surface stress, depending on the surface deformation generated. The responsible phenomena for this change is not fully understood due to the complex equilibrium mechanism of sensing layer, the molecules of surrounding liquid phase, and the probe and target molecules. However, it has been demonstrated that some factors play important roles in the deformation of microcantilevers, including the interactions generated by hydration; electrostatic, Van der Waals, and steric forces; changes in the surface hydrophilia; or conformational changes of the adsorbed molecules (Alvarez and Lechuga 2010; Fritz 2008). The most common model to study microcantilevers deflections to surface stress is based on the work of



**Fig. 1** Schematic diagram of a microcantilever deformation induced by surface stress

**Fig. 2** Dynamic operation of a microcantilever



G. Stoney in 1909 (Lavrik et al. 2004; Stoney 1909), where the relation of the difference in surface stress  $\Delta\sigma$  between the top and bottom sides of a microcantilever; the microcantilever deflection,  $\Delta z$ ; Young’s modulus,  $E$ ; the Poisson coefficient,  $\nu$ ; and the dimensions of the microcantilever, the length  $L$  and the thickness  $t$ , can be expressed by

$$\Delta\sigma = \frac{E \cdot t^2}{3 \cdot (1 - \nu) \cdot L^2} \cdot \Delta z \tag{1}$$

For the biochemical sensing in the static mode, there is only one surface of the microcantilever to be functionalized, especially employing microcantilever arrays to perform the detection task.

In contrast to the case of static mode, the dynamic mode relies on binding-induced changes in cantilever frequency shift caused by stiffness or mass change. So the surface biochemical sensing is directly measured because the shift frequency is a function of the total mass adsorbed on the both surfaces of the microcantilever. Assuming an effective adsorbed mass  $m_0$  and a spring constant  $k$ , the fundamental resonance frequency of microcantilever,  $f_0$ , in the absence of damping can be approximated as

$$f_0 = \frac{1}{2\pi} \cdot \sqrt{\frac{k}{m_0}} \tag{2}$$

When working in the dynamic mode, the frequency shift of the microcantilever,  $\Delta f$ , is determined by the quality factor  $Q$  that quantifies the energy dissipation and is often defined as the relation between the dissipated per vibration cycle and the mechanical energy accumulated. Under liquid phases, the quality factor shifts  $Q$  toward much lower values than in vacuum or air, which decreases sharply the

overall sensitivity. Thus, the dynamic mode of operation of microcantilever is difficult to implement the biochemical detection in liquid phases. Conversely, the environmental medium is not a limiting factor that can be strongly influenced, and most of the microcantilever-based biochemical sensors are based on the surface stress mode.

## Readout Methods

The readout system capable of monitoring the nanomechanical motion of microcantilever induced by biochemical reaction or absorption is crucial for biochemical sensing applications. The most extend readout operations are optical and piezoresistive methods. The optical methods used to measure the microcantilever deflections include optical beam deflection and optical interferometry (Koev et al. 2010; Kosaka et al. 2014). The optical readout system consists of a laser and a deflection sensor. The laser beam is focused on the backside of the cantilever and reflected onto a position sensitive laser diode. When the cantilever changes its deflection, the laser beam is reflected onto a different spot on the deflection sensor (Thaysen 2001). Optical method shows an ultrasensitive and linear response with subangstrom resolution. However, this readout technique requires external equipment to measure the deflection of microcantilever, so that its calibration is very time consuming. In addition, the bulk and stiff price of optical instrument also limit the onsite detections of biochemical agents.

To overcome these problems, some researchers have employed microcantilever with piezoresistors integrated in it to perform biochemical sensing. Typical piezoresistors are fabricated by a boron-doped single-crystal silicon or polysilicon. When a microcantilever with a silicon piezoresistor placed at appropriate region is deformed, the change in the piezoresistor reflects the extent of the deformation. The change in piezoresistance can be measured by containing microcantilever into a Wheatstone bridge with a DC bias voltage. When a DC voltage is supplied over the Wheatstone bridge, a differential output of the Wheatstone bridge can be measured.

Compared to the optical readout method, piezoresistive readout also has several advantages:

- (a) Piezoresistive microcantilever can work in opaque liquid phase or gas phase environments.
- (b) Technical expertise for operating the instruments and alignment of the laser system is not needed.
- (c) They are compatible with high-throughput array and miniaturization and thereby obtaining a candidate as a point-of-care devices.

For all these reasons, the piezoresistive microcantilevers are employed more to perform the onsite biochemical detection (Carrascosa et al. 2006; Tamayo et al. 2013).



The main disadvantage of piezoresistive readout is the intrinsic noise that directly affects the signal-to-noise rate (SNR) comparing to the optical readout. Aiming to decrease the intrinsic noise and improve the SNR for the resolution and sensitivity, the optimizations of piezoresistive microcantilevers need to be carried out, including optimizing the design and the fabrication process of the device, which can make the piezoresistive microcantilever-based sensor achieves lower detection limit.

## Design Approach

### Sensitivity

For a piezoresistive microcantilever, relation between the relative resistance change and the surface stress suggest clearly that increasing the stress level will increase the relative resistance change, and therefore the sensitivity of the cantilever-based sensor. The deflection sensitivity is defined as a ratio of cantilever resistance change ( $\Delta R/R$ ) to the end deflection of the cantilever ( $\Delta z$ ). For a rectangular cantilever, the deflection sensitivity of a piezoresistive cantilever is described by the following expression (Yang et al. 2003a)

$$\frac{\Delta R/R}{\Delta z \Pi} = \frac{\sigma_1 - \sigma_2}{\Delta z} \quad (3)$$

where  $\Pi = \beta \cdot (3\pi_L(1 - \nu)/t)$  is the piezoresistive coefficient,  $\pi_L$  is the longitudinal piezoresistive coefficient of silicon,  $\sigma_1$  and  $\sigma_2$  are the longitudinal stress and transverse stress, respectively,  $t$  is the thickness of microcantilever,  $\nu$  is the Poisson ratio, and  $\beta$  is a factor that adjusts for the thickness of the piezoresistor. When the cantilever is bent by a force  $F$  acting on the end, and assuming that the cantilever is deflected only by the longitudinal stress and the cantilever resistor located at maximum stress area, the deflection sensitivity can be expressed by Yu et al. (2002):

$$\left(\frac{\Delta R}{R}\right)/\Delta z = \frac{3\pi Et(L - l/2)}{2L^3} = K \cdot \frac{3t(L - l/2)}{2L^3} \quad (4)$$

where  $K = E\pi$  denotes the gauge factor of the piezoresistor, and  $L$  and  $l$  are the length of microcantilever and piezoresistor. This equation shows that the deflection sensitivity is only dependent on the thickness of the layers, gauge factor of the piezoresistor, and the length of cantilever and piezoresistor.

### Noise

For a piezoresistive microcantilever, one important performance parameter is noise. There are two main types of intrinsic noises affecting the resolution of cantilever-based sensors (Yu et al. 2002), which are frequency-independent Johnson noise and

low-frequency  $1/f$  noise. The Johnson spectral power noise  $S_{VJ}$  for a resistance  $R$  is defined as

$$S_{VJ} = 4k_B \cdot T \cdot R \quad (5)$$

where  $k_B$  is the Boltzmann constant and  $T$  is temperature. The  $1/f$  noise in homogeneous materials is dependent on the total number of carriers in a resistor, which was put forth in 1969 by Hooge (1969). Thus, the  $1/f$  noise is described by an empirical formula (Hooge et al. 1981)

$$S_{VH} = \frac{\alpha \cdot V^2}{f \cdot N} \quad (6)$$

where  $S_{VH}$  is the spectral power noise density corresponding to voltage fluctuation,  $V$  is the biased voltage across a piezoresistor with a total number of carriers  $N$ ,  $f$  is frequency, and  $\alpha$  is a Hooge factor but a device dimension-independent parameter whose value lies in the range of  $10^{-7}$  to  $10^{-3}$  (Vandamme 1983).  $(S_{VH}f)^{1/2}$  is used to indicate the  $1/f$  voltage noise power, which can be written as Yu et al. (2002)

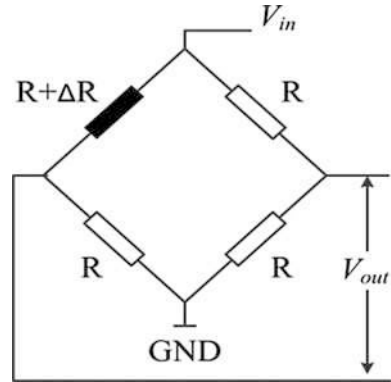
$$(S_{VH}f)^{1/2} = \left( \frac{\alpha}{plwt_p} \right)^{1/2} V_{in} \quad (7)$$

where  $p$  is the carrier's concentration, and  $l$ ,  $w$ , and  $t_p$  are the piezoresistor length, width, and thickness, respectively. The Hooge factor then can be calculated with the linear relation between  $S_{VH}f$  and surface area of piezoresistors. Unlike the Johnson noise source mentioned above, which is well understood, the origin of  $1/f$  noise is still an active research area. Many authors have proven by their experiments that  $1/f$  noise is a fluctuation of the mobility. The mobility fluctuation has been well interpreted in terms of the lattice scattering model, whereas the impurity scattering has no appreciable contribution to the noise (Chen et al. 1999; De Graaff and Huybers 1983; Hooge 1994; Hooge et al. 1981; Vandamme 1983; Vandamme and Oosterhoff 1986).

## Resolution

The minimum detectable deflection (MDD), which represents the resolution of a cantilever, is determined by the noise level and the sensitivity. According to the working principle of the Wheatstone bridge and deflection sensitivity of a cantilever, the MDD is calculated by replacing the output voltage signal of the Wheatstone bridge with the total noise. For a rectangular piezoresistive microcantilever with a Wheatstone bridge configuration outlined in Fig. 3, if one of the four piezoresistors is acted on by a force  $F$ , the relation between the resistance change  $\Delta R$  and the output signal  $V_{out}$  is given by

**Fig. 3** Schematic diagram of arrangement of four piezoresistive microcantilevers as a Wheatstone bridge circuit for signal read out



$$\frac{\Delta R}{4R} = \frac{V_{out}}{V_{bias}} \tag{8}$$

where  $V_{bias}$  is the bias voltage, and the factor 1/4 appears because only one of the four piezoresistors undergoes a stress change. The noise from the piezoresistors will limit the detection of the output signals. When  $V_{out}$  is substituted by the total noise, the MDD ( $z_{min}$ ) of the microcantilever at a given measured bandwidth ( $f_{max} - f_{min}$ ) can finally be expressed as

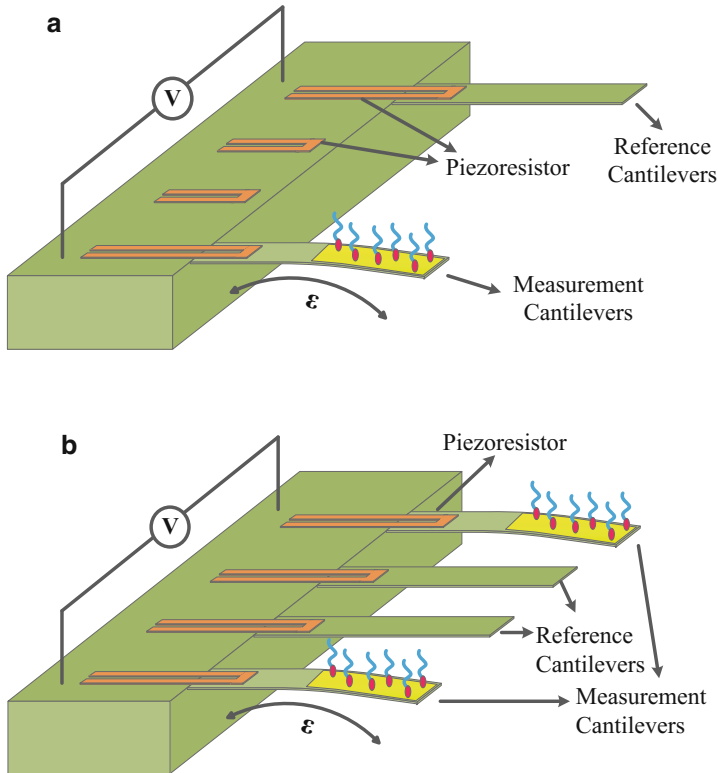
$$z_{min} = \frac{4}{V_{bias}} \left[ \frac{\alpha V^2}{N} \ln \frac{f_{max}}{f_{min}} + 4k_B T R (f_{max} - f_{min}) \right]^{1/2} \cdot \left[ K \frac{3(L-l/2)t}{2L^3} \right]^{-1} \tag{9}$$

### Design Method

An electrical readout system of piezoresistive cantilevers usually designed in a Wheatstone bridge form, where a change in resistance is transformed into a voltage signal. One of the configurations is an onchip Wheatstone bridge with two of the resistors placed on cantilevers and two resistors placed on the substrate (Fig. 4a). In this mode, one of the cantilevers is then used as a measurement cantilever, while the other is used as a reference. During measurement, the two cantilevers are exposed to the same physical environment; therefore changes in the resistance due to changes in the physical environment are then cancelled in the Wheatstone bridge (Thaysen 2001).

Another design is that four identical cantilevers with four piezoresistors embedded form a Wheatstone bridge, where two of them are sensing cantilevers and the other two are reference cantilevers (Fig. 4b). Due to the identical mechanical structure, this design can suppress and counteract the mechanical and external environmental noise effectively.

The optimal design with regard to the dimensions of the cantilever and piezoresistor is found by minimizing the minimum detectable deflection. For the

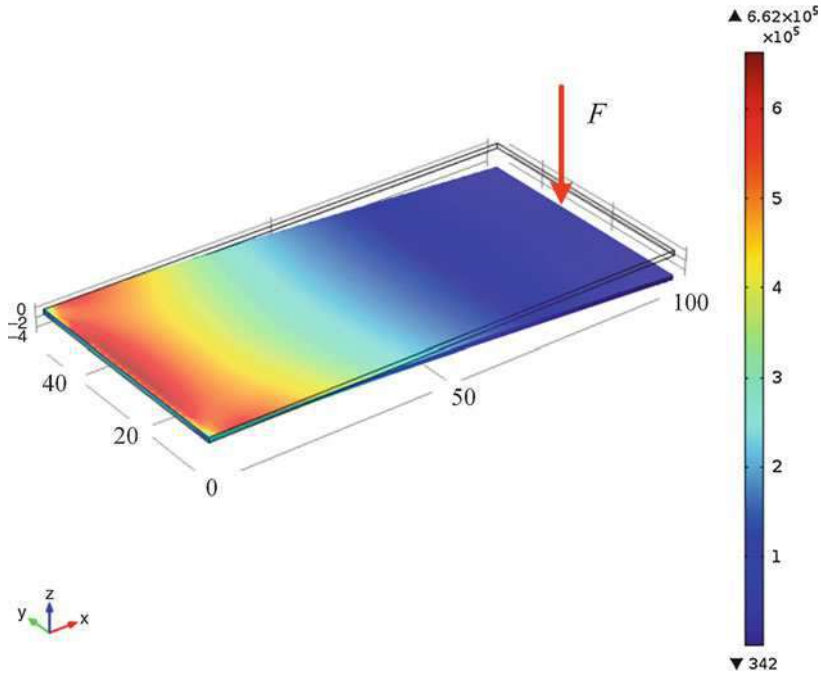


**Fig. 4** Schematic diagram of two Wheatstone bridges

biochemical application of piezoresistive microcantilever, the typical length and width are 100~500  $\mu\text{m}$  and 30~70  $\mu\text{m}$  respectively, and the thickness is usually thinner than 1  $\mu\text{m}$ , so as to obtain sufficient surface area. It is seen from the expression of the deflection sensitivity in Eq. 4 that the sensitivity is proportional to the gauge factor  $K$ . It is, therefore, an advantage to obtain a gauge factor as high as possible in order to maximize the sensitivity. The highest gauge factor is obtained from single crystalline silicon. Therefore, silicon piezoresistors are often used as transducers in micromechanical components. It can be seen from the finite element simulation of Fig. 5 that the surface stress concentrates at one-third of the cantilever root, therefore the piezoresistors should be embedded in this region.

## Fabrication

Piezoresistive microcantilevers are usually fabricated on a silicon-on-insulator (SOI) wafer using microfabrication process since thinner single crystalline silicon resistor can be obtained from the device layer of the SOI wafer and the buried layer can serve



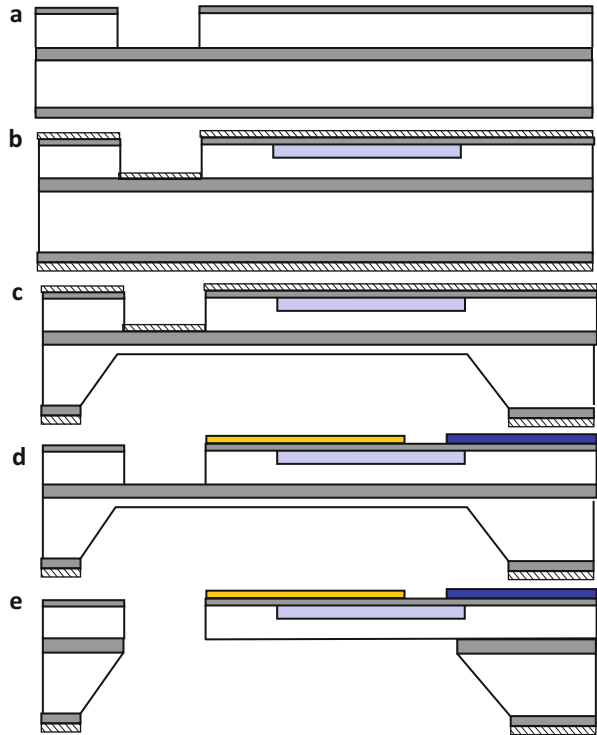
**Fig. 5** Finite element simulation results of a cantilever with a force  $F$  applied at its end

as a bottom protective layer of the piezoresistors. The microfabrication process includes thin layer deposition, photolithographic pattern, wet and dry etching, bulk and surface micromachining, etc. Herein, two micromachining processes are introduced in detail, and the main difference lies in the releasing methods.

### Fabrication Process with Backside Releasing Technique

Fabrication processes for the cantilevers with backside releasing technique is shown in Fig. 6. The process starts from a SOI wafer using a series of front side definition and backside wet/dry etching. The device layer thickness of the SOI wafer is p-type in  $\langle 100 \rangle$  orientation, which controls the piezoresistors type. First a thinner layer of silicon oxide is grown on the wafer, which is then patterned with a cantilever mask and defined under RIE system (Fig. 6a). After the etching of top silicon layer, boron ion implantation with energy of 35 keV and a dosage of  $2 \times 10^{14} \text{ cm}^{-2}$  for the piezoresistor followed by a heavy doping of  $5 \times 10^{15} \text{ cm}^{-2}$  for the contact pad are performed. Electrical activation is achieved using rapid thermal annealing (RTA) at 1000 °C for 30 s followed by low-pressure chemical vapor deposition (LPCVD) silicon nitride, which is used as a mask layer for backside KOH deep etching (Fig. 6b).

**Fig. 6** Fabrication processes for the cantilevers with backside releasing technique

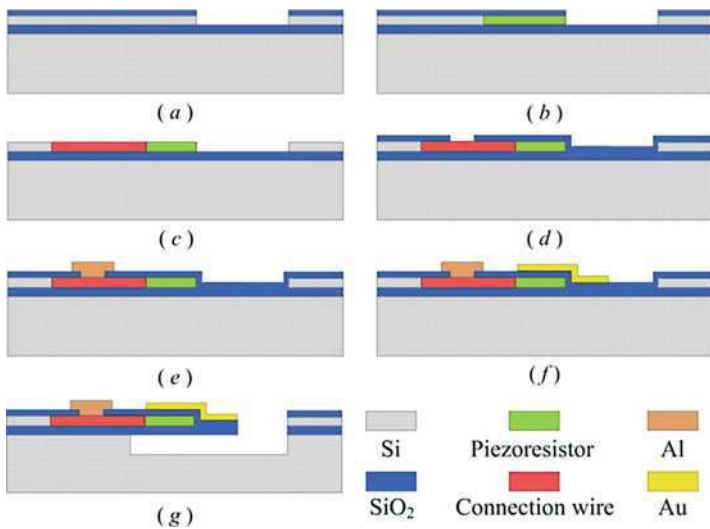
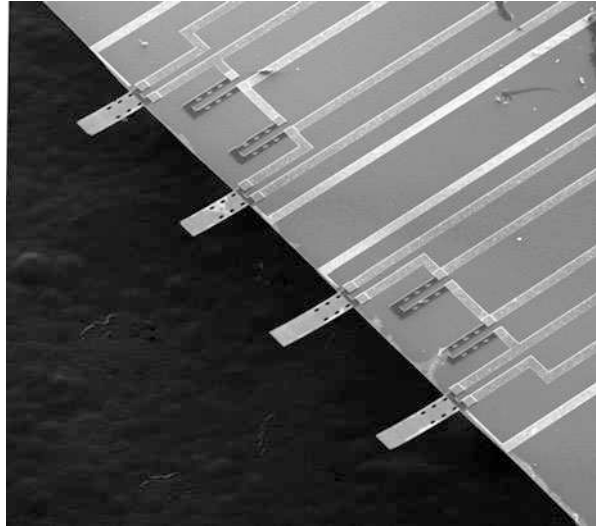


Since the cantilevers are to be used as biosensor, two metal layers need to be deposited on the cantilever top surface. One is 300 nm Al metal wire which is used as the connection of piezoresistor, the other is a 50 nm Au layer that is deposited for the immobilization of biomolecule on the cantilever surface. Considering the influence on the two metals, the silicon backside etching in KOH liquid is performed beforehand, and the etching is continued until about a 20  $\mu\text{m}$  thick handle silicon is left (Fig. 6c). After contact holes are opened, a 300 nm thick Ti/Al layer is first deposited by sputtering and patterned to obtain electrical contact for the resistors. Then 50 nm Cr/Au is deposited and patterned with wet etching technology (Fig. 6d). Finally, the backside is subjected to DRIE dry etching to remove the left 20  $\mu\text{m}$  silicon and followed by a silicon oxide etching in RIE system to release the cantilever (Fig. 6e). Figure 7 is an example of SEM photo of a cantilever array fabricated by the backside realizing technique.

### Fabrication Process with Frontside Releasing Technique

The wet etching releasing technique can provide low cost, high yield, and repeatability of the fabrication process, but the contamination of potassium ions caused by KOH wet etching will give rise to the unsteadiness of output signal for the piezoresistive microcantilevers. Meanwhile, the fabrication reliability is greatly

**Fig. 7** SEM photo of a cantilever array fabricated by backside realizing technique



**Fig. 8** Fabrication process of the piezoresistive microcantilever based on dry etching releasing technique

influenced due to the presence of punch-through structure. To overcome these negative elements, a dry etching method is employed to release the microcantilever, whose fabrication process is sketched in Fig. 8 and described as follows:

- (a) The substrate is first oxidized to generate a 30 nm SiO<sub>2</sub> layer on the top of device layer as the doping protective layer by dry oxidization and then inversely doped

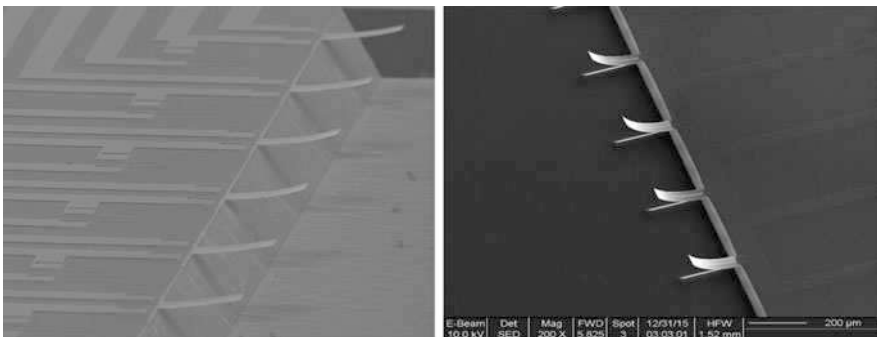
phosphorus ions with a dosage of  $5 \times 10^{13} \text{ cm}^{-2}$  at 100 keV to generate an  $\text{N}^+$  inversion layer. After that, active areas are patterned by reactive ion etching (RIE) technique.

- (b) To form the piezoresistors, a boron ion implantation is processed with a dosage of  $4 \times 10^{14} \text{ cm}^{-2}$  and an energy of 50 keV. The corresponding resistance of piezoresistor is 7 k $\Omega$  for the designed dimension.
- (c) The electrical wiring between piezoresistors and pads are realized by a boron heavy doped device layer with a dosage of  $8 \times 10^{15} \text{ cm}^{-2}$ . Then the  $\text{SiO}_2$  protective layer is removed in buffered hydrofluoric acid (BHF) solution.
- (d) After depositing a 300 nm  $\text{SiO}_2$  layer as a passivation layer by low pressure chemical vapor deposition (LPCVD) to fully encapsulate the piezoresistors and electrical wires together with the buried oxide layer, the contact holes are opened by BHF wet etching of the  $\text{SiO}_2$  layer.
- (e) An 800 nm aluminum layer is sputtered and patterned, which are then annealed at 470 °C for 30 min in nitrogen environment to form an ohmic contact.
- (f) The Cr/Au modified layer with the thickness of 10/50 nm are sequentially sputtered by physical vapor deposition (PVD) and patterned on the top surface of the sensing cantilevers by a lift-off process.
- (g) After the cantilevers are patterned by a hybrid etching of  $\text{SiO}_2$  layers, which include 300 nm passivation layer and 400 nm thick buried oxide layer, with BHF and RIE, a hybrid process of  $\text{SF}_6$  anisotropic and isotropic dry etching is adopted to release the cantilevers with a 4  $\mu\text{m}$  thick photoresist as a mask. Figure 9 shows the SEM photos of cantilever array fabricated by the frontside realizing process.

## Characteristics of Piezoresistive Micrcantilever

### Spring Constant

The theoretical spring constant for the rectangular cantilevers can be calculated by the following equation (Yu et al. 2000)



**Fig. 9** SEM photos of cantilever array fabricated by frontside realizing technique



$$k = \frac{E \cdot W \cdot t^3}{4 \cdot L^3} \tag{10}$$

where  $W$  is the width of the cantilever. As we know, the geometries of the microcantilever and Young’s modulus have appreciable contributions to the spring constant and the fabrication process may cause inaccuracy for the prior calculation of the spring constant. To measure the spring constant of the tested microcantilever, a calibration microcantilever is applied by pressing the end of tested cantilever with a transmitted force. The spring constant  $k_{cal}$  of the calibration microcantilever should be in the same order of magnitude with the tested microcantilever, and the measured transmitted force is given as (Gates et al. 2011; Pandya et al. 2014)

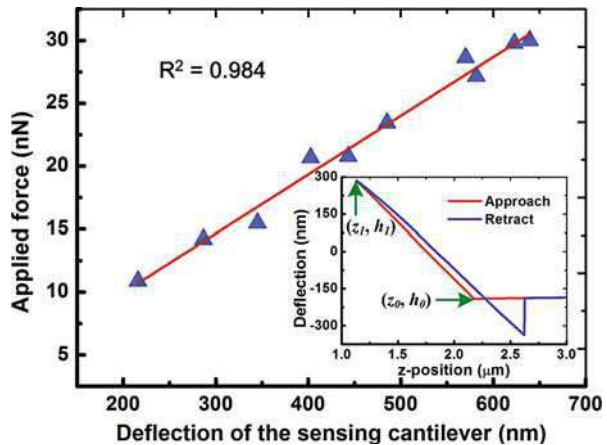
$$F = k_{cal} \cdot (h_1 - h_0) \tag{11}$$

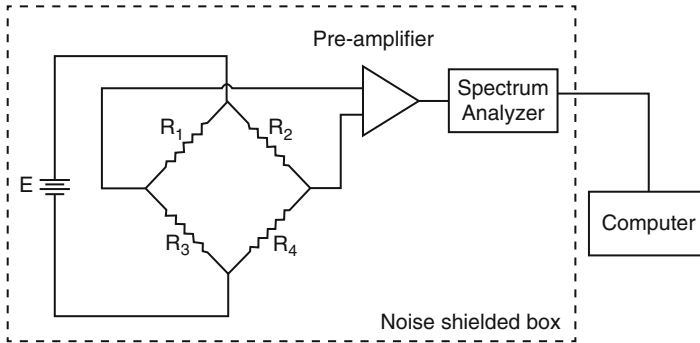
where  $h_0$  and  $h_1$  are the initial and final deflections of the calibration microcantilever detected by a photodiode, respectively. The deflection of the tested cantilever,  $d$ , is calculated by the difference of the net deflection of the calibration microcantilever and the total  $z$ -travel, as given below (Pillarisetti et al. 2011)

$$d = |z_1 - z_0| - |h_1 - h_0| \tag{12}$$

The typical measured result is shown in Fig. 10, which plots the relation between the transmitted force,  $F$ , and the deflection,  $d$ , of the tested microcantilever. The calibrated spring constant of the tested cantilever is extracted by the slope of the linear fit of  $F$  and  $d$ .

**Fig. 10** Applied force versus deflection for measuring the spring constant of the sensing cantilever. Inset plots AFM force curves on the sensing cantilever of the fabricated sensor. The approach curve is used for calibration process





**Fig. 11** Simplified diagram of noise measurement setup

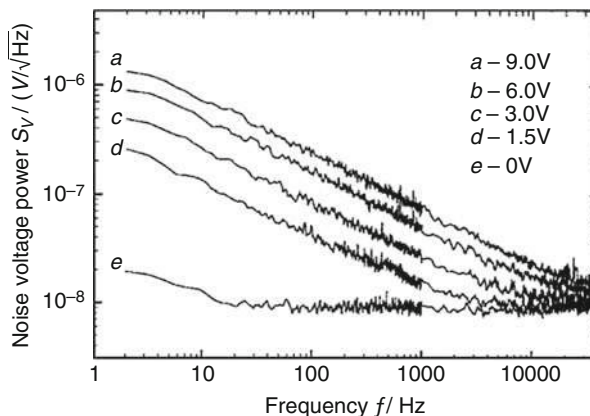
## Noises

The measurements of the noises can be carried out at room temperature using a spectrum analyzer with different DC bias voltages. Since the noise from the spectrum analyzer is a factor of ten times higher than the noise of the piezoresistors, the output signals from the Wheatstone bridge must be amplified by a hundred or a thousand times with a low-noise amplifier (Yu et al. 2001). Figure 11 shows a schematic diagram of the measurement setup. The noise now measured at the output is a summation of many contributions from different sources. The external noise must be reduced as much as possible. Firstly, the microcantilever is shielded from the measurement instruments. Then all the measurement instruments except the computer are enclosed inside a noise-shielded box. After carefully grounding and shielding, the external noises can be minimized significantly. Selecting the right measurement band width, 2~1000 Hz and 1~30 kHz, is also important for obtaining the ideal noise spectra. During the measurement, all the cables should be as short as possible.

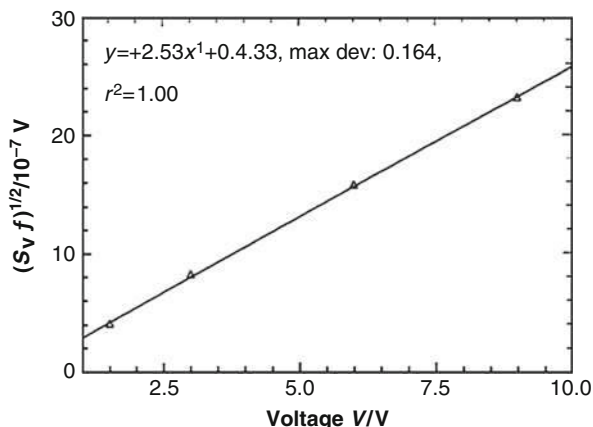
Figure 12 shows a group of typical noise spectra of polysilicon microcantilever measured at different DC voltages, with a resistance of 6.25 k $\Omega$ . It can be seen clearly that the noise at high frequencies is Johnson noise, which was measured to be  $1 \times 10^8$  V/Hz<sup>1/2</sup> at 0 V bias voltage. This measurement result is in fairly good agreement with the theoretical Johnson noise level. As the bias voltage is increased, the Johnson noise becomes a little higher than that at 0 V voltage, and the  $1/f$  noise and the corner frequency (the intersection frequency of Johnson noise and  $1/f$  noise) increase at the same time. The low-frequency noise spectra have been fitted, and a  $-0.5$  slope was obtained for every noise spectrum. Thus, the conclusion is that the dominant noise at low frequencies is  $1/f$  noise. The relations of  $(S_{VH}, f)^{1/2}$  versus voltage is plotted in Fig. 13. The Hooge factor  $\alpha$  can be calculated using the slope of the straight line of  $(S_{VH}, f)^{1/2}$  versus voltage.

Figure 14 shows a group of typical noise spectra of single crystal silicon piezoresistive microcantilever at bias DC voltages of 0, 3, 6, and 9 V, with a resistance of 7 k $\Omega$ . The  $1/f$  noises increase as the bias voltage increased, and the

**Fig. 12** Noise spectra of a polysilicon piezoresistive microcantilever measured at 0, 1.5, 3.0, 6.0, and 9.0 V bias voltages



**Fig. 13** Dependence of  $(S_V f)^{1/2}$  on bias voltages

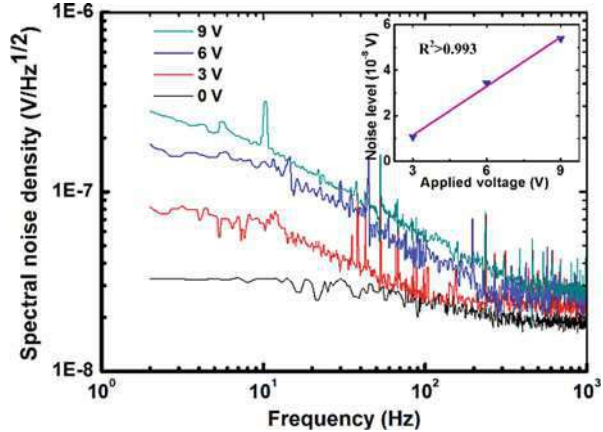


Hooge factor  $\alpha$  can be found by fitting the noise level  $(S_{VH}f)^{1/2}$  to the bias voltage in the inset of Fig. 14. It can be found that the Hooge factor of single crystal silicon microcantilever is two to three orders lower than that of polysilicon microcantilever.

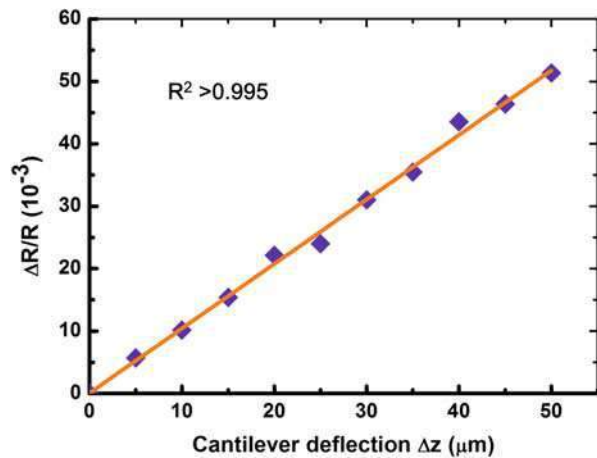
### Sensitivity

The deflection sensitivity of piezoresistive microcantilever can be obtained by measuring the relative resistance change with a force applied at the cantilever end. The vertical tip deflection of the microcantilever could be controlled in 5~10  $\mu\text{m}$  steps with a precision stage. Figure 15 shows a typical relation between the relative resistance change ( $\Delta R/R$ ) and the vertical deflection ( $\Delta z$ ) of the microcantilever, with a resistance of 7 k $\Omega$  and a dimension of 200  $\mu\text{m} \times 50 \mu\text{m} \times 1 \mu\text{m}$ , where a deflection sensitivity ( $\Delta R/R \cdot \Delta z^{-1}$ ) can be obtained by the slope of the fitted

**Fig. 14** Noise spectra curves corresponding to different voltages applied to the Wheatstone bridge. Insert shows the dependence of noise level  $(S_{VHf})^{1/2}$  on bias voltages



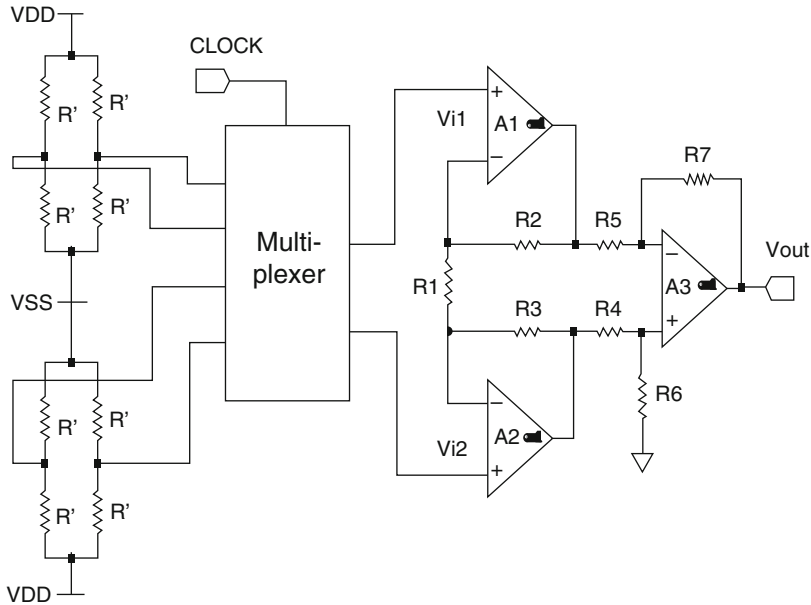
**Fig. 15** Relative resistance change of piezoresistor measured for tip deflection of 50  $\mu\text{m}$  in 5  $\mu\text{m}$  steps, the deflection sensitivity is given by using the fitted line slope



line. Substituting Hooke factor, gauge factor, and cantilever dimensions into the calculation formula of MDD, the MDD of the rectangular cantilevers can be calculated smaller than 0.1 nm at a 1000 Hz measurement bandwidth and a 3 V bias voltage (Xiaomei Yu 2001).

## Monolithic Integration of Piezoresistive Microcantilever with Readout Circuits (Xiaomei Yu et al. 2008)

Piezoresistive transducers are widely adopted in measuring cantilever bending due to the simple electrical output measurement and easier to be integrated with integrated circuits (ICs). The need for an integration of the MEMS devices with the integrated circuitry is crucial for communicating and transducing the minute signals to the macroscopic world. There is a documented evidence of microcantilevers integrated



**Fig. 16** Schematic diagram of an integrated sensor system

with a signal conditioning circuitry for detecting different kinds of physical or chemical properties (Colinge 2000; Li et al. 2003). Generally speaking, MEMS-IC integration can be categorized into two groups, including hybrid integration and monolithic integration. In the case of hybrid integration, MEMS parts and circuit parts are fabricated separately and then packaged together by wafer bonding or other packaging technologies. With regard to the micro-to-nanometer-scale signals of cantilever sensors, monolithic integration is preferred. The main contributions of monolithic integration are the decrease of the system bulk and cost, the increase of the device reliability, and the elimination of parasitic capacitance introduced by the external bonding pads and wires. The complementary metal-oxide-semiconductor (CMOS) circuitry for the readout of the cantilever deflection integrated together with the cantilever by using a monolithic technology will permit in situ and smart detections of the minute information. Herein, we provide a monolithic integration of the microcantilevers and the CMOS circuitry by using both the silicon-on-insulator (SOI) CMOS and the SOI micromachining technologies.

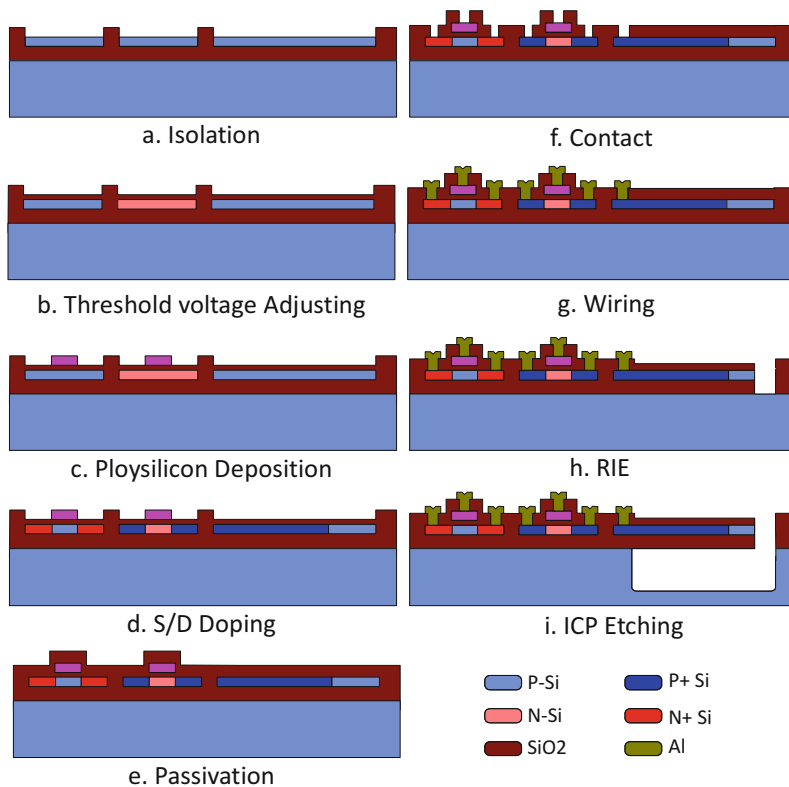
Figure 16 shows the schematic diagram of a monolithically integrated sensor system, which includes two Wheatstone bridges, a multiplexer, and an instrumentation amplifier. The differential voltage signals from the Wheatstone bridges are first input into a time-division multiplexer. The multiplexer is controlled by a digital clock works as a function of collecting multisignals from different sensors and transfers them into the amplifier. Therefore, those signals from different sensor channels can be detected with the instrument amplifier at a different time. A two-stage cascade symmetrical operational transconductance amplifier (OTA) is selected

to implement the voltage amplification because of its relatively high amplification coefficient, high output swing, low noise, and simple architecture. The first stage of the two-stage instrumentation amplifier is made up of the subamplifier A1, subamplifier A2, and their proportional resistors R1, R2, and R3, while subamplifier A3 and its proportional and feedback resistors R4, R5, R6, and R7 build the second stage of the amplifier.

A SOI wafer with a device layer thickness of 200 nm is chosen to fabricate this integrated system in considering the requirement of cantilever sensitivity, and partial depletion (PD) type SOI CMOS technology together with the SOI micromachining processes was adopted in our integration process. Body contacts, which will lead the holes at the body region into the source end, were designed in order to avoid the kink effect in PD-SOI NMOS transistors. PSPICE and ISE DIOS were used to simulate the characteristics of SOI CMOS transistors and circuits at different substrates and gate biasing voltages, and some of the device and process parameters were extracted from the simulated results. To obtain  $\pm 0.8$  V threshold voltages with  $V_{BS} = 0$ , the optimum doping concentrations in polysilicon gate were simulated to be  $2.2 \times 10^{12} \text{ cm}^{-2}$  and  $2.5 \times 10^{12} \text{ cm}^{-2}$  for SOI NMOS and SOI PMOS transistors respectively, which correspond a doping concentration of  $1.1 \times 10^{17} \text{ cm}^{-3}$  in transistor channel. For the transistor with a width/length ratio of 50/20 and a gate oxide thickness of 20 nm, the simulated leakage current is  $\pm 250$  nA. These process parameters are applied to the CMOS circuit designs and will give us a basic guidance to the following process design at the same time.

P-type single-crystalline silicon has the biggest piezoresistive coefficient in  $\langle 110 \rangle$  orientation, so a SOI wafer with a 200-nm-thick device layer (p-type, in (100) crystal surface) and 400-nm-thick box oxide layer is used for the fabrication, and the fabrication processes are depicted in Fig. 17. Apart from the cantilever releasing, a standard 1- $\mu\text{m}$  CMOS technology in the Micro/Nanofabrication Lab of Peking University was applied to all processes, which include: (a) Defining the active and field regions by fully oxidizing the device layer with the first mask. The piezoresistors were also defined in this step, therefore to form a  $\text{SiO}_2$  sidewall encapsulation layer around the piezoresistors; (b) Light doping with boron and phosphor ions with mask 2 and 3 in order to adjust the threshold voltage of CMOS devices; (c) Depositing polysilicon layer and patterning the polysilicon gate electrodes with mask 4; (d) Heavily doping with boron and phosphor ions to form the source and drain regions with mask 5 and 6, as well as the piezoresistive Wheatstone bridges; (e) Insulating the circuits with a  $\text{SiO}_2$  film and electrically activating the dopants for above implantations; (f) Opening of the contact holes for the metallization with mask 7; (g) Sputtering and patterning the Al metal wires with mask 8 both for the CMOS circuits and the Wheatstone bridges; (h) After a photolithography with mask 9, reactive ion etching the top oxide layer, the device layer, and the bottom oxide layer to define the cantilevers again; (i) At last, a silicon isotropy etching under an advanced silicon etching (ASE) system was applied to release the cantilevers from the front side.

Since the cantilever dimension defined in step (h) is 1  $\mu\text{m}$  wider than the first definition with mask 1, the field oxidizing layer together with the buried oxide and

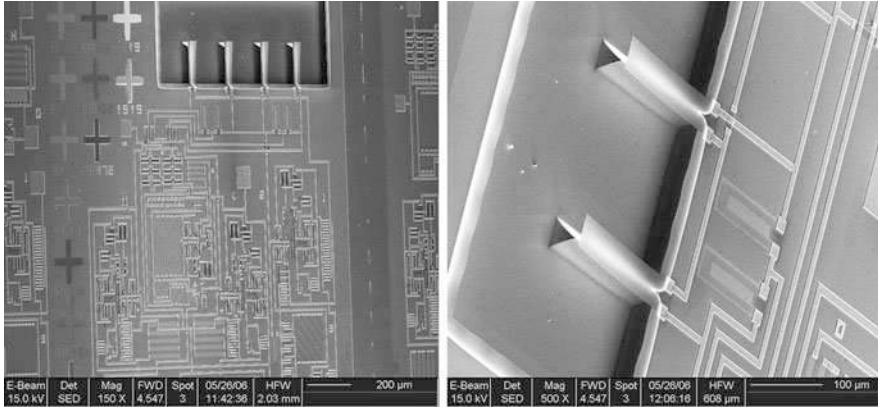


**Fig. 17** Fabrication processes for the monolithically integrated piezoresistive microcantilever with post-MEMS processes

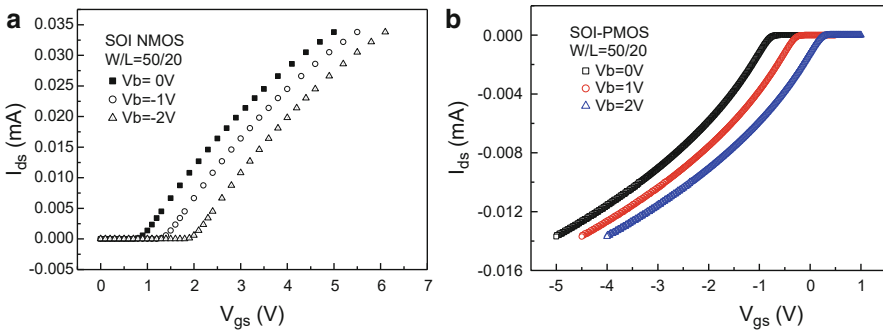
the SiO<sub>2</sub> insulating layer of CMOS circuits will serve as the sidewall, the bottom, and the top encapsulation layer of the piezoresistors, which will protect the piezoresistors to be dry etched in the ASE system.

A SEM photo of the finished integrated piezoresistive microcantilever chip is shown in Fig. 18a, while a close-up view, where two released cantilevers can be clearly seen, is shown in Fig. 18b. Owing to the feature of isotropic etching from the front side, silicon prisms are left at the bottom of etched cavity beneath the cantilevers, and these prisms should have no significant influence on the sensor performances.

We tested the transistors and circuits characteristics by a HP 4156 precision semiconductor parameter analyzer. The transfer characteristic curves of a PMOS transistor and a NMOS transistor with the body contact designed and at different substrate biasing voltages are shown in Fig. 19, and the threshold voltages were extracted to be 0.82 V and 0.84 V for the NMOS and PMOS transistors respectively, which are almost consistent with the  $\pm 0.8$  V designed values. The transfer characteristic curves show that the transistors work well as expected, and the kink effect on the SOI NMOS device is totally avoided with the body contact designed.



**Fig. 18** SEM photographs of the monolithically integrated piezoresistive microcantilever



**Fig. 19** Transfer characteristic curve for (a) a SOI NMOS transistor with body contact designed and (b) a SOI PMOS transistor

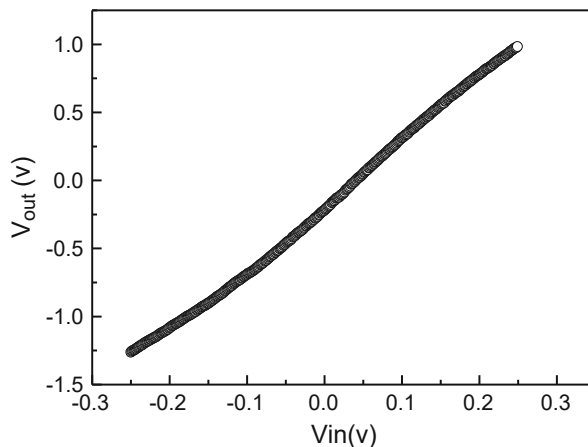
A close-loop DC sweep curve of a preamplifier is shown in Fig. 20. It is observed from the figure that when the input has a change of 0.5 V (from  $-0.25$  to  $0.25$  V), the output voltage is amplified to 2.25 V (from  $-1.25$  to  $1.0$  V), with the gain of 4.5 and a linearity of 1%. The input-output characteristic curve of an instrument amplifier is shown in Fig. 21. When the input has a change of 0.45 V (from 1.2 to 1.65 V), the output is amplified to 9.5 V (from  $-4.7$  to  $4.8$  V). The gain magnitude of the instrument amplifier is measured to be 20, which is consistent with the calculated and simulated results.

## Chemical Sensors

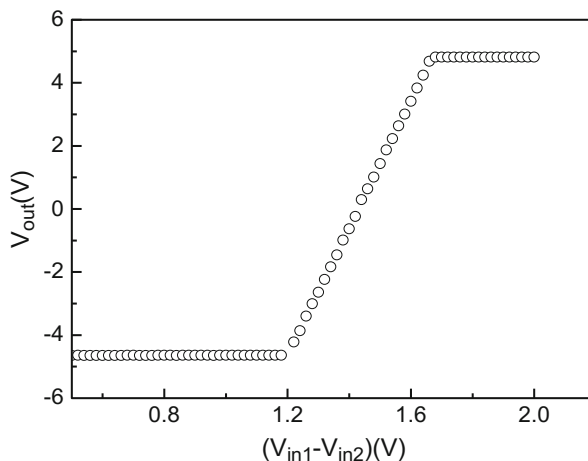
Microcantilever with gold-coated layer was first used to detect volatile compounds in the 1990s of last century (Lang et al. 1999). Since then, a lot of work was reported. In order to pursue a wide spectrum of applications and higher selectivity, several



**Fig. 20** Close-loop DC sweep characteristic curve of an OTA

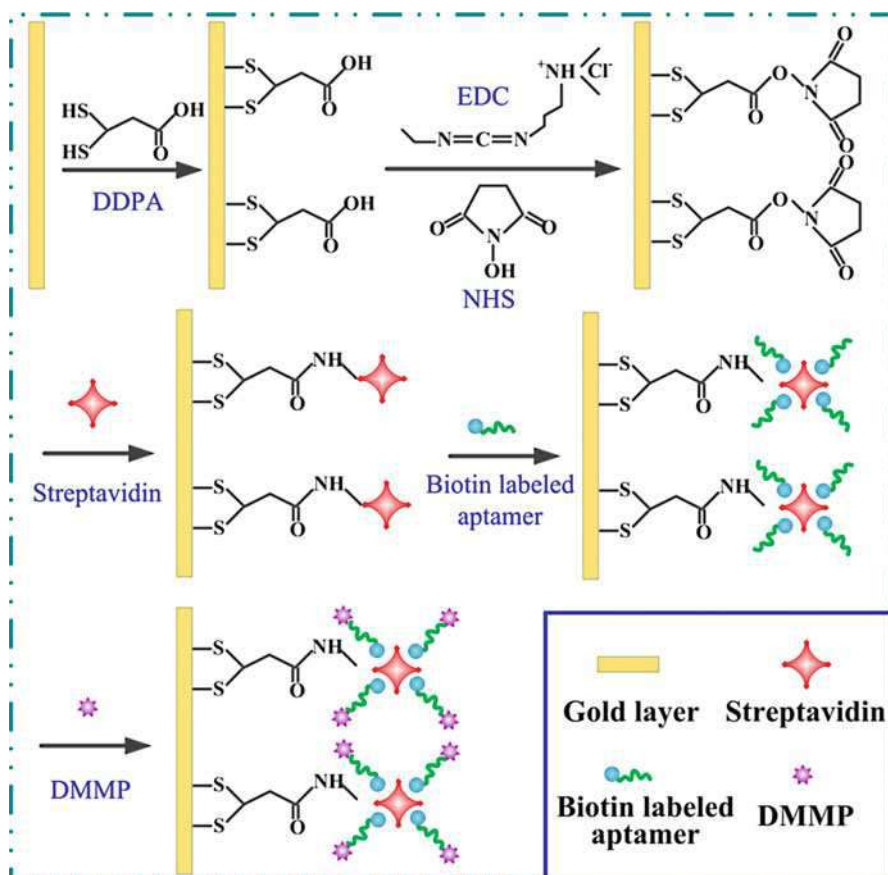


**Fig. 21** DC input-output characteristic curve of an instrument amplifier



polymer coatings were introduced to perform the chemical detections, such as polyethylene glycol (PEG) and poly-*N*-vinyl pyrrolidinone (PVP) spun on microcantilever to measure the relative changes of humidity, and polymethyl methacrylate (PMMA) tethered to a functionalized overlayer on microcantilever by self-assembly to detection polar organic vapor (McCaig et al. 2014). In the field of explosive detections, Chen et al. reported the measurement of trinitrotoluene (TNT) using a piezoresistive microcantilever with the methoxy group of the siloxane molecule (Chen et al. 2010).

Detections on organic phosphorus compounds and heavy-metal ions, in general, have also been studied. Dutta et al. used single and binary mixtures of different



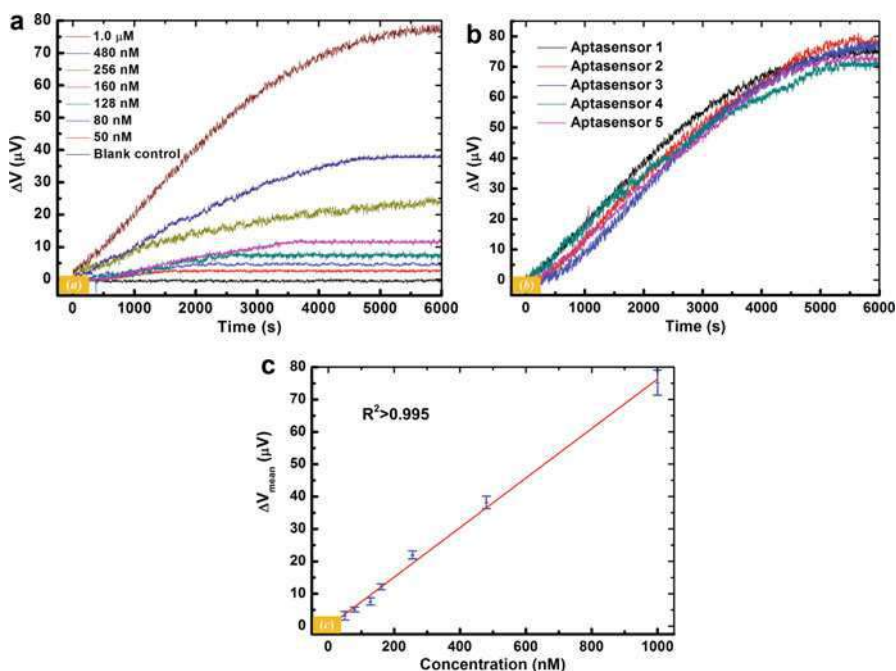
**Fig. 22** Schematic diagram of anti-DMMP immobilization on the sensing cantilevers and DMMP molecules binding with anti-DMMP

thiolated ligands as self-assembled monolayers to detect  $\text{Co}^{2+}$ ,  $\text{Cs}^{2+}$ ,  $\text{Fe}^{3+}$ , etc. with the concentration of 0.1 mM (Dutta et al. 2005). Some researchers proposed biological analytical methods based on  $\text{Cu}^{2+}$ /L-cysteine layer (Yang et al. 2003b), acetylcholine esterase (Yan et al. 2004), and organophosphorus hydrolase (Karnati et al. 2007) to perform the detections in aqueous, and these methods showed relatively different specificity.

In 2004, aptamers began to attract the attention of researchers as new binding elements for biochemical sensors. They are synthesized short single-stranded oligonucleotides that bind to specific molecules tightly with the comprehensive advantages, such as chemical stability, high affinity, and wide target library in the liquid phase (Du et al. 2012; Song et al. 2008; Zhao et al. 2014). In particular, the conformation of aptamers often undergoes significant changes with high

selectivity. All these unique properties make aptamers able to offer remarkable flexibility in designing novel biological and chemical sensors with high sensitivity. Yu et al. showed a piezoresistive cantilever-based aptasensor for the detection of dimethyl methylphosphonate (DMMP), a stimulant of nerve agents, in aqueous matrices. Figure 22 schematically shows the anti-DMMP immobilization on the sensing cantilevers and DMMP molecules binding with anti-DMMP (Zhao et al. 2017).

Figure 23 plots a group of typical output voltages of the piezoresistive microcantilever-based sensors as the analyte concentration increased. It is clear that all the responses can reach steady states but with different response magnitude, this can be explained that the induced surface stress is correlated with analyte concentrations. In order to evaluate the specificity of the chemical sensors, comparison experiments should be implemented by exposing the sensors to analyte and analogs at the same concentration. The reproducibility statistics is also very important for the chemical sensors, in which the steady-state responses versus the analyte concentrations should show a linear relation in the detection range of analyte.



**Fig. 23** (a) The measured results of DMMP at concentrations of 50, 80, 128, 160, 256, 480 and 1000 nM respectively; (b) The response curves of five reproducibility tests at a concentration of 1.0  $\mu M$ ; (c) The steady-state responses of aptasensors versus the DMMP concentrations, which show a linear relation

## Biosensors

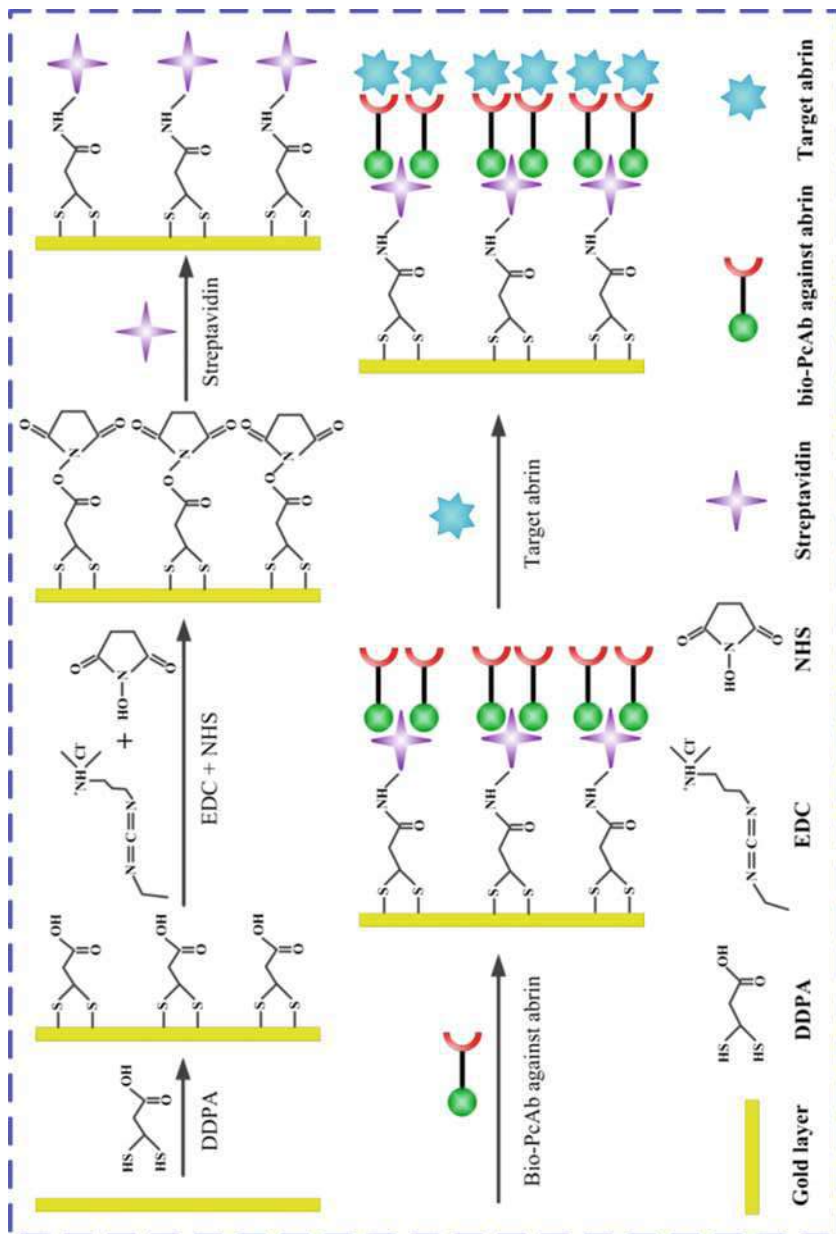
The significant milestone in developing microcantilever-based biosensor was the demonstration of their capability to analyze DNA hybridization and conformational changes of protein. Fritz et al. used arrays of functionalized microcantilever and optical readout technique to monitor the specific hybridization of single-stranded DNA in real-time (Fritz et al. 2000). Wu et al. reported on sensitive and specific monitoring of conformational changes of protein using microcantilever and optical readout of its deflection (Wu et al. 2001). After more than a decade, diverse functionalization schemes have been tried to detect antigen-antibody binding, DNA hybridization, enzyme catalysis, bacteria, biotoxin, and pathogen. Recently, developing efficient biosensors for onsite detection with high sensitivity and reliability have attracted much attention, especially after the terrorist attacks on September 11, 2001. Silicon-based piezoresistive microcantilevers, as one of the typical surface stress-based cantilevers, are highly attractive as biological platforms for their promising low cost, portability, and real-time performance. However, they generally present relatively lower sensitivity with SNR difficult to be improved.

Zhao et al. reported an immunological detection of abrin with the limit detection is 1.2 pM based on the biotin-streptavidin method (Zhao et al. 2015). Figure 24 shows the schematic diagram of an antibody against abrin functionalization processes and the target abrin binding with probes, and the real-time monitoring of the functionalized piezoresistive microcantilever biosensors with the addition of different concentrations of abrin are shown in Fig. 25. The binding of the antibody against abrin and the target abrin induced a change in surface stress, which caused a differential output signal of piezoresistive microcantilever biosensors in the range 3~80  $\mu\text{V}$ , depending on the abrin concentration (1.2~600 pM).

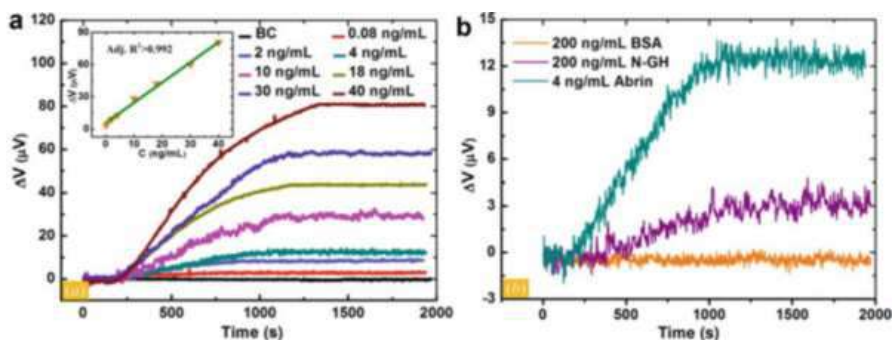
---

## Summary

In this chapter, biochemical sensors based on piezoresistive microcantilever have been discussed. The chapter starts out by presenting static and dynamic operation modes and the readout methods of the piezoresistive microcantilever, which is important for the design of microcantilever-based biochemical sensors. Then the noise, deflection sensitivity, and resolution are briefly discussed in order to give a deep understanding of the important parameters involved in the piezoresistive microcantilever design. After that, two fabrication processes and a monolithic integration technique of piezoresistive microcantilever with readout circuit are provided, and some characteristics were measured at the same time. Finally, we discussed the applications of piezoresistive microcantilever in biochemical sensing and indicate that microcantilever provides a viable platform for the development of biochemical sensor.



**Fig. 24** Schematic diagram of antibody against abrin functionalization processes and the target abrin binding with probes



**Fig. 25** (a) Real-time monitoring of the functionalized PZR cantilever biosensors with the addition of different concentrations of abrin. Inset presents linear relation of output and detecting concentration. (b) Specific comparison detections utilizing two analogs at concentrations of 200 ng/mL and abrin at a concentration of 4 ng/mL

## References

- Alvarez M, Lechuga LM (2010) Microcantilever-based platforms as biosensing tools. *Analyst* 135:827–836
- Binnig G, Quate CF, Gerber C (1986) Atomic force microscope. *Phys Rev Lett* 56:930
- Butt H-J (1996) A sensitive method to measure changes in the surface stress of solids. *J Colloid Interface Sci* 180(1):251–260
- Carrascosa LG, Moreno M, Álvarez M, Lechuga LM (2006) Nanomechanical biosensors: a new sensing tool. *TrAC Trends Anal Chem* 25:196–206
- Chen GY, Thundat T, Wachter EA, Warmack RJ (1995) Adsorption-induced surface stress and its effects on resonance frequency of microcantilevers. *J Appl Phys* 77(8):3618–3622
- Chen X, Salm C, Hooge F, Woerlee P (1999)  $1/f$  noise in polycrystalline SiGe analyzed in terms of mobility fluctuations. *Solid State Electron* 43:1715–1724
- Chen Y, Xu P, Li X (2010) Self-assembling siloxane bilayer directly on  $\text{SiO}_2$  surface of microcantilevers for long-term highly repeatable sensing to trace explosives. *Nanotechnology* 21:265501
- Colinge J-P (2000) SOI devices and circuits. In: *Microelectronics, 2000. Proceedings. 2000 22nd international conference on*. IEEE, pp 407–414
- De Graaff H, Huybers M (1983)  $1/f$  noise in polycrystalline silicon resistors. *J Appl Phys* 54:2504–2507
- Du Y, Li B, Wang E (2012) “Fitting” makes “sensing” simple: label-free detection strategies based on nucleic acid aptamers. *Acc Chem Res* 46:203–213
- Dutta P, Chapman PJ, Datskos PG, Sepaniak MJ (2005) Characterization of ligand-functionalized microcantilevers for metal ion sensing. *Anal Chem* 77:6601–6608
- Fritz J (2008) Cantilever biosensors. *Analyst* 133:855–863
- Fritz J, Baller M, Lang H, Rothuizen H, Vettiger P, Meyer E, Güntherodt H-J, Gerber C, Gimzewski J (2000) Translating biomolecular recognition into nanomechanics. *Science* 288:316–318
- Gates RS, Reitsma MG, Kramar JA, Pratt JR (2011) Atomic force microscope cantilever flexural stiffness calibration: toward a standard traceable method. *J Res Natl Inst Stand Technol* 116:703
- Hooge FN (1969)  $1/f$  noise is no surface effect. *Phys Lett A* 29:139–140
- Hooge F (1994)  $1/f$  noise sources. *IEEE Trans Electron Devices* 41:1926–1935
- Hooge F, Kleinpenning T, Vandamme L (1981) Experimental studies on  $1/f$  noise. *Rep Prog Phys* 44:479

- Karnati C, Du H, Ji H-F, Xu X, Lvov Y, Mulchandani A, Mulchandani P, Chen W (2007) Organophosphorus hydrolase multilayer modified microcantilevers for organophosphorus detection. *Biosens Bioelectron* 22:2636–2642
- Koev ST, Bentley WE, Ghodssi R (2010) Interferometric readout of multiple cantilever sensors in liquid samples. *Sens Actuators B Chem* 146:245–252
- Kosaka PM, Pini V, Ruz J, da Silva R, González M, Ramos D, Calleja M, Tamayo J (2014) Detection of cancer biomarkers in serum using a hybrid mechanical and optoplasmonic nanosensor. *Nat Nanotechnol* 9:1047–1053
- Lang HP, Ballera MK, Bergerc R, Gerberc C, Gimzewskic JK, Battistonb FM, Fornarob P, Ramseyerb JP, Meyerb E, Güntherodtb HJ (1999) An artificial nose based on a micromechanical cantilever array. *Anal Chim Acta* 393(1–3):59–65
- Lavrik NV, Sepaniak MJ, Datskos PG (2004) Cantilever transducers as a platform for chemical and biological sensors. *Rev Sci Instrum* 75:2229–2253
- Li Y, Vancura C, Hagleitner C, Lichtenberg J, Brand O, Baltes H (2003) Very high Q-factor in water achieved by monolithic, resonant cantilever sensor with fully integrated feedback. In: *Sensors, 2003. Proceedings of IEEE. IEEE*, pp 809–813
- McCaig HC, Myers E, Lewis NS, Roukes ML (2014) Vapor sensing characteristics of nano-electromechanical chemical sensors functionalized using surface-initiated polymerization. *Nano Lett* 14:3728–3732
- O'Shea SJ, Welland ME, Brunt TA, Ramadan AR, Rayment T (1996) Atomic force microscopy stress sensors for studies in liquids. *J Vac Sci Technol B Microelectron Nanometer Struct* 14(2):1383–1385
- Pandya H, Kim HT, Roy R, Desai JP (2014) MEMS based low cost piezoresistive microcantilever force sensor and sensor module. *Mater Sci Semicond Process* 19:163–173
- Pillarsetti A, Desai JP, Ladjal H, Schifffmacher A, Ferreira A, Keefer CL (2011) Mechanical phenotyping of mouse embryonic stem cells: increase in stiffness with differentiation. *Cell Reprogram (Former "Cloning Stem Cells")* 13:371–380
- Raiteri R, Butt H-J (1995) Measuring electrochemically induced surface stress with an atomic force microscope. *J Phys Chem* 99(43):15728–15732
- Song S, Wang L, Li J, Fan C, Zhao J (2008) Aptamer-based biosensors. *TrAC Trends Anal Chem* 27:108–117
- Stoney GG (1909) The tension of metallic films deposited by electrolysis. *Proc R Soc Lond Ser A Contain Pap Math Phys Character* 82:172–175
- Tamayo J, Kosaka PM, Ruz JJ, San Paulo Á, Calleja M (2013) Biosensors based on nanomechanical systems. *Chem Soc Rev* 42:1287–1311
- Thaysen J (2001) Cantilever for bio-chemical sensing integrated in a microliquid handling system. PhD thesis, Technical University of Denmark
- Vandamme L (1983) Noise in physical systems and 1/f noise. In: Savelli M, Lecoy G, Nougierfeds J-P (eds) *Proceedings of the 7th international conference on "noise in physical systems" and the 3rd international conference on "1/f noise," Montpellier, May 17–20, 1983*. Elsevier Science Publishers BV/North Holland, p 183
- Vandamme L, Oosterhoff S (1986) Annealing of ion-implanted resistors reduces the 1/f noise. *J Appl Phys* 59:3169–3174
- Wu G, Datar RH, Hansen KM, Thundat T, Cote RJ, Majumdar A (2001) Bioassay of prostate-specific antigen (PSA) using microcantilevers. *Nat Biotechnol* 19:856–860
- Xiaomei Yu (2001) Optimization of piezoresistive cantilever with respect to signal and noise. PhD thesis, Beihang University
- Xiaomei Yu, Yaquan Tang, Haitao Zhang (2008) Monolithic integration of micromachined sensors and CMOS circuits based on SOI technologies. *J Micromech Microeng* 18:037002. (7pp)
- Yan X, Tang Y, Ji H, Lvov Y, Thundat T (2004) Detection of organophosphates using an acetyl cholinesterase (AChE) coated microcantilever. *Instrum Sci Technol* 32:175–183
- Yang M, Zhang X, Vafai K, Ozkan CS (2003a) High sensitivity piezoresistive cantilever design and optimization for analyte-receptor binding. *J Micromech Microeng* 13:864

- Yang Y, Ji H-F, Thundat T (2003b) Nerve agents detection using a  $\text{Cu}^{2+}$ /L-cysteine bilayer-coated microcantilever. *J Am Chem Soc* 125:1124–1125
- Yu M-F, Lourie O, Dyer MJ, Moloni K, Kelly TF, Ruoff RS (2000) Strength and breaking mechanism of multiwalled carbon nanotubes under tensile load. *Science* 287:637–640
- Yu X, Jiang X, Thaysen J, Hansen O, Boisen A (2001) Noise and sensitivity in polysilicon piezoresistive cantilevers. *Chin Phys* 10:918
- Yu X, Thaysen J, Hansen O, Boisen A (2002) Optimization of sensitivity and noise in piezoresistive cantilevers. *J Appl Phys* 92:6296–6301
- Zhao R, Wen Y, Yang J, Zhang J, Yu X (2014) Aptasensor for *Staphylococcus* enterotoxin B detection using high SNR piezoresistive microcantilevers. *J Microelectromech Syst* 23:1054–1062
- Zhao R, Ma W, Wen Y, Yang J, Yu X (2015) Trace level detections of abrin with high SNR piezoresistive cantilever biosensor. *Sens Actuators B Chem* 212:112–119
- Zhao R, Jia D, Wen Y, Yu X (2017) Cantilever-based aptasensor for trace level detection of nerve agent simulant in aqueous matrices. *Sens Actuators B Chem* 238:1231–1239





# Micro-Heater-Based Gas Sensors

Tie Li, Lei Xu, and Yuelin Wang

## Contents

Overview .....	718
Basic Design Principles of Micro-heaters .....	719
Structure Design .....	719
Materials .....	720
FEM Analysis .....	721
2D Micro-heaters .....	722
Fabrication Process .....	723
Two-Beam Micro-heaters .....	724
Optimization of Temperature Distribution .....	725
3D Micro-heaters .....	727
Fabrication Process .....	728
Comparison of 3D Micro-heater and 2D Micro-heater .....	730
FEM Analysis for 2D/3D Micro-heaters .....	731
Catalytic Gas Sensors .....	732
Sensing Principle .....	733
Fabrication Process .....	734
2D Catalytic Gas Sensors .....	734
3D Catalytic Gas Sensors .....	735
Comparison of 3D and 2D Catalytic Gas Sensors .....	738
Semiconductor Gas Sensors .....	741
Sensing Principle .....	742
Structure Design .....	742
Fabrication Process .....	743

---

T. Li (✉) · Y. Wang

Science and Technology on Micro-system Laboratory, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China

e-mail: [tli@mail.sim.ac.cn](mailto:tli@mail.sim.ac.cn); [ylwang@mail.sim.ac.cn](mailto:ylwang@mail.sim.ac.cn)

L. Xu

School of Information Science and Technology, University of Science and Technology of China, Hefei, China

e-mail: [okxulei@ustc.edu.cn](mailto:okxulei@ustc.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

717

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,  
[https://doi.org/10.1007/978-981-10-5945-2\\_21](https://doi.org/10.1007/978-981-10-5945-2_21)

Wafer-Level Fabrication of Micro-/Nano-gas Sensors .....	745
Properties of Micro-/Nano-gas Sensors .....	748
Conclusions .....	749
References .....	751

### Abstract

Microelectromechanical system (MEMS) technology has been promoting development of sensors. Micro-heaters, as typical MEMS devices, have been extensively researched for high-performance gas sensors in the past two decades. This chapter presents the design, fabrication, and characterization of micro-heaters with low power consumption, well-controlled temperature distribution, high mechanical strength, and long-term stability. Catalytic gas sensors and semiconductor gas sensors will be introduced with novel design and high performance. The goal of this chapter is to provide the reader with a general overview of micro-heaters and broader ways of designing better micro-heater-based gas sensors.

### Keywords

Gas sensor · Micro-heater · Micro-hot platform (MHP) · FEM analysis · Micro-/nano fabrication · Catalytic combustion · Metal oxide semiconductor (MOS) · Lower explosive limit (LEL)

## Overview

Gas sensors have been widely used in areas such as industrial production, environmental monitoring, medical devices, and domestic appliances. Most gas sensors like catalytic gas sensors and semiconductor gas sensors need a heater to provide appropriate temperature for gas sensing. However, such heater-based gas sensors with traditional fabrication process can hardly meet the increasing demands in low cost, small size, low power consumption, high sensitivity, and long lifetime.

MEMS is a promising technology which will overcome the difficulties of conventional fabrication process. The introduction of micro-heaters fabricated by MEMS technology will tremendously improve the performance of gas sensors. To achieve a micro-heater-based gas sensor with good properties, structure design, material selection, thermal analysis, and fabrication process have to be taken into consideration.

This chapter reviews the research and developments of the authors established so far in the field of micro-heater-based gas sensors. It includes general concepts, basic design principles, novel structures, and new fabrication process of micro-heaters and gas sensors. The general concepts and basic design properties of micro-heaters such as structures, materials, and analysis are mainly described in section “[Basic Design Principles of Micro-heaters](#).” Two-dimensional (2D) micro-heaters with low power consumption, well-controlled temperature distribution, high mechanical strength, and long-term stability are introduced in section “[2D Micro-heaters](#).” Then, a

novel three-dimensional (3D) micro-heater with high heating efficiency is presented in section “[3D Micro-heaters.](#)” Based on these micro-heaters, catalytic gas sensors and semiconductor gas sensors, which are widely used for detecting explosive gases and volatile organic compounds (VOC), are presented in sections “[Catalytic Gas Sensors](#)” and “[Semiconductor Gas Sensors.](#)”

---

## Basic Design Principles of Micro-heaters

A micro-heater generally consists of a resistor supported by a dielectric membrane within a silicon rim for thermal isolation. To fabricate micro-gas sensors, micro-heaters with functional elements have to be designed (Simon et al. 2001). This is complicated for micro-heaters that consist of thin films with different structures and materials. This section describes typical designs and some criteria for the fabrication of micro-heaters with low power consumption, well-controlled temperature distribution, high mechanical strength, and long-term stability.

Low power consumption is necessary for gas sensors, especially for the applications in wireless sensing networks and wearable devices with batteries. Since the sensing properties of micro-gas sensors are strongly temperature dependent, temperature distributions over the sensing layer have to be well controlled (Kunt et al. 1998). High mechanical strength is desired not only for all processing steps to fabricate micro-heaters but also for micro-gas sensors during the deposition of sensing film and subsequent processing. Long-term stability is required for practical applications with reliable sensitivity and long lifetime.

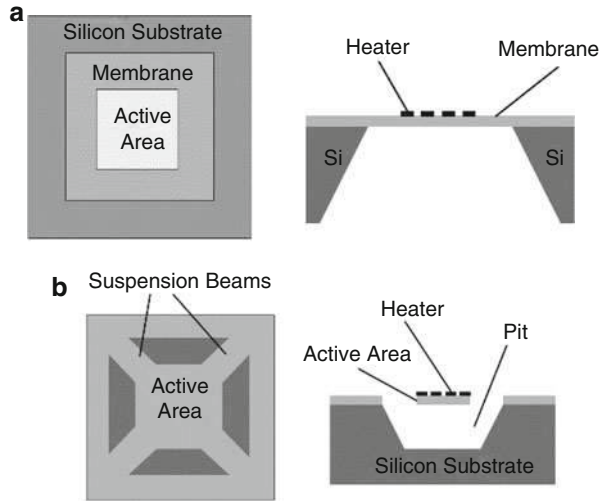
### Structure Design

Basically there are two different structures of micro-heaters according to the dielectric membrane. One is the closed-membrane-type micro-heater and the other the suspended-membrane-type micro-heater (Simon et al. 2001). Schematic views of these structures are presented in Fig. 1.

The closed-membrane-type micro-heater is formed by anisotropic etching of silicon from the back side. Wet etchants like KOH or TMAH (tetramethylammonium hydroxide) are often used. The disadvantage of closed-membrane-type micro-heater lies in the need of double-side alignment for the bulk silicon etching from the back side and the sloped sidewalls obtained by anisotropic etching. The sloped sidewalls make the lateral dimensions needed to form a membrane a factor  $\sqrt{2}$  larger (Simon et al. 2001).

The suspended-membrane-type micro-heater is completely fabricated from the front side. The suspended membrane is either formed by anisotropic etching with KOH or TMAH from the front side or by sacrificial etching. In case of anisotropic etching, the membrane is first formed by dielectric layers like silicon oxide and/or nitride and patterned to form etch windows. Sacrificial etching of porous silicon is another possibility to obtain suspended membranes (Lim et al. 1998).

**Fig. 1** Schematic views of the closed-membrane-type micro-heater (a) and the suspended-membrane-type micro-heater (b)



## Materials

Basically, a micro-heater consists of a membrane as a supporting layer and a resistor generating heat while current passes through it. The selection and combination of materials highly influence power consumption, mechanical strength, and stability of micro-heaters.

Silicon oxide and silicon nitride, which are of low thermal conductivity, are commonly selected as supporting membrane with a total thickness between 1 and 2  $\mu\text{m}$  for low power consumption. Silicon nitride layers generally have large tensile stress, and silicon oxide layers are compressive. Nevertheless, by appropriate choice of deposition and annealing steps, multilayer systems of silicon oxide and/or silicon nitride have been proven mechanically stable (Simon et al. 2001). Residual stress and thermal stress should be well controlled under an acceptable level. Residual stress is generally much larger than thermal stress, and the control of residual stress of multilayer systems seems to be crucial for membrane stability (Low et al. 1998). The resultant residual stress  $\sigma_r$  of a stacked membrane can be approximated by

$$\sigma_r = \frac{\sigma_1 d_1 + \sigma_2 d_2 + \dots}{d_1 + d_2 + \dots} = \frac{\sum_{i=1}^n \sigma_i d_i}{\sum_{i=1}^n d_i} \quad (1)$$

where  $d_i$  and  $\sigma_i$  are thickness and intrinsic stress of the different membrane layers, respectively. Rossi et al. (1998) suggest an acceptable range for resultant residual stresses as  $-0.1 \text{ GPa} < \sigma_r < +0.1 \text{ GPa}$ . A typical dielectric layer for micro-heater is made of a layer of 300 nm thermal oxide  $\text{SiO}_2$  and a layer of 1000 nm LPCVD low-stress  $\text{SiN}_x$  (Xu et al. 2011a).

The resistor is often formed by polysilicon or metals. Polysilicon is compatible in IC process, but not stable when temperature goes higher than 400  $^\circ\text{C}$ . Platinum, which has a stable temperature conference of resistance (TCR) and stable in high

temperature, is widely used for resistors and electrodes in micro-heater-based gas sensors. The temperature of Pt heater element can be extracted by the following formula:

$$T_2 = T_1 + \frac{R_2 - R_1}{R_1 \times TCR} \quad (2)$$

where  $R_1$  and  $R_2$  are the heater resistances at temperature  $T_1$  and  $T_2$ , respectively. TCR is the temperature coefficient of Pt resistance. Although it cannot give the exact temperature of the active area, this method is still a convenient and valuable way to estimate the temperature.

To achieve good adhesion of the metal resistors, generally a thin adhesive layer is deposited before Pt, such as titanium (Ti), tantalum (Ta), and chromium (Cr) (Hamann et al. 2008; Guidi et al. 1998). But Ti, Ta, and Cr can diffuse into Pt at high temperature (typically 400 °C). Not only the stability of micro-heaters working at high temperature is decreased, but also the TCR of the resistance is changed, which will influence the temperature extraction. A stable adhesive layer of  $Al_2O_3$  that performs good adhesion and stability at high temperature is demonstrated in our previous work (Xu et al. 2011a).

## FEM Analysis

The thermal and mechanical characteristics of micro-heaters have to be optimized if the sensor is intended to be of low power consumption, well-controlled temperature distribution, and high mechanical strength.

Heat transfer occurs due to heat conduction, heat convection, and radiation (Simon et al. 2001). Figure 2 shows the heat fluxes of a micro-heater. Along the membrane one has to deal with heat conduction. Above and below the membrane, heat transfer occurs due to heat conduction and heat convection through the surrounding atmosphere. Additionally, radiation has to be taken into account.

In a micro-heater, the total heat flow can be expressed as:

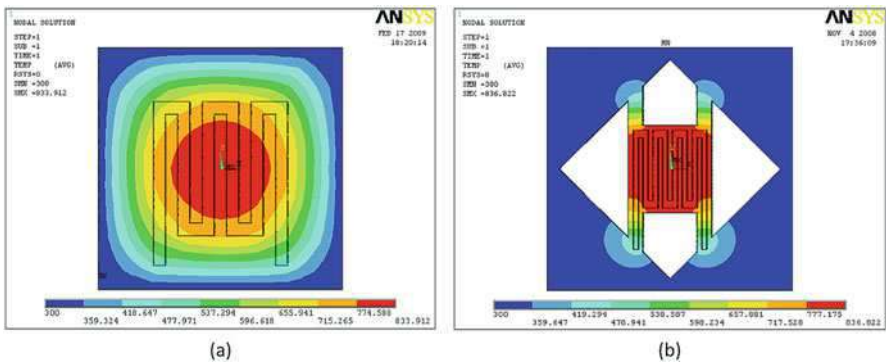
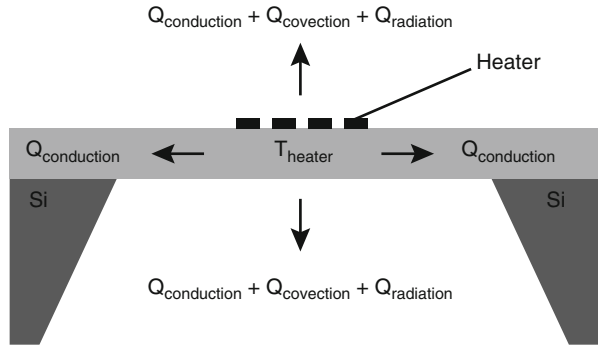
$$Q_{total} = Q_{conduction} + Q_{convection} + Q_{radiation} + \Delta_x \quad (3)$$

Here  $Q_{conduction}$  describes the heat conduction through closed membrane or suspended beams,  $Q_{convection}$  the heat losses through the air,  $Q_{radiation}$  heat losses due to radiation, and  $\Delta_x$  accounts for unknown heat losses including free convection.

According to the different pathways of heat transfer, the determination of the total heat loss and temperature distribution is not an easy task. Simple models for basic understanding are necessary, and, to minimize the costs of development, extensive simulation studies like finite element method (FEM) analysis are required. Commercially available FEM software like ANSYS and Coventor are commonly used for MEMS design. Figure 3 shows the temperature distribution of a closed-membrane-type micro-heater and a suspended-membrane-type micro-heater obtained by ANSYS.

Since micro-heaters consist of membranes with different thermal expansion coefficients, displacement and stress caused by temperature must also be considered

**Fig. 2** Heat fluxes of a micro-heater



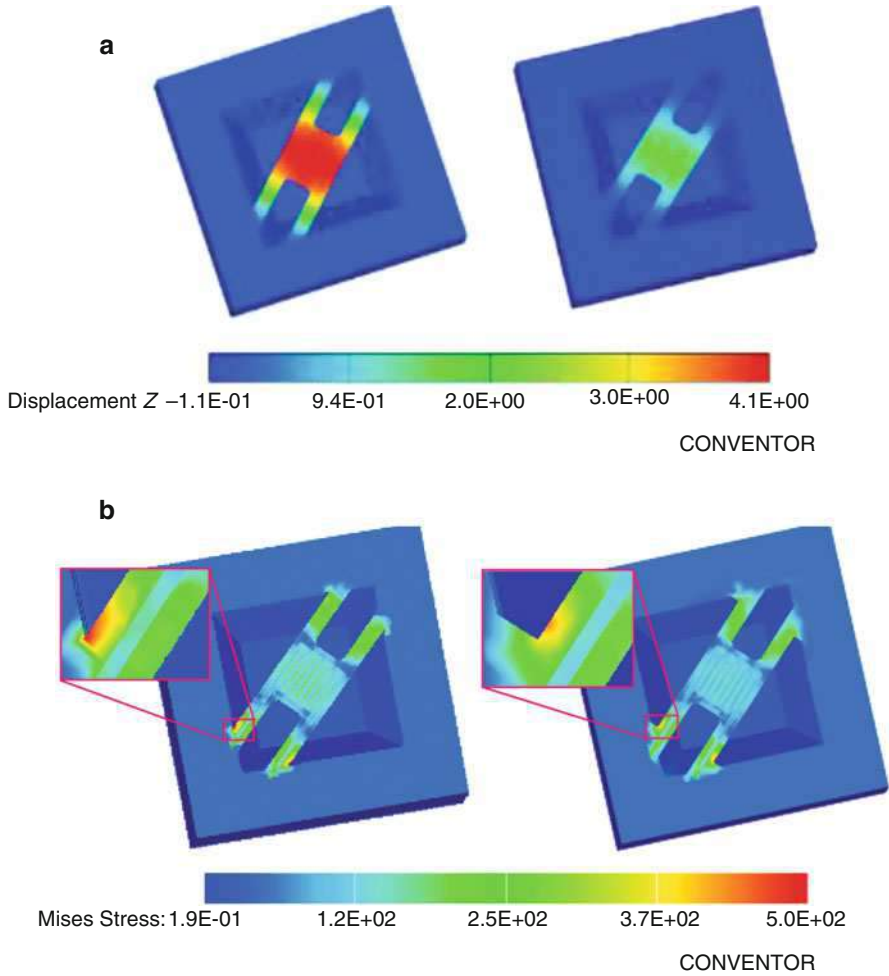
**Fig. 3** Temperature distribution of a closed-membrane-type micro-heater (a) and a suspended-membrane-type micro-heater (b) obtained by ANSYS

and optimized for high mechanical strength, especially for suspended-membrane-type micro-heaters. Figure 4 shows the FEM analysis results of displacement and stress of a four-beam suspended-membrane-type micro-heater obtained by Coventor.

The FEM analysis results suggest that the total power consumption can be minimized by constructing thin membranes with materials of low thermal conductivity, using suspended beams rather than a closed membrane, increasing the length-to-width ratio of suspended beams, and decreasing the active area (heated area). Mechanical strength of suspended-membrane-type micro-heaters can be effectively improved by adding corner compensations (Xu et al. 2011a).

## 2D Micro-heaters

A micro-heater with a plat supporting membrane, no matter it is a closed-membrane-type or suspended-membrane-type, can be defined as a 2D micro-heater. This section introduces the fabrication and optimization of 2D micro-heaters.

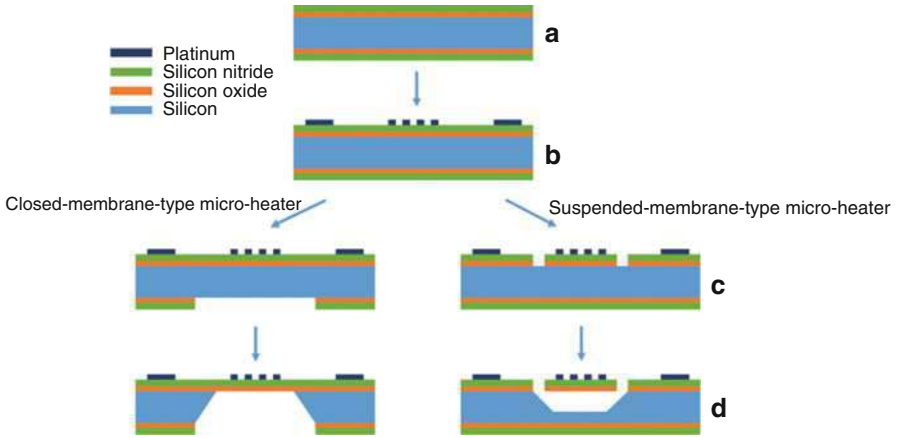


**Fig. 4** Displacement (a) and stress (b) of a four-beam suspended-membrane-type micro-heater obtained by Coventor

### Fabrication Process

Closed-membrane-type micro-heaters are of large active area and high mechanical strength. Suspended-membrane-type micro-heaters are of low power consumption and small size. Typically, the fabrication process is based on  $\langle 100 \rangle$  Si wafer with anisotropic etching to release supporting membrane. The fabrication process is shown in Fig. 5 and listed as follows:

- (a) It started with a double-side polished n-type  $\langle 100 \rangle$  oriented silicon wafer with a layer of  $\text{SiO}_2$  (300 nm) thermally grown at  $1100^\circ\text{C}$ . Then a layer of low-stress



**Fig. 5** Fabrication process of closed-membrane-type micro-heaters and suspended-membrane-type micro-heaters

SiNx (1000 nm) was deposited on each side of the silicon substrate by low-pressure chemical vapor deposition (LPCVD) at 800 °C.

- The Pt/Ti electrodes and bonding pads were patterned by lift-off process on the front side of silicon wafer.
- Then the SiO<sub>2</sub>/SiNx membrane was selectively etched by RIE etching to form releasing windows: on the back side of silicon wafer for closed-membrane-type micro-heater and on the front side of silicon wafer for suspended-membrane-type micro-heater separately.
- At last the membrane was released in a solution of TMAH (25 wt. %) at 80 °C.

Figure 6 shows the photos of a closed-membrane-type micro-heater and three suspended-membrane-type micro-heaters with different beams and structures

## Two-Beam Micro-heaters

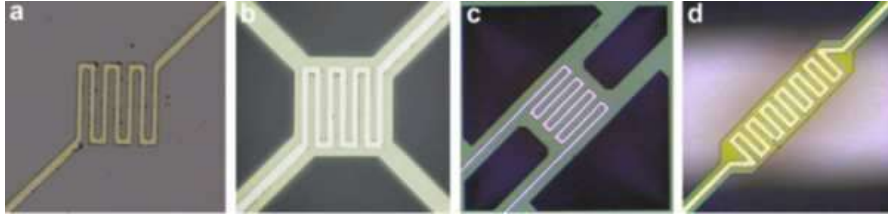
For further power consumption reduction and higher heating efficiency, a rectangular micro-heater suspended over the bulk silicon chip just via two slender beams is introduced here. Figure 7 presents the structure design and temperature distribution of the two-beam micro-heater. A rectangular membrane of 210 × 70 μm with a serpentine Pt resistor is linked to the bulk silicon chip via two 120 × 25 μm beams. Heat loss via thermal conduction was restricted by the two slender beams.

The heat loss through conduction obtains in the case of two suspended beams are

$$Q_{\text{two-beam}} = 2K_{Pt}A_{Pt}(T_{\text{hot}} - T_{\text{cold}})/L + 2K_{\text{beam}}A_{\text{beam}}(T_{\text{hot}} - T_{\text{cold}})/L \quad (4)$$

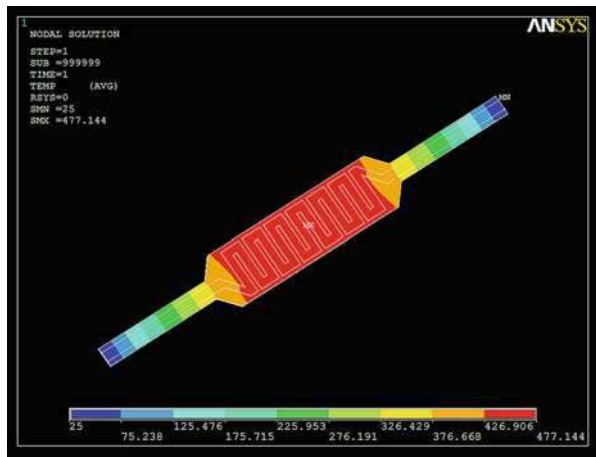
and in the case of four suspended beams:





**Fig. 6** Photos of a closed-membrane-type micro-heater (a) and three suspended-membrane-type micro-heaters (b, c, and d)

**Fig. 7** Structure design and temperature distribution of a two-beam micro-heater



$$Q_{\text{four-beam}} = 2K_{Pt}A_{Pt}(T_{\text{hot}} - T_{\text{cold}})/L + 4K_{\text{beam}}A_{\text{beam}}(T_{\text{hot}} - T_{\text{cold}})/L \quad (5)$$

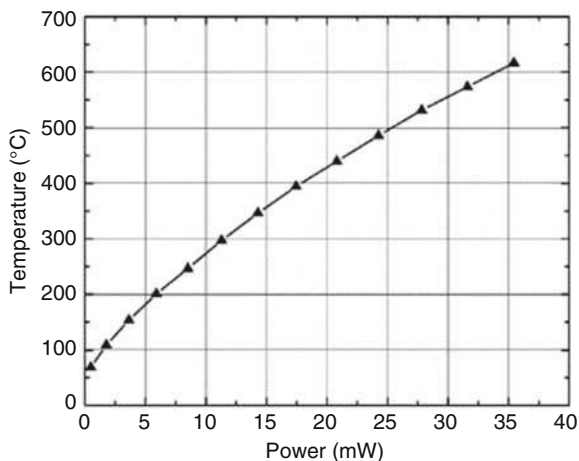
where  $K_{Pt}$  and  $K_{\text{beam}}$  are thermal conductivities of Pt and the stacked membrane beams, respectively,  $A_{Pt}$  and  $A_{\text{beam}}$  are section areas of Pt and beams,  $T_{\text{hot}}$  and  $T_{\text{cold}}$  are temperature of the hot points closed to the rectangular membrane and temperature of the bulk silicon, and  $L$  refers to the length of the beams, respectively. So heat loss through conduction of the micro-heaters with only two beams can be further reduced by approximately 44% compared to that of the micro-heaters with four suspended beams when the heating areas are heated to the same temperature.

Temperature versus power consumption of the two-beam micro-heater is show in Fig. 8.

### Optimization of Temperature Distribution

In the application of catalytic gas sensors, good temperature distribution in active area is required for better sensing performance (Gotz et al. 1997; Graf et al. 2006; Briand et al. 2000). Two-beam micro-heater is very suitable for catalytic gas sensors

**Fig. 8** Temperature versus power consumption of a two-beam micro-heater



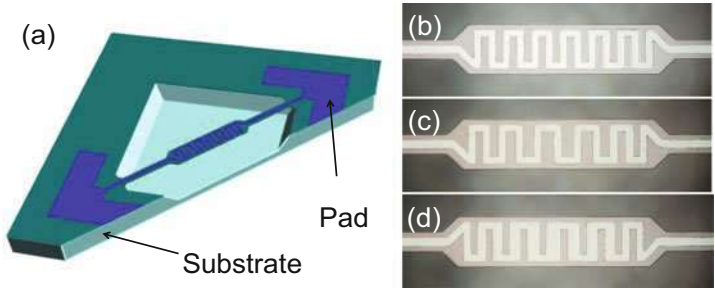
and easy to obtain good temperature distribution just by redesigning the geometry of Pt resistor.

Figure 9 shows the structure design of the two-beam micro-heater. A  $70 \times 270 \text{ um}^2$  rectangular membrane is connected to two suspension beams of  $30 \times 200 \text{ um}^2$  by two trapezoid elements of  $1000 \text{ um}^2$ . A serpentine track of Pt resistor on the rectangular membrane serves as the heater. For thermal isolation, a silicon pit under the suspension structure was formed by wet-chemical anisotropic etching. Based on the above rectangular micro-heater, three resistor geometries are designed and shown in Fig. 9b, c, d. The difference is the width and spacing of the Pt wires.

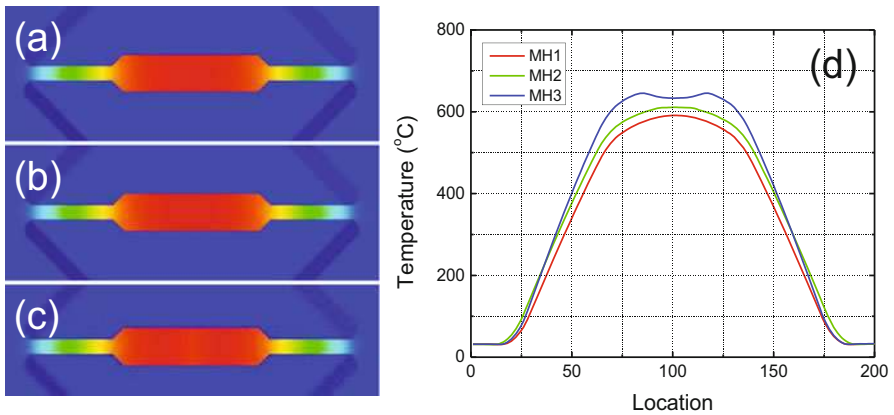
In the original micro-heater (MH1), shown in Fig. 9b, the width and spacing are both 10  $\text{um}$ . This design tended to suffer from an uneven temperature distribution across the active area. This will sequentially cause an uneven reaction and hence reduce the performance of catalytic gas sensors. For better temperature distribution, the original design has been optimized by changing the width or spacing of the Pt wires. In MH2, the width is 10  $\text{um}$ , with spacing linearly decreasing from the center to both ends of the hot plate. In MH3, the spacing is 10  $\text{um}$ , while the width of Pt wire linearly decreases from the center to both ends of the micro-heater. Resistances of the three micro-heaters were measured at room temperature as MH1,  $R_1 = 25.7 \text{ }\Omega$ ; MH2,  $R_2 = 24.4 \text{ }\Omega$ ; and MH3,  $R_3 = 21.3 \text{ }\Omega$ .

The FEM analysis is carried out by Coventor. Figure 10 shows the results of temperature distribution obtained by electrothermomechanical analysis. It suggests that the peak in MH1 is sharper than that in MH2, and also, there are two peaks in MH3.

The active area thermal distribution of the devices has been investigated using an infrared microscope Vistec 100. Figure 11 shows a series of infrared photos of the three heaters under different voltages. In MH1, a central hotspot was observed, and the thermal profile is almost the same as predicted by the simulation results. Thermal profile for MH2 showed better thermal homogeneity than MH1. More so, it indicates



**Fig. 9** The design of two-beam micro-heaters: sectional view of the chip (a), MH1 (b), MH2 (c), and MH3 (d)



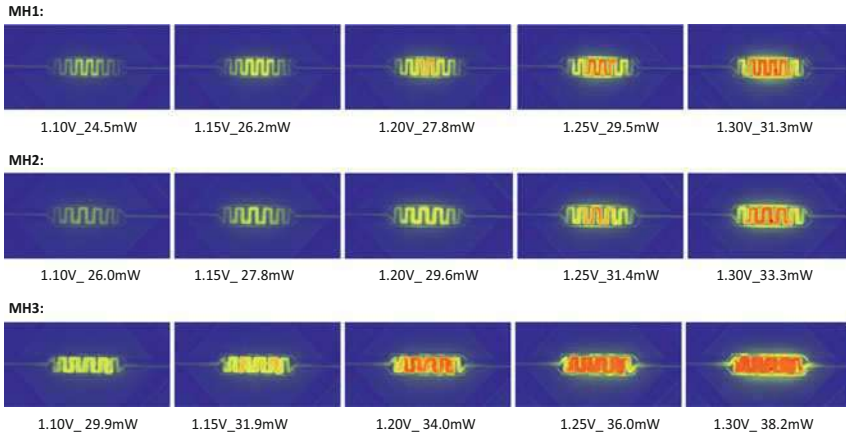
**Fig. 10** Temperature distribution obtained by Coventor: MH1 (a), MH2 (b), MH3 (c), and the temperature along the rectangular hot plate (d)

that the design of MH3 has successfully reduced the hot spot and created a larger isothermal active area.

Three micro-heaters with different arrangements of Pt heater supported by a rectangular hot plate with the same size were fabricated by micromachining technology. By increasing the width of the central heating electrodes, temperature distribution of a rectangular micro-heater can be effectively improved.

### 3D Micro-heaters

The most attractive advantage of micro-heater for the application in gas sensors is its low power consumption. A well-known method for power reduction is to reduce heat loss through heat conduction by using slender beams with high length-to-width ratio (Lee et al. 2008; Xu et al. 2011a; Elmi et al. 2008). However, the increase in length-to-width ratio of the beams usually leads to low mechanical strength.



**Fig. 11** Infrared photos of the two-beam micro-heaters

Another potential way for further power reduction is to decrease active area which can reduce the heat loss via heat convection and radiation. Since most (bio)chemical sensors have better sensing mechanisms with larger active area, sensitivity decreases when using a small active area. Researchers have to face this dilemma and try to keep a balance among the power consumption, mechanical strength, and sensitivity.

## Fabrication Process

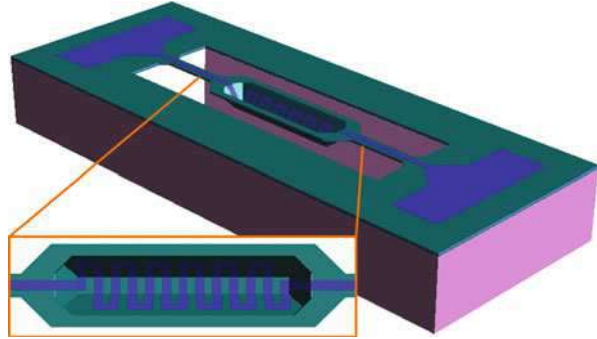
In order to realize micro-heaters with low power consumption and large active area for (bio)chemical sensing applications, we demonstrated a novel 3D micro-heater with a special designed concave membrane.

Figure 12 shows the design and structure of the 3D micro-heater. The structure is quite simple; it has a concave-shaped membrane with a Pt heater embedded in it, which is supported by two beams. For thermal isolation, the whole structure is suspended over a silicon pit. Twice of wet-chemical anisotropic etching were used in the fabrication process. One is for forming the concave-shaped membrane, and the other is for releasing the whole structure.

The fabrication process of 3D micro-heater mainly includes film deposition, photoresist-spray-coating-based lift-off process, deep reactive-ion etching (DIRE), reactive-ion etching (RIE), and wet-chemical anisotropic etching (Xu et al. 2011b, 2012). The fabrication process steps are shown in Fig. 13 and detailed as follows:

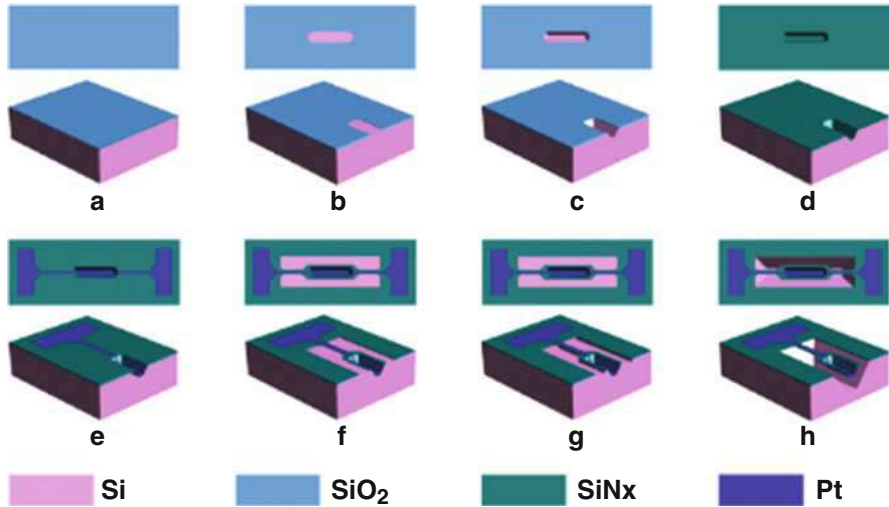
- (a) The fabrication process started from a 4-inch n-type (100) silicon wafer with a resistivity of 1–10  $\Omega$  cm. Then a layer of 300-nm-thick silicon oxide was thermally grown on the front and back side of the wafer.

**Fig. 12** The design and structure of a 3D micro-heater



- (b) A rectangular opening with the four right angles cut was formed by RIE. The four sides of the rectangular opening are Si  $\langle 110 \rangle$  oriented.
- (c) Then, the first anisotropic etching was processed. A concave cavity with the depth of 30–40  $\mu\text{m}$  was formed by KOH etching through the rectangular etching opening.
- (d) Then a concave-shaped dielectric membrane was formed by deposition of a layer of 300-nm-thick silicon oxide and a layer of 1000-nm-thick low-stress silicon nitride. Due to the four missing right angles of the rectangular opening, four suspended triangle membranes were also formed to enhance the robustness of the concave-shaped membrane.
- (e) Using a commercial spray-coating machine EVG-101, a layer of 4- $\mu\text{m}$ -thick photoresist was spray-coated on the surface of both the concave dielectric membrane bottom/sidewalls and the top dielectric membrane on the substrate. Then a 200-nm-thick Pt heating resistor and bonding pads with adhesive layer buried beneath Pt were formed by lift-off process. Due to the gap between the cavity and the Cr mask, a deviation of the pattern size was observed in the lithography. The width of the pattern size increases with the increasing of the height of the gap. The width of the Pt resistors on the bottom of the cavity is about 10% wider than those on the surface of the beams.
- (f) The stacked film of  $\text{SiO}_2/\text{SiN}_x$  was etched by RIE to define the active area and suspension beams with two etching openings formed at the same time.
- (g) Then the silicon substrate was etched by DRIE with a depth as about 25  $\mu\text{m}$  so as the two supporting beams and the concave-shaped membrane can be released by wet-chemical anisotropic etching.
- (h) Finally, the second anisotropic etching was processed. The whole structure was released using a TMAH solution at 80  $^\circ\text{C}$ .

Figure 14a shows the SEM image of the fabricated 3D micro-heater. Figure 15 shows the average temperature of the 3D micro-heater versus power consumption. It can reach up to 400  $^\circ\text{C}$  at a power of 30 mW. By applying an electrical power of 60 mW, the 3D micro-heater can be brought to emit a red glow, shown in Fig. 14c.



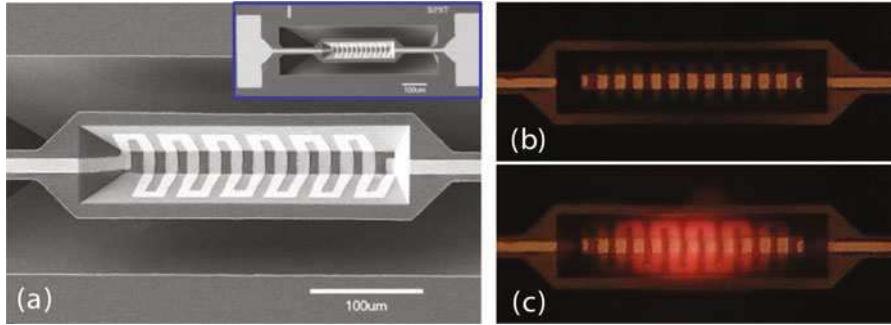
**Fig. 13** Fabrication process of the 3D micro-heater

### Comparison of 3D Micro-heater and 2D Micro-heater

Table 1 shows the comparison between the 3D micro-heater (Xu et al. 2011b) and three other 2D micro-heaters (Lee et al. 2008; Xu et al. 2011a; Elmi et al. 2008). The size of the 3D micro-heater presented is bigger than the comparative 2D micro-heaters. And the cross-sectional area of the suspension beam is also larger than the other three. These are the reasons why the power consumption is high among the four micro-heaters. Heat efficiency of a micro-heater cannot be determined by low power consumption. The heat efficiency of micro-heaters is determined by power and area. If a micro-heater can heat a large active area with a relatively lower power, the micro-heater has a better capability of heat isolation.

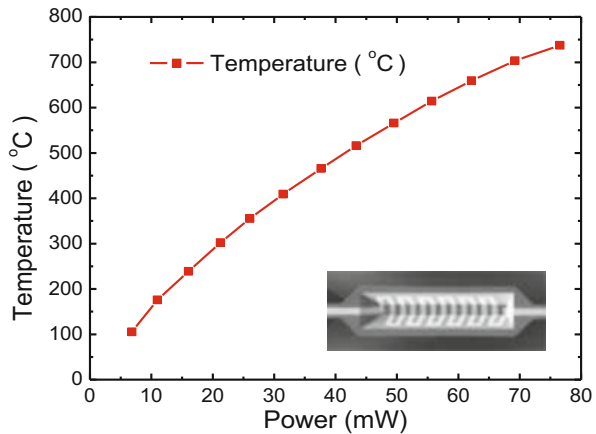
In order to evaluate the heating capability, power per active area (PPAA) is introduced. PPAA is defined as the ratio of electrical power to active area when heated to a certain temperature (typically 400 °C). Table 1 also indicates that PPAA of the 3D micro-heater is only half or less of that in the other 2D micro-heaters.

Unlike the 2D micro-heaters with flat membrane, the heating resistor of this 3D micro-heater is embedded in a concave membrane. By concentrating heat in the concave membrane, heat loss via convection can be reduced. Due to its concave structure, the velocity of airflow will be reduced and the rate of heat loss also decreased. Therefore, convection coefficients of the membranes in the concave are lower than that of an open plat membrane. Furthermore, with such concave membrane, this 3D micro-heater feathers small effective radiation area; therefore, heat loss through radiation will be diminished.



**Fig. 14** Pictures of the 3D micro-heater: SEM image of the fabricated device (a); optical photo of the 3D micro-heater without electrical power (b); optical photo of the 3D micro-heater under an electrical power of 60 mW (c)

**Fig. 15** Temperature versus power consumption of the 3D micro-heater





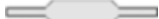

### FEM Analysis for 2D/3D Micro-heaters

FEM analysis has been done by Coventor. Two models were built: the 3D micro-heater and a comparative 2D micro-heater with the same active area. Some parameters of thin films are different from those in bulk materials (Simon et al. 2001; Lee et al. 2003). The parameters and dimensions used in this simulation can be found in the previous work (Xu et al. 2011b).

Temperature distribution of the 3D micro-heater and the comparative 2D micro-heater with the same size of the active area has been simulated under the same power consumption as shown in Fig. 16.

Figure 16a, b shows the simulation results with all heat losses through conduction, convection, and radiation. It can be seen that the 3D micro-heater reaches higher temperature than the comparative 2D micro-heater. In other words, if heating the same size of active area to the same temperature (typically 400 °C), the 3D micro-heater consumes lower power than the 2D micro-heater. It suggests that the 3D micro-heater

**Table 1** Dimensions of four micro-heaters with different structures

Micro-heaters	3D micro-heater with a concave active area presented in this letter (Xu et al. 2011b)	2D micro-heater with a square active area in reference (Lee et al. 2008)	2D micro-heater with a rectangular active area in reference (Xu et al. 2011a)	2D micro-heater with a circular active area in reference (Elmi et al. 2008)
Schematic drawing of heated membranes and suspension beams				
Active area ( $\mu\text{m}^2$ )	46,960	10,000	17,070	4780
Power @ 400 °C (mW)	30	12	18	9
Power per active area (PPAA) ( $\text{mW}/\mu\text{m}^2$ )	6.39e-4	1.2e-3	1.05e-3	1.88e-3

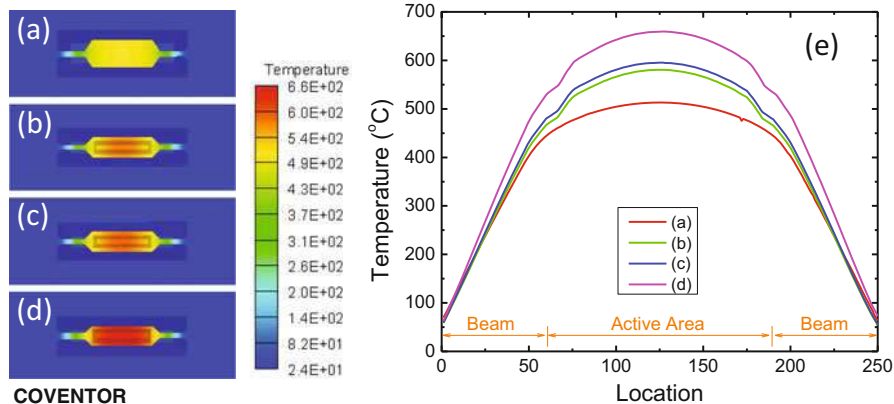
with concave active area features low PPAA. By neglecting the heat loss via radiation (Fig. 16c) and the heat losses via both air and radiation (Fig. 16d), simulation results suggest that heat losses via air and radiation, especially the former one, are significant in a small device with high temperature. So power consumption can be further reduced, if heat losses through air and radiation are effectively decreased.

A 3D micro-heater with Pt heating resistor embedded in a concave active area has been fabricated and tested. By decreasing heat loss via thermal convection and radiation, the 3D micro-heater showed lower power per active area (PPAA) compared to current 2D micro-heaters. With such a novel structure, micro-heater with low power consumption and large active area can be achieved.

## Catalytic Gas Sensors

Catalytic gas sensors, which often referred as catalytic pellistor sensors, have been widely used for detecting combustible gases and vapors in air for several decades (Hierlemann et al. 2003). It is sensitive to most combustible gases and is suitable for the measurement of % LEL (lower explosive limit) on a wide range of combustible gases. Conventional catalytic gas sensors usually consist of a detector and a compensator. The detector has a catalytic surface constructed around a platinum coil that heats the catalyst to a sufficiently high temperature, at which any flammable gas molecules present can burn and release combustion heat (Moseley 1997), while the compensator consists of a coil of fine platinum wire embedded in a bead of alumina which is treated to inhibit oxidation. The platinum coil not only acts as a heater, but also as a temperature





**Fig. 16** Temperature ( $^{\circ}\text{C}$ ) distribution obtained by FEM simulation: the comparative 2D micro-heater with all heat losses involved (a), the 3D micro-heater with all heat losses involved (b), the 3D micro-heater with heat loss through radiation neglected (c), the 3D micro-heater with heat losses through air and radiation neglected (d), node temperature across the beam-active area-beam region (e)

sensor. Gas concentrations can be detected by monitoring resistance change of the Pt coil resulting from temperature increases produced by combustion heat.

With the development of wireless sensor networks, low-power gas sensors are in great demand. However, power consumption of current commercial catalytic gas sensors is usually in a range from 150 to 600 mW. In order to reduce power consumption, instead of platinum wire coils, micro-heaters have been introduced to supply heat in those gas sensors (Advani and Jordan 1980; Sasahara et al. 2005; Briand et al. 2000). In this section, several micro-heater-based catalytic gas sensors will be introduced with structure design, FEM analysis, and fabrication process.

## Sensing Principle

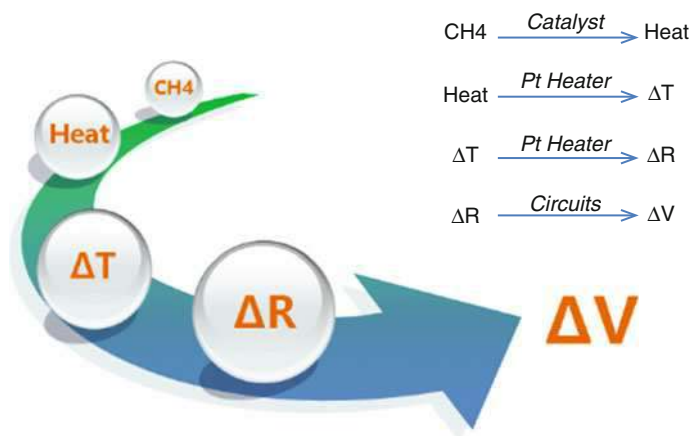
Catalytic gas sensors detect flammable gases by monitoring resistance change of the Pt heater resulting from temperature increase produced by combustion heat. The detecting mechanism can be divided into four sub-steps as shown in Fig. 17.

### Step 1.

By applying electric voltage, the Pt heater reaches a high temperature, normally  $400^{\circ}\text{C}$  to  $600^{\circ}\text{C}$ . Combustible gases (typically methane) mixed with air burn and induce heat with the help of catalyst. In this step, chemical energy is transformed into thermal energy.

### Step 2.

Part of the combustion heat is lost to the environment through conduction, convection, and radiation, while part of the combustion heat transfers into Pt heater and causes a temperature increase ( $\Delta T$ ) of the Pt heater.



**Fig. 17** Sensing mechanism of catalytic gas sensors

Step 3.

Due to the nature of platinum, resistance of Pt heater increases ( $\Delta R$ ) with the increasing of temperature ( $\Delta T$ ).

Step 4.

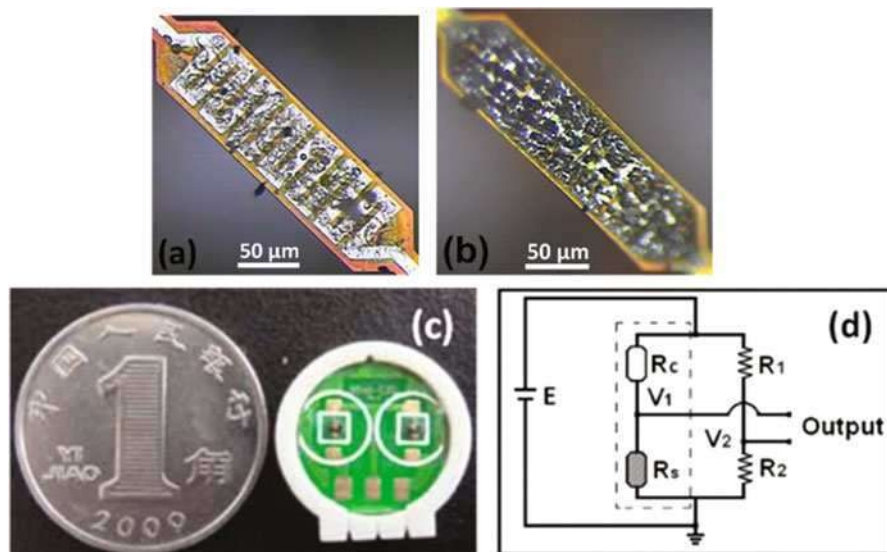
By a Wheatstone bridge circuit, the resistance increase ( $\Delta R$ ) then induces a voltage increase ( $\Delta V$ ). Thus, flammable gases can be detected by measuring the voltage change.

## Fabrication Process

The sensing material is synthesized by sol-gel process. By introducing catalytic metal (15 wt.% Pd-Pt) supported by  $\gamma$ -Al<sub>2</sub>O<sub>3</sub> as sensing material and  $\gamma$ -Al<sub>2</sub>O<sub>3</sub> as compensation material, palladium and platinum chloric acid mixed with alumina aerogel and alumina aerogel were spin coated separately on micro-heaters. After electrical heating at 650 °C for 5 min in air, the alumina aerogel was sintered into mesoporous  $\gamma$ -Al<sub>2</sub>O<sub>3</sub> support, while high activity palladium and platinum particles were also formed.

## 2D Catalytic Gas Sensors

Figure 18a, b shows a compensator and a detector based on two-beam micro-heaters. Methane (CH<sub>4</sub>) is a typical flammable gas which should be monitored under industries especially the mining industry. Figure 19a shows the original output signal of sensor response to methane with different concentrations. Response time of the catalytic gas sensor is less than 10 s. The sensitivity to 50% LEL CH<sub>4</sub> is 2.4 mV/ % CH<sub>4</sub> which is higher than that of four-beam gas sensor. Linear fit of the sensor output



**Fig. 18** Catalytic gas sensors based on two-beam micro-heaters: compensator (a), detector (b), PCB with two elements (c), bridge circuits for test (d)

to methane with different concentrations is shown in Fig. 19b. Output signal has a fairly linear relationship to methane concentration. And the sensor has a 2 mV output voltage with a good signal-to-noise ratio (SNR) to 10% LEL  $\text{CH}_4$ .

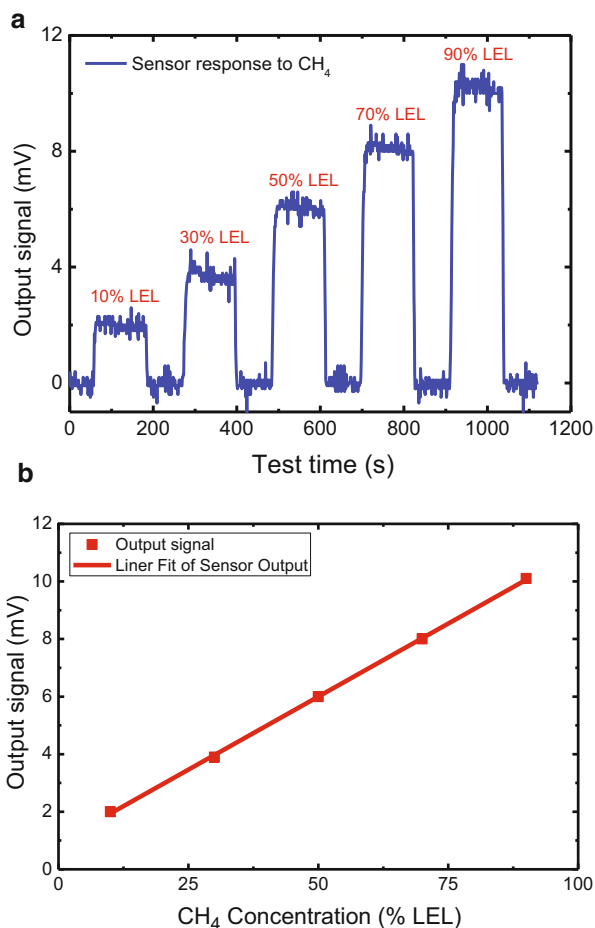
### 3D Catalytic Gas Sensors

The 3D micro-heater presented in section “3D Micro-heaters” has higher heating efficiency than current 2D micro-heaters. Catalyst and porous alumina can be synthesized right into the concave active area by sol-gel process. The size-increased active area with more catalyst loaded can generate more heat during combustion reaction. More so, the concave-shaped structure can help increase the utilization of combustion heat by reducing heat loss via convection and radiation.

Thermal analysis has been carried out by ANSYS. Figure 20 shows three models of micro-heaters, H1, H2, and H3. H1 is a common 2D micro-heater with a rectangular active area supported by two beams. H2 is also a rectangular micro-heater with the same length but larger width than that of H1. The size of active area in H2 is 1.4 times of that of H1. Sensitivity of H2 should be also about 1.4 times of that of H1, if catalyst loaded on the active areas have the same catalytic efficiency. H3 is the 3D micro-heater with a concave-shaped active area whose size is as large as H2. How about the relation of sensitivities between H2 and H3?

Details about the analysis are described in the previous work (Xu et al. 2012). Figure 21 shows the temperature distribution of the sensors. Temperature increase of

**Fig. 19** Sensor response to methane of the two-beam gas sensor: output signal (a) and linear fit of output signal (b)

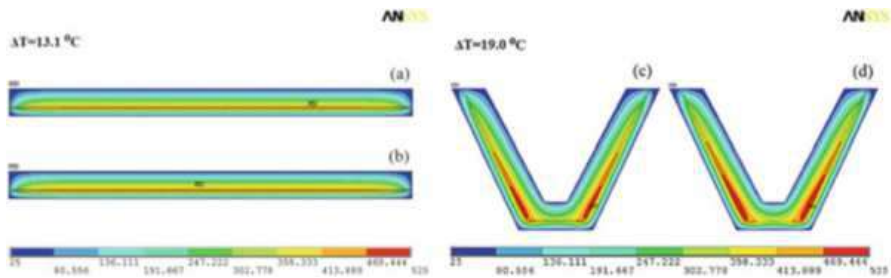
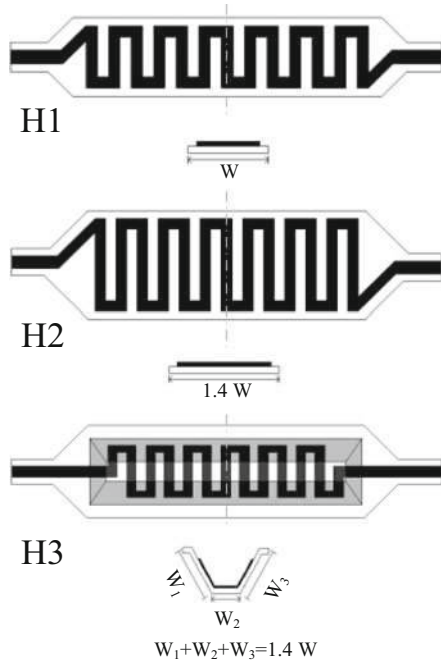


Pt heater in H2 is 13.1 °C. Temperature increase of Pt heater in H3 is 19.0 °C which is 1.45 times that of H2. This indicates that the concave-shaped active area can concentrate heat, making better utilization of combustion heat. If we assume that the increase of combustion heat is proportional to the increase of active area, sensitivity of H2 should be 1.4 times of that of H1. Therefore, the sensitivity of H3 is about twice of that of H1.

It is discovered that, in step 1, heat generation added in H3 is only 64% of that in H2 when heating the Pt heater to 400 °C, which predicts power consumption of H3 could be lower than H2. In other words, H3 has higher heating efficiency than H2; therefore, the higher the heating efficiency of the micro-heaters, the higher the sensitivity of the sensors.

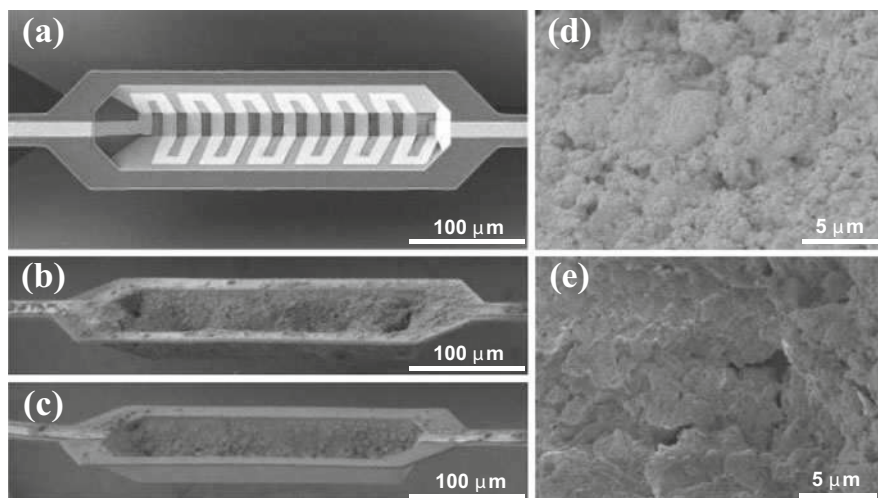
The sensing material of 3D catalytic gas sensor was also developed by sol-gel process (Xu et al. 2012) and shown in Fig. 22. The package of the sensor is shown in Fig. 23.

**Fig. 20** Schematic drawing of three kinds of micro-heaters. *H1*, *H2*, rectangular 2D micro-heaters with the same length but different widths, *H3*, the 3D micro-heater



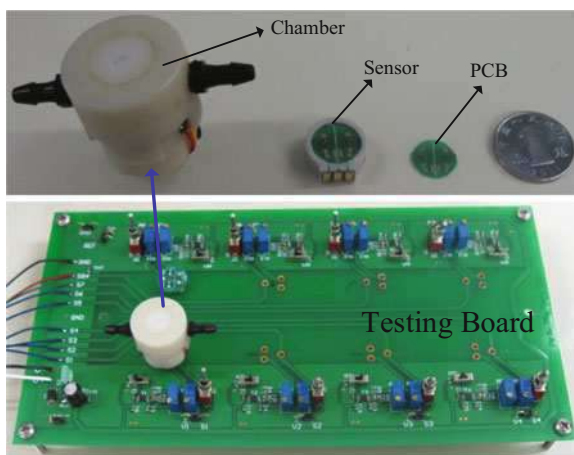
**Fig. 21** Results of the thermal analysis: average temperature of the Pt heater in H2 reaches up to 400 °C by adding HGEN 1 (a), HGEN 2 to the catalyst/alumina in H2 and record the temperature increase ( $\Delta T$ ) of the Pt heater (b), average temperature of the Pt heater in H3 reaches up to 400 °C by adding HGEN 1 (c), add HGEN 2 to the catalyst/alumina in H3 and record the temperature increase ( $\Delta T$ ) of the Pt heater (d)

Figure 24 shows the original signal of sensor response to methane from 10% LEL to 90% LEL and the linear fit of output voltage to methane concentration. The gas sensor was sensitive to 10% LEL  $\text{CH}_4$  with an output voltage as 12 mV. Output voltage to 50% LEL  $\text{CH}_4$  is 30 mV. Output voltage has a good linear relation to methane concentration.



**Fig. 22** Catalytic gas sensors based on 3D micro-heaters: SEM images of the 3D micro-heater (a), detector (b), compensator (c), Pd catalyst (d), porous  $\text{Al}_2\text{O}_3$  (e)

**Fig. 23** The chamber, sensor, and testing board



### Comparison of 3D and 2D Catalytic Gas Sensors

In order to make a comparison, we fabricated two catalytic gas sensors (only detectors) based on a 3D micro-heater and a comparative 2D micro-heater separately. The schematic drawing of the 3D micro-heater and the comparative 2D micro-heater is shown in Fig. 25. The 3D micro-heater has a concave membrane with a rectangular opening as large as the rectangular membrane of the 2D micro-heater. A common catalyst Pd-Pt is introduced and synthesized in/on the active areas of the micro-heaters by sol-gel process.

**Fig. 24** Sensor response to methane of the 3D micro-heater-based catalytic gas sensor: the original output signal to methane from 10% LEL to 90% LEL (a) and linear fit of output voltage to methane concentration (b)

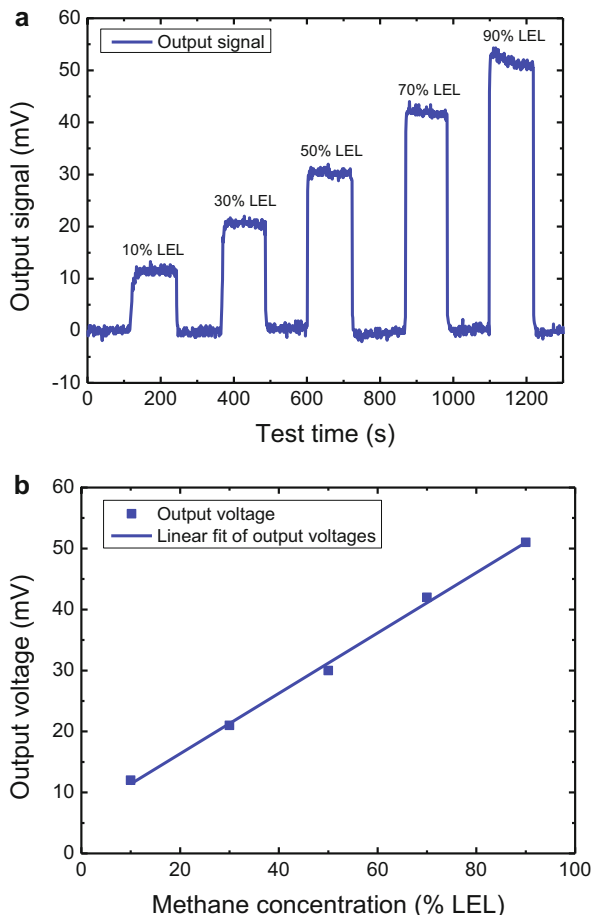
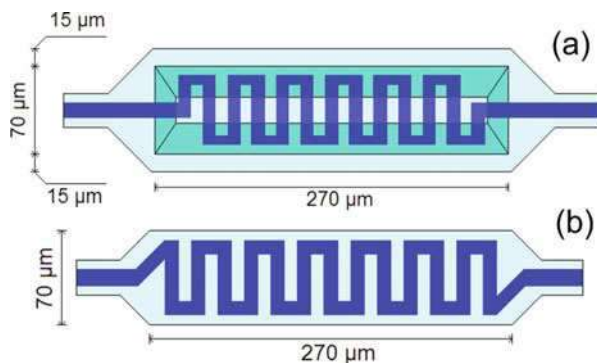


Figure 26 shows the detectors based on the 3D micro-heater and the 2D micro-heater. And the average thickness of the catalyst is shown in Table 2.

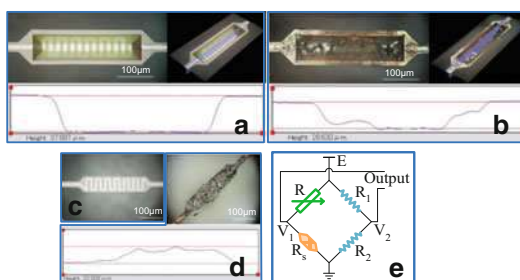
Figure 27a shows the sensor response to 50% LEL  $\text{CH}_4$  at temperature of 400 °C. At the initial stage, there is no gas flowing into or out of the chamber. And the output signal is set as zero by adjusting the resistance of R to that of the sensors. When air flowed through the chamber, heat loss to air by convection increased. Therefore, resistance of sensor decreased; hence, a negative output signal is observed. When 50% LEL  $\text{CH}_4$  mixed in air flowed through the chamber, resistance of the sensor increased because of the heat generated by catalytic combustion. Thus a positive output signal is observed.

From the inset of Fig. 27a, we can see that the negative output signal of the 2D micro-heater-based gas sensor is larger than that of the 3D micro-heater-based gas sensor. So, the 2D micro-heater-based sensor is more influenced by airflow. More so,

**Fig. 25** Schematic drawing of the 3D micro-heater (a) and the comparative 2D micro-heater (b)



**Fig. 26** Catalyst coating process and testing circuit: the bare 3D micro-heater (a), the 3D micro-heater with catalyst (b), the bare 2D micro-heater (c), the 2D micro-heater with catalyst (d), and the bridge circuit for gas detection (e)



the output signal of the 2D micro-heater-based sensor has a lower SNR (signal-to-noise ratio) over the 3D micro-heater-based sensor, shown in Table 2.

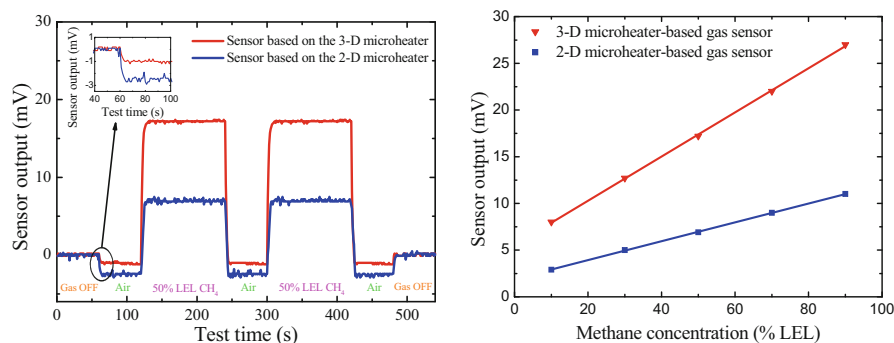
At the working temperature of 400 °C, when methane molecules are absorbed and burnt, combustion heat will be released on the surface of catalyst with the same catalytic efficiency. Firstly, the actual active area of the 3D micro-heater is about 1.4 times of that of the 2D micro-heater. More combustion heat is released inside the concave membrane. Secondly, due to the unique capability of concentrating heat, the 3D micro-heater made better use of the combustion heat over the 2D micro-heater. Thus, temperature increase ( $\Delta T$ ) of the 3D micro-heater-based gas sensor is higher over the 2D micro-heater-based gas sensor, listed in Table 2. As a result, sensitivity of the 3D micro-heater-based gas sensor is about 2.5 times of that of the 2D micro-heater-based gas sensor.

Sensor response to different concentrations of methane is shown in Fig. 27b, although both of the sensors have a fairly linear relation to methane concentration. Hence, the slope of the output signal of the sensors based on 3D micro-heater is higher. In the 3D micro-heater, catalytic combustion reaction occurs inside the concave membrane, while it is on the surface of the sensor based on the 2D micro-heater. Heat lost to air of the 2D micro-heater-based sensor will be much easier, especially at a high methane concentration. Test results indicate that this 3D micro-heater has a capability of concentrating the combustion heat, which in turn improves the overall performance of the catalytic gas sensor.



**Table 2** Characteristics of the catalytic gas sensors based on the 3D micro-heater and 2D micro-heater

Sensor type	Average thickness of the catalysts THK ( $\mu\text{m}$ )	Power @ 400 °C P (mW)	Temperature increase of the Pt heater $\Delta T$ ( $^{\circ}\text{C}$ )	$\Delta T/P$ ( $^{\circ}\text{C}/\text{mW}$ )	Output signal $\Delta V$ (mV)	Sensitivity (mV/% $\text{CH}_4$ )	$\Delta V/P$ (mV/mW)	Response time $T_{90}$ (s)	SNR (dB)
Based on the 2D micro-heater	15	22	3.9	0.18	6.9	2.76	0.31	5.6	8.5
Based on the 3D micro-heater	16	24	8.6	0.36	17.2	6.88	0.72	5.8	18.4

**Fig. 27** Sensor response to methane of the detector based on a 2D/3D micro-heater: original output signal to 50% LEL  $\text{CH}_4$  (a) and sensor response to methane with different concentrations (b)

## Semiconductor Gas Sensors

Due to the remarkable sensitivity and large detection range to a variety of gases, metal oxide semiconductor (MOS) gas sensors have been extensively researched in recent decades for the application in domestic, industries, aerospace, and military (Collins et al. 2000; Liu et al. 2011; Yang et al. 2010; Hoshi and Lahiri 2004).

Since it was first proposed in 1962, tin dioxide ( $\text{SnO}_2$ ) has undergone extensive research and development for gas sensors (Seiyama et al. 1962). Due to the high sensitivity, low operating temperature, and low cost, it has become the dominant choice for solid-state gas sensors. As an n-type semiconductor, conductivity of tin dioxide increases in the presence of reducing gases and decreases in the presence of oxidizing gases.

In this section, we present semiconductor gas sensors based on micro-hot platform with small size, low power, fast response, and high sensitivity. Wafer-level fabrication process of micro-/nano-gas sensors will also be introduced.

## Sensing Principle

When oxygen from air adsorbed onto the surface of  $\text{SnO}_2$  particle, electrons from the surface of  $\text{SnO}_2$  are transferred to the adsorbed oxygen. Therefore, an electron-depleted region, also called the space-charge layer, is formed near the surface of  $\text{SnO}_2$  particle (Miller et al. 2005). Depending on the temperature of the sensor,  $\text{O}_2^-$  was adsorbed at lower temperature (below 175 °C), and  $\text{O}^-$  and  $\text{O}_2^-$  are adsorbed at higher temperatures (above 175 °C) (Barsan and Weimar 2001). So the control of working temperature is critical for the performance of gas sensors.

When exposed to a reducing gas ( $\text{H}_2$ ,  $\text{CO}$ ,  $\text{C}_2\text{H}_5\text{OH}$ ), surface reactions release electrons back to  $\text{SnO}_2$ , leading to a decrease in resistance of the space-charge layer (Barsan and Weimar 2001; Miller et al. 2005). Decreasing the crystallite size can highly improve the sensitivity; however, the small dimensions are difficult to achieve in practical application (Miller et al. 2005; Yamazoe 1991).

## Structure Design

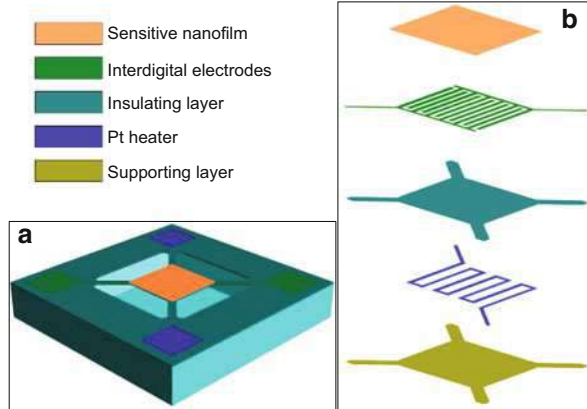
For semiconductor gas sensors, not only micro-heaters are required for appropriate temperature, but also interdigital electrodes (IDEs) are necessary for measuring resistance of sensing film. We named the micro-heater with IDEs as micro-hot platform (MHP).

A suspended-membrane-type MHP is designed for the semiconductor gas sensors as shown in Fig. 28. A square  $\text{SiO}_2$ -supporting membrane (active area) with a platinum heater on it is suspended by four beams over a silicon cavity. An insulating layer of  $\text{SiN}_x$  covers the Pt heater. IDEs are placed on the insulating layer, and sensing materials are deposited on top of the IDEs.

These device offers certain advantages: (i) due to the design, working temperature of the sensor can be well controlled by applying appropriate voltages on the Pt heater; (ii) by isolating heat in the active area, low power consumption can be achieved (approximately 25 mW is sufficient power to heat a sensor to 300 °C); and (iii) because of the low thermal mass, this low-power sensor features fast response time on the order of milliseconds.

Since many semiconductor materials have strong temperature-dependent sensitivity and cross-sensitivity, temperature gradient on the active area will be as small as possible. Figure 29 shows the temperature distribution of the MHP. Temperature gradient on the supporting beams is high, which means that a lot of heat flows through the beams from the center heated membrane to the substrate. The active area achieves a homogenous temperature distribution with an average temperature gradient of about 0.14 °C/ $\mu\text{m}$ .

**Fig. 28** Structure of a micro-hot platform-based semiconductor gas sensor: 3D model of the sensor chip (a) and explosion view of the active area (b)



## Fabrication Process

The fabrication process of the MHP starts from a Si (100) wafer, as shown in Fig. 30.

- (a) A layer of  $\text{SiO}_2$  with a thickness of 350 nm is thermally grown on both sides of the silicon substrate.
- (b) Then a layer of  $\text{SiN}_x/\text{SiO}_2$  with a thickness of 500 nm is deposited by a low-pressure chemical vapor deposition process (LPCVD) on the  $\text{SiO}_2$  layer.
- (c) A layer of Pt/Ti film patterned by lift-off process is used for micro-heater in order to increase the temperature and to better activate the sensing materials.
- (d) An insulating layer of  $\text{SiN}_x$  with a thickness of 600 nm is deposited on the heater layer and patterned by reactive-ion etching (RIE).
- (e) A sensing electrode layer of Pt/Ti is deposited using a sputtering process and then patterned by lift-off process.
- (f) Then the dielectric stacked membrane is selectively etched by a RIE process to form the active area and supporting beams. After that, the whole membrane is released by wet-chemical anisotropic etching using a solution of TMAH.

The fabrication process of sensitive materials on the MHP is shown in Fig. 31. The  $\text{SnO}_2$  ordered porous films are synthesized by a sol-gel method and the colloidal monolayer template, as previously described (Duan et al. 2007).

- (a) Large-area (more than  $1 \text{ cm}^2$ ) ordered polystyrene colloidal monolayer templates, with the sphere diameter of 500 nm, are prepared on a cleaned glass slide by spin-coating method.
- (b) Then the glass slide covered with the colloidal monolayer is immersed into the precursor solution ( $0.1 \text{ M SnCl}_4 \cdot 5\text{H}_2\text{O}$ ), and the colloidal monolayer is then lifted off from the glass substrate and floated on the surface of the precursor solution.
- (c) Then, the floated colloidal monolayer is picked up by a MHP chip and covers the surface of the active area. At the same time, the interstitial space between the monolayer and the substrate is filled with solution.

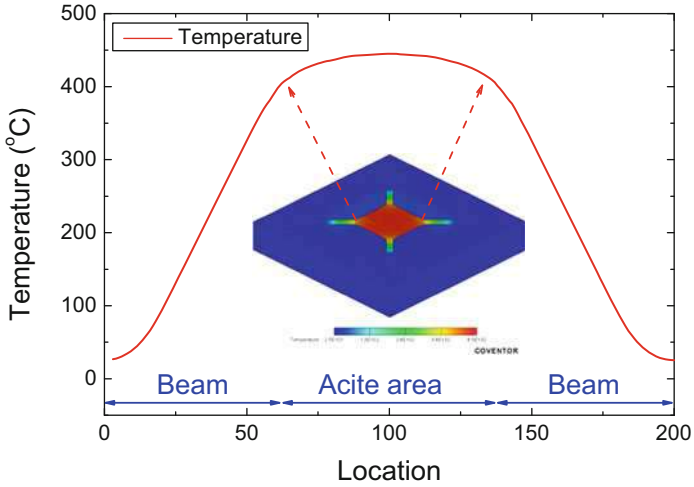


Fig. 29 Temperature distribution of the sensor obtained by electrothermal analysis

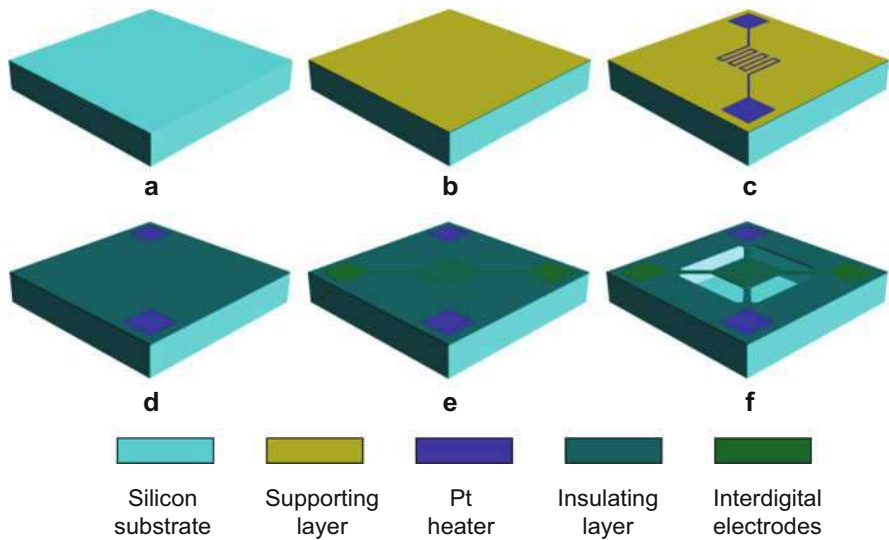
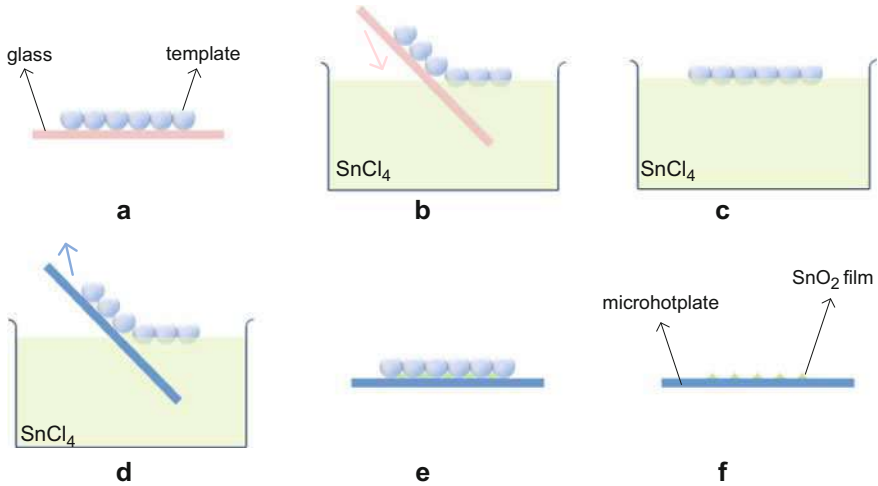


Fig. 30 Fabrication process of micro-hot platform: (a) it was started from a n-type <100> silicon substrate, (b) the supporting layer was deposited by LPCVD, (c) the Pt heater was fabricated by a lift-off process, (d) the insulating layer was deposited by PECVD and patterned by RIE, (e) the interdigital electrodes was fabricated by a lift-off process, and (f) the whole structure was released by wet-chemical anisotropic etching in a TMAH solution



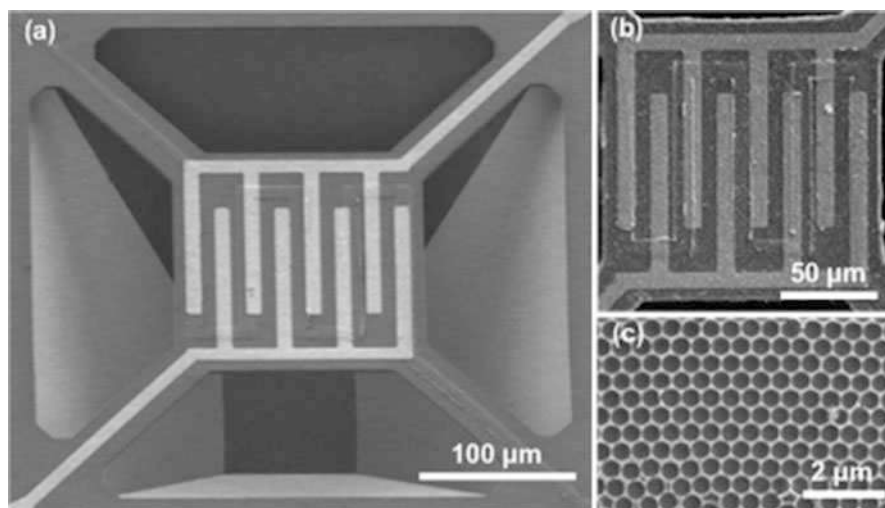
**Fig. 31** Fabrication process of SnO<sub>2</sub> porous nanofilm on the micro-hot platform: (a) monolayer template on a glass substrate, (b) the glass slide covered with the colloidal monolayer was immersed into the precursor solution, (c) the colloidal monolayer floats onto the surface of the solution, (d) using a MHP to pick up the colloidal monolayer, (e) the interstitial space between the monolayer and the MHP was filled with SnCl<sub>4</sub> solution, and (f) ordered porous film was fabricated after removal of template by annealing

- (d) Through drying at a temperature slightly above the glass-transition temperature of PS (120 °C), the solute gradually deposits on the substrate.
- (e) After annealed at 400 °C for one hour to burn off the polystyrene colloidal template, a layer of ordered porous SnO<sub>2</sub> film is thus produced on the MHP.

Figure 32 shows the structure and morphology of the MHP-based semiconductor gas sensor. The MHP has a square active area (150 × 150 μm) with 8 Pt fingers (10 × 100 μm). A monolayer of SnO<sub>2</sub> porous film was in situ synthesized on the active area with a diameter of 500 nm. Stripe-like patterns indicate that the pore array is highly ordered.

## Wafer-Level Fabrication of Micro-/Nano-gas Sensors

Here we propose a novel strategy toward the goal of in situ wafer-level fabrication of high-performance gas sensing chips. The in situ wafer-level fabrication process is to fully combine micromachining process and nanofabrication process and naturally integrate MHP with nano-pore array (NPA). As shown in Fig. 33a, a wafer of MHP fabricated by micromachining process, a wafer of polystyrene (PS) colloidal monolayer template, and a precursor solution for synthesis of gas sensing material are, respectively, prepared in advanced. The template-directed ordered nano-pore arrays will exhibit good advantages in the controllable microstructure, homogenous



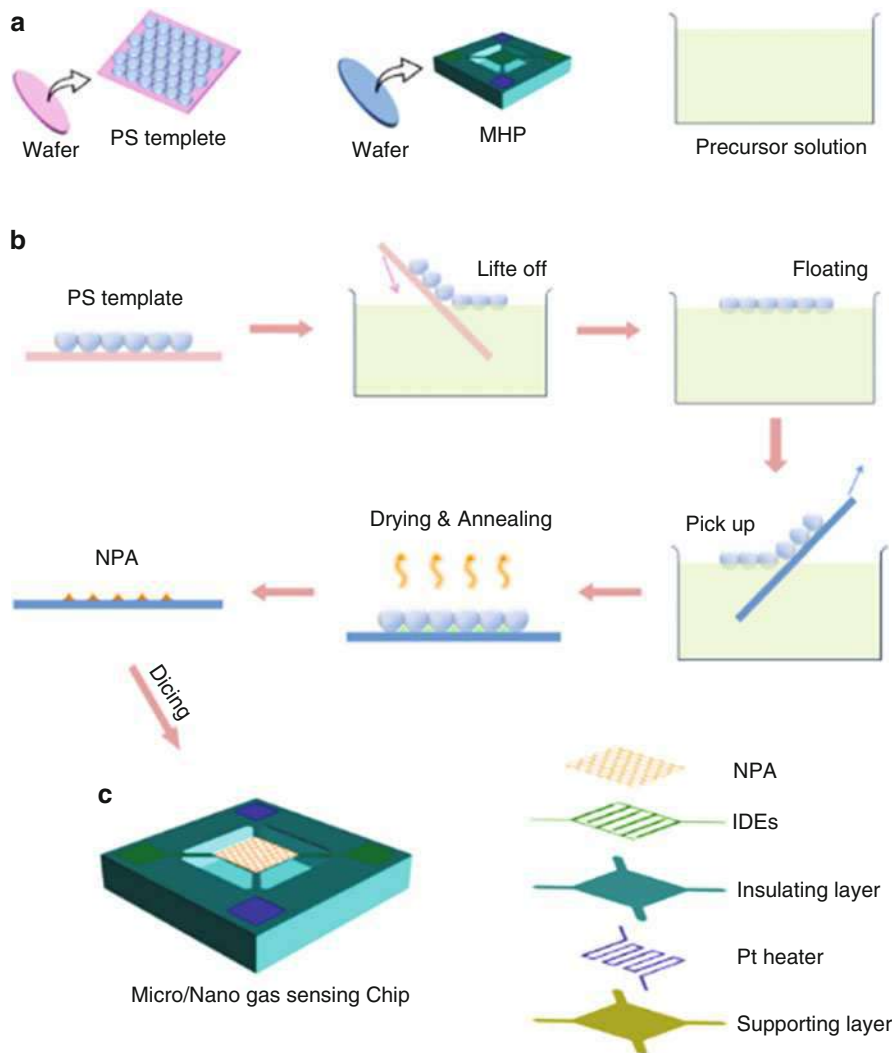
**Fig. 32** SEM images of the MHP (a), the active area of the sensor (b), and the porous nanofilm (c)

thickness, and reproducible fabrication, which is beneficial to fabricate high-performance sensors.

The fabrication process is demonstrated in Fig. 33b. Tin dioxide is taken for an example. First, a self-organized PS-sphere monolayer template (Jia and Cai 2010; Dai et al. 2012, 2013) is lifted off from a glass wafer and then floated on the surface of precursor solution ( $\text{SnCl}_4$ ). Then, such floated monolayer is transferred to the MHP wafer by a simple picking-up process. Due to the capillary effect, the PS monolayer on the wafer also contains the precursor solution in the interstitials among PS spheres and wafer (Duan et al. 2007). Along with subsequent drying and annealing, the organic PS template can be removed, and ordered  $\text{SnO}_2$  NPA is thus formed on the MHP wafer. In this case, a great many MHP-NPA integrated sensors are simultaneously fabricated and arranged on the wafer. After dicing, massive sensor chips can be obtained. Such a semiconductor gas sensor through micro-/nano-integrated in situ wafer-level fabrication process is named as micro-/nano-gas sensor.

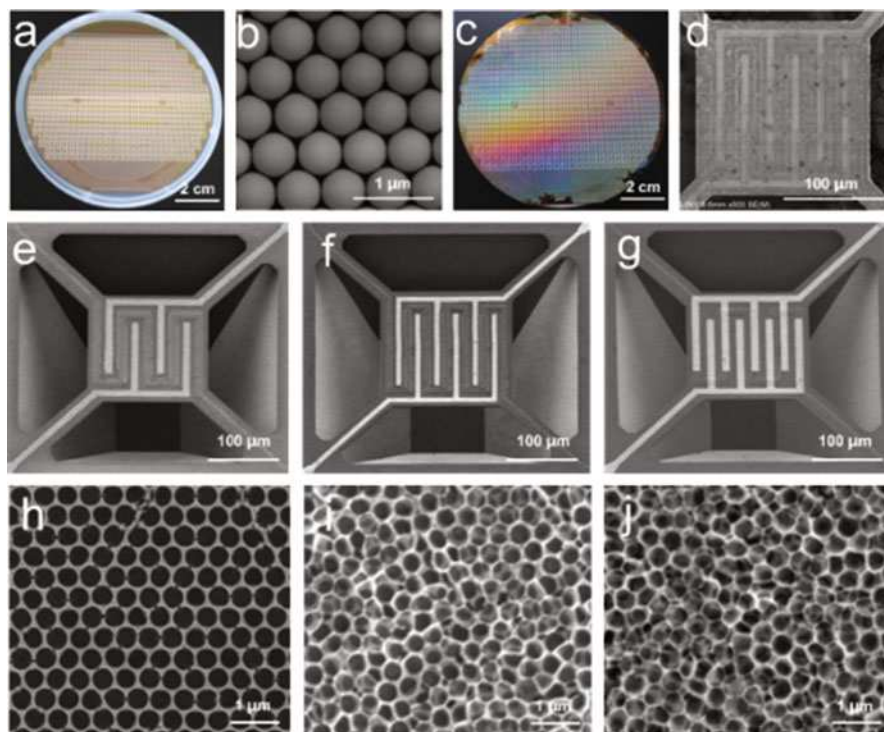
All the process in this fabrication method is under well control. The nanostructure, material, and thickness of NPA are defined by the size of PS template, material, and concentration of the precursor solution. Furthermore, multilayer NPA can be fabricated by repeating the above procedures. In addition, micro-/nano-gas sensing chips with different MHP designs can also be realized in one batch. Most importantly, such strategy makes a reality of the combination of micro- and nano-fabrication and the integration of MPH with NPA for mass production of gas sensing chips.

Figure 34 shows how micro-/nano-gas sensors are fabricated. Based on a wafer of MHPs with three different structures and a wafer of PS template with a diameter of 500 nm, micro-/nano-gas sensors with three kinds of  $\text{SnO}_2$  NPA (monolayer, double



**Fig. 33** The micro-/nano-gas sensors and the strategy of in situ wafer-level fabrication process: (a) a wafer of MHP, a wafer of polystyrene (PS) colloidal monolayer template, and a precursor solution are, respectively, prepared in advanced. (b) In situ synthesis of  $\text{SnO}_2$  NPA on the MHP. (c) Sensor chip and the active area of the micro-/nano-gas sensor

layer, and triple layer) have been fabricated, respectively. The three MHPs (Fig. 34e–g) have the same Pt micro-heater ( $8\ \mu\text{m}$  in width and  $20\ \mu\text{m}$  in spacing) but different IDEs. The spacing of fingers of MHP1, MHP2, and MHP3 are 24, 18, and  $10\ \mu\text{m}$ . By using the same PS template and precursor solution, monolayer, double-layer, and triple-layer  $\text{SnO}_2$  NPA (Fig. 34h–j) are in situ synthesized on the MHP wafers, respectively.



**Fig. 34** Micro-/nano-gas sensors, a 4-inch wafer of MHP; chip size,  $3 \times 3$  mm (a); PS template, diameter of the PS ball is 500 nm (b); the whole 4-inch wafer of MHP covered with a monolayer of NPA (c); active area of the micro-/nano-gas sensing chip (d); MHP1, spacing of fingers is 24  $\mu\text{m}$  (e); MHP2, spacing of fingers is 18  $\mu\text{m}$  (f); MHP3, spacing of fingers is 10  $\mu\text{m}$  (g); monolayer  $\text{SnO}_2$  NPA (h); double-layer  $\text{SnO}_2$  NPA (i); and triple-layer  $\text{SnO}_2$  NPA (j)

## Properties of Micro-/Nano-gas Sensors

Figure 35a shows the 3D plot of the sensitivity ( $S$ ) to 1 ppm ethanol as a function of the working temperature and types of MHP. It clearly indicates that gas sensor has the highest sensitivity at a working temperature of 350  $^{\circ}\text{C}$ . The sensor based on MHP3 has higher sensitivity than other two sensors. MHP3 has narrow gap electrodes that induce a higher collection efficiency of the electrical signal and less influenced by free charges. Besides, MHP3 has more pairs of fingers that can reduce the contact resistance between Pt and  $\text{SnO}_2$ . In addition, nanostructured  $\text{SnO}_2$  has defects such as inactive particles, fractured parts, and  $\text{SnO}_2$  film without nano-pores. These defections may be not sensitive or less sensitive to gases.

Based on MHP3, we then evaluated the properties of sensors with different layers of NPA. Figure 35b shows the sensitivities of sensors with monolayer, double-layer, and triple-layer  $\text{SnO}_2$  NPA. Sensors with double-layer or triple-layer



NPA have lower sensitivity and larger error than sensors with monolayer NPA. The space-charge layer, which is formed near the surface of  $\text{SnO}_2$  particle, can change its resistance when exposed to a reducing gas. Surface reactions release electrons back to  $\text{SnO}_2$ , leading to a decrease in resistance of the space-charge layer. In a sensor with multiple layer of  $\text{SnO}_2$ , gases can hardly penetrate into the layers beneath the top layer of  $\text{SnO}_2$ . Besides, we also found that the quality of the single layer of  $\text{SnO}_2$  is better than that of two or three layers, shown in Fig. 35h, i, j. Therefore, the relative resistance change of a single layer of  $\text{SnO}_2$  is more significant than that of multiple layers of  $\text{SnO}_2$ . Sensitivity of a single layer is higher than that of multiple layers. The error comes from fabrication process. The MEMS process of fabricating micro-hot platforms has high yield. While in this work, the solution process was conducted by hands, which leads to lower yield and consistency. Based on the results of these test, sensors on MPH 3 with monolayer NPA and working temperature of  $350^\circ\text{C}$  have high sensitivity.

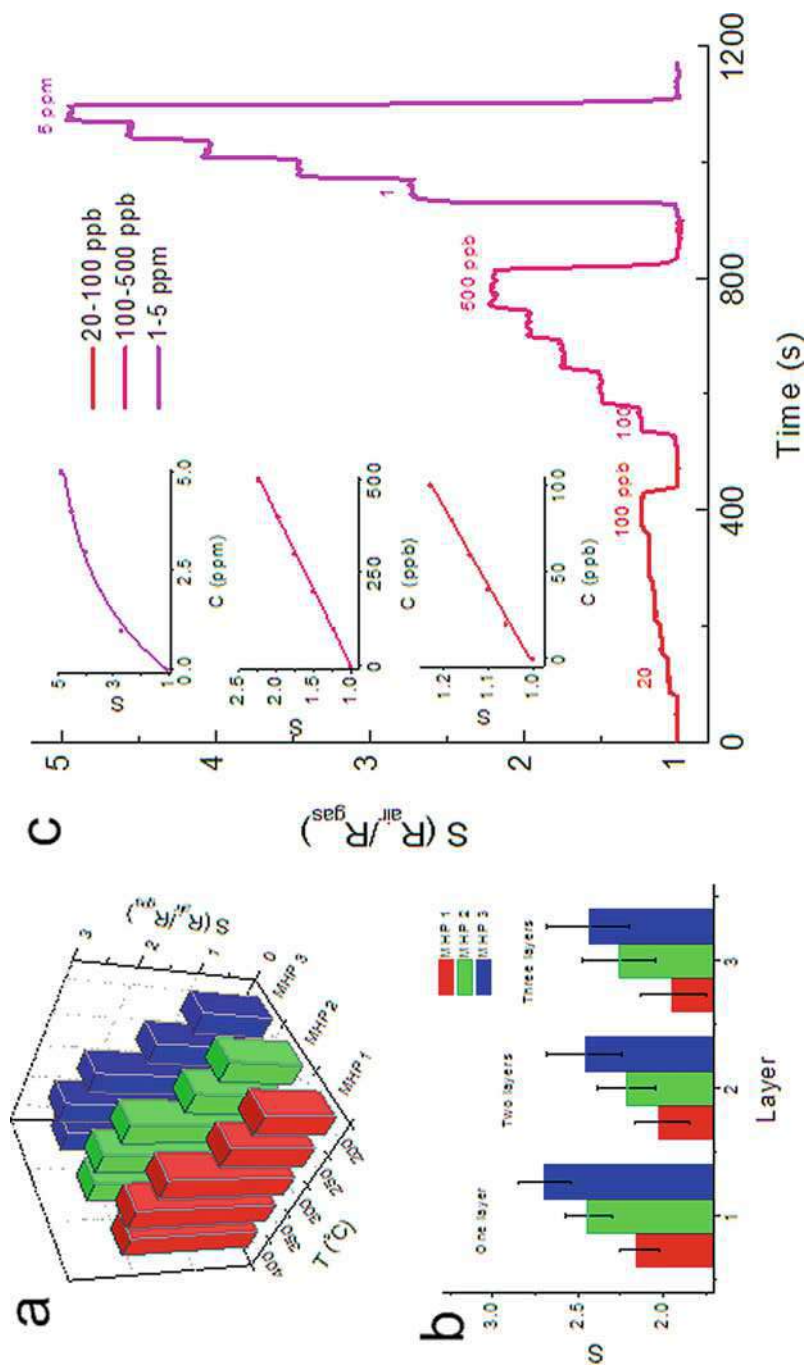
Figure 35c shows the sensor response to ethanol with three different levels of concentrations: from 20 to 100 ppb, from 100 to 500 ppb, and from 1 to 5 ppm. Response time and recover time are defined as the times for the sensor to reach up to 90% of its steady value and back down to 10% of the value, respectively. The response time and recover time to ethanol are around 2 s with a concentration less than 100 ppb and decrease to around 1 s and less than 1 s when the concentration increases to 100–500 ppb and 1–5 ppm, respectively.

In conclusion, a strategy toward in situ wafer-level fabrication process is presented and evaluated for the fabrication of micro-/nano-gas sensing chips with high performance and high throughput. The integration of the micromachining process and the nanofabrication process fully preserves the merits of both MHP and NPA to massively produce thousands of gas sensing chips with high sensitivity (down to 20 ppb), low power consumption (down to 30 mW), and fast response time (down to 1 s). It is also demonstrated that such fabrication process yields high reliability and high repeatability, which enables high throughput and low cost production for practical applications.

---

## Conclusions

In this chapter, the design method and fabrication process of micro-heaters and micro-heater-based gas sensors have been introduced. By structure optimization, thermal simulation, and process control, one can realize high-performance micro-heaters for gas sensors with low power consumption, high mechanical stress, and long lifetime. Catalytic gas sensors based on 2D/3D micro-heaters were presented. And the 3D micro-heater-based gas sensors demonstrated better performance. Semiconductor gas sensors based on micro-hot platform are introduced and are of low power consumption and high sensitivity. Wafer-level fabrication process not only improved the properties of sensors but also decreased the cost.



**Fig. 35** Sensor response to ethanol: the 3D plot of the sensitivity to 1 ppm ethanol as a function of the working temperature and types of MHP (a); the sensitivities of sensors based on MHP 3, with monolayer, double layer, and triple layer of NPA (b); and sensor response to ethanol with three different levels of concentrations: from 20 to 100 ppb, from 100 to 500 ppb, and from 1 to 5 ppm (c)

## References

- Advani GN, Jordan AG (1980) Thin-films of SnO<sub>2</sub> as solid-state gas sensors. *J Electron Mater* 9(1):29–49. <https://doi.org/10.1007/Bf02655213>
- Barsan N, Weimar U (2001) Conduction model of metal oxide gas sensors. *J Electroceram* 7(3):143–167. <https://doi.org/10.1023/A:1014405811371>
- Briand D, Krauss A, van der Schoot B, Weimar U, Barsan N, Gopel W, de Rooij NF (2000) Design and fabrication of high-temperature micro-hotplates for drop-coated gas sensors. *Sensor Actuat B-Chem* 68(1–3):223–233. [https://doi.org/10.1016/S0925-4005\(00\)00433-0](https://doi.org/10.1016/S0925-4005(00)00433-0)
- Collins PG, Bradley K, Ishigami M, Zettl A (2000) Extreme oxygen sensitivity of electronic properties of carbon nanotubes. *Science* 287(5459):1801–1804. <https://doi.org/10.1126/science.287.5459.1801>
- Dai ZF, Li Y, Duan GT, Jia LC, Cai WP (2012) Phase diagram, design of monolayer binary colloidal crystals, and their fabrication based on ethanol-assisted self-assembly at the air/water Interface. *ACS Nano* 6(8):6706–6716. <https://doi.org/10.1021/nn3013178>
- Dai ZF, Xu L, Duan GT, Li T, Zhang HW, Li Y, Wang Y, Wang YL, Cai WP (2013) Fast-response, sensitive and low-powered chemosensors by fusing nanostructured porous thin film and IDEs-microheater chip. *Sci Rep-UK* 3. ARTN 1669. <https://doi.org/10.1038/srep01669>
- Duan G, Cai W, Luo Y, Sun F (2007) A hierarchically structured Ni(OH)<sub>2</sub> monolayer hollow-sphere array and its tunable optical properties over a large region. *Adv Funct Mater* 17(4):644–650. <https://doi.org/10.1002/adfm.200600568>
- Elmi I, Zampolli S, Cozzani E, Mancarella F, Cardinali GC (2008) Development of ultra-low-power consumption MOX sensors with ppb-level VOC detection capabilities for emerging applications. *Sensor Actuat B-Chem* 135(1):342–351. <https://doi.org/10.1016/j.snb.2008.09.002>
- Gotz A, Gracia I, Cane C, Lora-Tamayo E, Horrillo MC, Getino J, Garcia C, Gutierrez J (1997) A micromachined solid state integrated gas sensor for the detection of aromatic hydrocarbons. *Sensor Actuat B-Chem* 44(1–3):483–487. [https://doi.org/10.1016/S0925-4005\(97\)00171-8](https://doi.org/10.1016/S0925-4005(97)00171-8)
- Graf M, Barrettino D, Kirstein KU, Hierlemann A (2006) CMOS microhotplate sensor system for operating temperatures up to 500 degrees C. *Sensor Actuat B-Chem* 117(2):346–352. <https://doi.org/10.1016/j.sub.2005.11.012>
- Guidi V, Cardinali GC, Dori L, Faglia G, Ferroni M, Martinelli G, Nelli P, Sberveglieri G (1998) Thin-film gas sensor implemented on a low-power-consumption micromachined silicon structure. *Sensor Actuat B-Chem* 49(1–2):88–92. [https://doi.org/10.1016/S0925-4005\(98\)00039-2](https://doi.org/10.1016/S0925-4005(98)00039-2)
- Hamann S, Ehmann M, Thienhaus S, Savan A, Ludwig A (2008) Micro-hotplates for high-throughput thin film processing and in situ phase transformation characterization. *Sensor Actuat A-Phys* 147(2):576–582. <https://doi.org/10.1016/j.sna.2008.05.007>
- Hierlemann A, Brand O, Hagleitner C, Baltes H (2003) Microfabrication techniques for chemical/biosensors. *Proc IEEE* 91(6):839–863. <https://doi.org/10.1109/Jproc.2003.813583>
- Hoshi T, Lahiri S (2004) Oxygen sensing: It's a gas! *Science* 306(5704):2050–2051. <https://doi.org/10.1126/science.1107069>
- Jia LC, Cai WP (2010) Micro/nanostructured ordered porous films and their structurally induced control of the gas sensing performances. *Adv Funct Mater* 20(21):3765–3773. <https://doi.org/10.1002/adfm.201001216>
- Kunt TA, McAvoy TJ, Cavicchi RE, Semancik S (1998) Optimization of temperature programmed sensing for gas identification using micro-hotplate sensors. *Sensor Actuat B-Chem* 53(1–2):24–43. [https://doi.org/10.1016/S0925-4005\(98\)00244-5](https://doi.org/10.1016/S0925-4005(98)00244-5)
- Lee SM, Dyer DC, Gardner JW (2003) Design and optimisation of a high-temperature silicon micro-hotplate for nanoporous palladium pellistors. *Microelectron J* 34(2):115–126. [https://doi.org/10.1016/S0026-2692\(02\)00153-2](https://doi.org/10.1016/S0026-2692(02)00153-2). Pii S0026-2692(02)00153-2
- Lee J, Spadaccini CM, Mukerjee EV, King WP (2008) Differential scanning calorimeter based on suspended membrane single crystal silicon microhotplate. *J Microelectromech S* 17(6):1513–1525. <https://doi.org/10.1109/Jmems.2008.2006811>
- Lim JW, Kim HH, Kang BH, Lee DD, Chung WY (1998) Fabrication and characteristics of suspended-type micro gas sensor. *J Korean Phys Soc* 33:S432–S435

- Liu N, Tang ML, Hentschel M, Giessen H, Alivisatos AP (2011) Nanoantenna-enhanced gas sensing in a single tailored nanofocus. *Nat Mater* 10(8):631–636. <https://doi.org/10.1038/Nmat3029>
- Low HM, Tse MS, Chiu MM (1998) Thermal induced stress on the membrane in integrated gas sensor with micro-heater. In: Proceedings of 1998 I.E. Hong Kong Electron Devices Meeting, pp 140–143
- Miller TA, Bakrania SD, Perez C, Wooldridge MS (2005) A new method for direct preparation of tin dioxide nanocomposite materials. *J Mater Res* 20(11):2977–2987. <https://doi.org/10.1557/Jmr.2005.0375>
- Moseley PT (1997) Solid state gas sensors. *Meas Sci Technol* 8(3):223–237. <https://doi.org/10.1088/0957-0233/8/3/003>
- Rossi C, Temple-Boyer P, Esteve D (1998) Realization and performance of thin SiO<sub>2</sub>/SiN<sub>x</sub> membrane for microheater applications. *Sensor Actuat a-Phys* 64(3):241–245
- Sasahara T, Kido A, Ishihara H, Sunayama T, Egashira M (2005) Highly sensitive detection of volatile organic compounds by an adsorption/combustion-type sensor based on mesoporous silica. *Sensor Actuat B-Chem* 108(1–2):478–483. <https://doi.org/10.1016/j.snb.2004.12.118>
- Seiyama T, Kato A, Fujiishi K, Nagatani M (1962) A new detector for gaseous components using Semiconductive thin films. *Anal Chem* 34(11):1502–1503. <https://doi.org/10.1021/ac60191a001>
- Simon T, Barsan N, Bauer M, Weimar U (2001) Micromachined metal oxide gas sensors: opportunities to improve sensor performance. *Sensor Actuat B-Chem* 73(1):1–26. [https://doi.org/10.1016/S0925-4005\(00\)00639-0](https://doi.org/10.1016/S0925-4005(00)00639-0)
- Xu L, Li T, Gao XL, Wang YL (2011a) Development of a reliable micro-hotplate with low power consumption. *IEEE Sensors J* 11(4):913–919. <https://doi.org/10.1109/Jsen.2010.2064765>
- Xu L, Li T, Wang YL (2011b) A novel three-dimensional microheater. *IEEE Electr Device L* 32(9):1284–1286. <https://doi.org/10.1109/Led.2011.2160145>
- Xu L, Wang YC, Zhou H, Liu YX, Li T, Wang YL (2012) Design, fabrication, and characterization of a high-heating-efficiency 3-D microheater for catalytic gas sensors. *J Microelectromech S21(6):1402–1409*. <https://doi.org/10.1109/Jmems.2012.2208219>
- Yamazoe N (1991) New approaches for improving semiconductor gas sensors. *Sensor Actuat B-Chem* 5(1–4):7–19. [https://doi.org/10.1016/0925-4005\(91\)80213-4](https://doi.org/10.1016/0925-4005(91)80213-4)
- Yang DJ, Kamienchick I, Youn DY, Rothschild A, Kim ID (2010) Ultrasensitive and highly selective gas sensors based on electrospun SnO<sub>2</sub> Nanofibers modified by Pd loading. *Adv Funct Mater* 20(24):4258–4264. <https://doi.org/10.1002/adfm.201001251>



# Picogram-Order Mass Sensors via Cantilever-Based Micro-/Nanostructures

Dong F. Wang, Xu Du, Xiaodong Li, Di Zhou, Cao Xia, Guowen Zheng, Shenglai Wan, and Xin Wang

## Contents

Approaching with Mode Localized Oscillation .....	754
Concept .....	754
Overview .....	754
Theory .....	755
Simulation Results and Experimental Verification .....	757
Conclusions and Further Reading .....	767
Approaching with Synchronized Oscillation .....	767
Overview .....	767
Frequency Locking in Coupled Cantilevers .....	770
Synchronization Region .....	771
Phase Noise .....	775
Picogram Mass Sensitivity .....	777
Conclusions and Further Reading .....	782
Summary .....	783
References .....	784

## Abstract

In recent years, MEMS, short for microelectromechanical systems, has drawn increasingly attention for its advantage on mass sensing. MEMS is the integration of mechanical elements, sensors, and electronics on a common silicon substrate through microfabrication technology. These devices replace bulky counterparts

D. F. Wang (✉)

Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, Jilin, China

Research Center for Ubiquitous MEMS and Micro Engineering, AIST, Tsukuba, Japan  
e-mail: [dongfwang@jlu.edu.cn](mailto:dongfwang@jlu.edu.cn)

X. Du · X. Li · D. Zhou · C. Xia · G. Zheng · S. Wan · X. Wang

Micro Engineering and Micro Systems Laboratory, School of Mechanical Science and Engineering, Jilin University, Changchun, Jilin, China

with micron-scale equivalent that can produce in large quantities by fabrication process used in integrated circuits in photolithography. Meanwhile, the devices reduce cost, bulk, weight, and power consumption greatly while increasing performance, production volume, and functionality by orders of magnitude. In particular, cantilever-based sensors fabricated by MEMS technology display high sensitivity as well as resolution in mass sensing.

This chapter reports two methods for picogram-order mass sensing by studying the response in frequency shifts and amplitude changes of a simplified cantilever array. Firstly, the mode localization phenomenon in the linear vibration region of a localized cantilever array is investigated for mass sensing with ultrahigh sensitivity. It is demonstrated by theoretical analysis and experimental verification that eigenstate shifts (amplitude shifts) are three to four times greater than corresponding ones in resonant frequency with a picogram-order mass perturbation. Secondly, the synchronization phenomenon in the nonlinear vibration region of a synchronized cantilever array is studied for mass sensing with high resolution. Synchronization via beam-shaped cantilever array and one U-shaped cantilever coupled with another beam shape cantilever has been researched for different application, respectively. Specifically, frequency multiplication and phase noise suppression of synchronized oscillation are demonstrated via the theoretical analysis and experimental verification of frequency locking and synchronized region for coupled cantilever array, which is beneficial to improve the quality factor and resolution of mass sensing.

---

**Keywords**

Picogram-order mass sensors · Cantilever-based micro-/nanostructures · Coupled array · Mode localization · Amplitude enhancement · Ultrahigh sensitivity · Synchronization · Frequency multiplication · Phase noise suppression · High resolution

---

## Approaching with Mode Localized Oscillation

### Concept

Mode localization is a phenomenon which was first found by Anderson (1958) in periodic systems. It describes a rule that the vibration energy can be gathered into some regions of the system when there exist a small perturbation. In recent years, such a concept has been proposed as a sensing mechanism based on eigenstate shift in coupled micro-cantilevers within various conditions, especially for mass detection on picogram level.

### Overview

In resonant frequency-based sensors, the output corresponds to a shift in the resonant frequency of a vibrating micromechanical structure when subjected to a small

perturbation in either its stiffness or mass. The most sensitive micro-cantilever based on mass detection experiments using the frequency shift approach have reported attogram-level detection in ultrahigh vacuum environment (Ekinici et al. 2004a; Davis and Boisen 2005) and femtogram level detection under ambient conditions (Ilic et al. 2001; Hosaka et al. 2006). Separately, the nano-mechanical properties of single-crystalline silicon cantilevers as resonating elements for ultimate sensing have been also investigated in an ultrahigh vacuum by flash heating (Wang et al. 2003), crystallographic orientation (Wang et al. 2003), as well as gas adsorption (Wang et al. 2004).

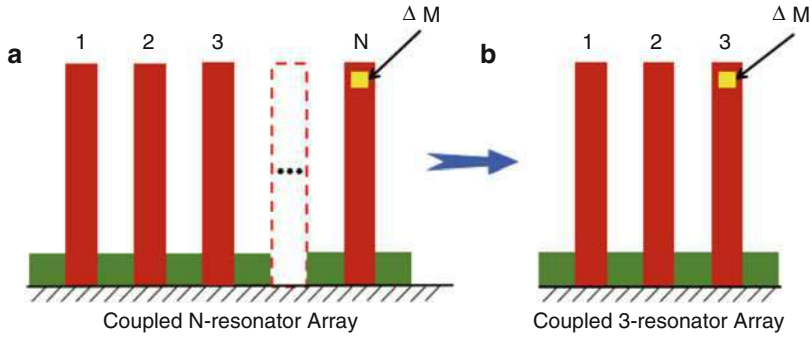
In contrast, in recent years, the concept of using Anderson or vibration mode localization (Hodges 1982; Pierre and Dowell 1987; Buks and Roukes 2002; Sato et al. 2003; Thiruvengatanathan et al. 2010, 2011) in any array of nearly identical coupled cantilevers has also been proposed as an eigenstate-shift-based sensing mechanism under ambient conditions. Some advantages of mode localized sensing can be listed below. Firstly, times or orders of magnitude in parametric sensitivity of micromechanical mass detection compared to the conventional frequency shift approach can be obtained. Secondly, such sensors can offer important advantages to intrinsic common mode rejection that renders it less susceptible to false-positive readings that frequency-shift-based sensors can cause. Thirdly, both ultrasensitive detection and analyte identification of small perturbations can be achieved at the same time with a single-coupled cantilever array.

In this chapter, the effect of different coupling designs (Chatani et al. 2011, 2012; Wang et al. 2012b) on the relative change (%) of amplitude was analytically studied using a coupled 15-cantilever array for each vibration mode. For simplifying complicated vibration characteristics, a coupled three-cantilever array was fabricated for preliminary examination with no mass perturbation (Wang et al. 2012b). When comparing experimental results with analytical ones, the main vibration characteristics, including the number of resonance modes, the high order of each cantilever at each resonance mode, as well as the vibration phase at each resonance mode, are found to be almost the same and correspond to each other. Both analytical and experimental results show that higher amplitudes could be obtained for the two side (outermost) cantilevers due to the symmetrical design. Further, by introducing a picogram mass perturbation (Wang et al. 2015a) into one side cantilever, shifts in the eigenstate can be orders of magnitude greater than shifts in the resonant frequency, which has been verified theoretically, analytically, as well as experimentally.

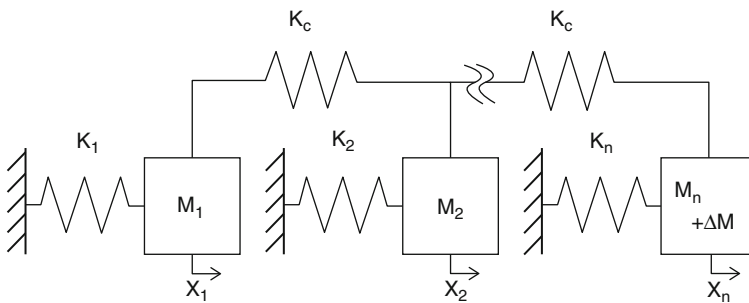
## Theory

### Mode-Localized Coupled N-Cantilever Array

A coupled N-cantilever array shown in Fig. 1a is usually simplified as a typical spring-mass model in Fig. 2. Each cantilever consists of a spring ( $K_i$ ), and a mass ( $M_i$ ) ( $i = 1, 2, 3$ ), and are coupled to its neighboring cantilever through coupling springs ( $K_c$ ). The matrix form of the equations of motion can be written in the following form (Spletzer et al. 2008):



**Fig. 1** Schematic of the coupled N-cantilever array (a) and three-cantilever array (b) with a mass perturbation placed at the end of one side beam-shaped cantilever



**Fig. 2** Simplified model of the coupled N-cantilever array, where  $K_i$  ( $=K$ ) and  $M_i$  ( $=M$ ) are, respectively, the bending stiffness and suspended mass of the identical cantilevers, while  $K_c$  is the stiffness of the overhang coupling the cantilevers.

$$\bar{M}^{-1} \bar{K} u = \omega^2 u \tag{1}$$

where  $u$  is a normalized eigenstate ( $|u_i| = 1$ ) of the system representing the tip amplitudes of each cantilever of the array at the corresponding eigenfrequency, ( $\omega$ ) is an eigenfrequency of the system, and  $M$  and  $K$  are the mass and stiffness matrices of the system, separately, given by the following equations, respectively:

$$\bar{M} = \begin{bmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_n + \Delta M \end{bmatrix} \tag{2}$$

$$\bar{K} = \begin{bmatrix} K_1 + K_c & -K_c & \dots & 0 \\ -K_c & K_2 + 2K_c & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & -K_c & K_n + K_c \end{bmatrix} \tag{3}$$



where  $M_i$  corresponds to beam-shaped cantilevers,  $M_i + \Delta M$  corresponds to the cantilever with mass perturbation, and  $K_i$  corresponds to the stiffness of each cantilever. Solving the mass matrix of the system when  $\Delta M = 0$  and  $M_i = M$ ,  $K_i = K$  yields  $n$  eigenstates of the initially perfectly balanced system, while the primary mode consists of all cantilevers vibrating in phase with equal amplitude.

### Mode-Localized Coupled Three-Cantilever Array

To further investigate the physics model of the coupled cantilever array, the  $N$ -cantilever array is simplified to a three-cantilever array which is displayed in Fig. 1b.

When a mass perturbation is introduced into the coupled system, the eigenvalue  $\lambda$  governing the undamped free oscillations of the system can be non-dimensionalized and written as follows (Spletzer et al. 2006; Wang et al. 2015a):

$$\begin{bmatrix} 1 + K_C/K_1 & -K_C/K_1 & 0 \\ -K_C/K_2 & 1 + 2K_C/K_2 & -K_C/K_2 \\ 0 & (-K_C/K_3)/(1 + \delta) & (1 + K_C/K_3)/(1 + \delta) \end{bmatrix} u = \lambda u \quad (4)$$

In the case of  $\delta = 0$ , both the eigenvalues and eigenstates of the coupled system can be expanded as follows (Wang et al. 2015a):

$$\begin{aligned} \lambda_1^0 = 1, u_1^0 &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \lambda_2^0 = 1 + K_C/K_3, u_2^0 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}; \lambda_3^0 = 1 + 3K_C/K_1, u_3^0 \\ &= \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \end{aligned} \quad (5)$$

When a small mass is added to one of the outmost cantilevers of the coupled three-cantilever system ( $\delta \neq 0$ ), the perturbed eigenstates  $u_i$  become localized, and the relative change in eigenstate  $u_i$  in this sensing paradigm can be written as follows, where  $u_{i,0}$  and  $u_i$  represent the  $i$ th eigenmode vector before and after mass is added:

$$|\Delta u_i| = |u_i - u_{i,0}| \quad (6)$$

It is extremely complex to expand the resulting three eigenvalues and three eigenstates without using an appropriate mathematical physics method. Therefore, appropriate computer software is necessary for simulating the vibration of coupled cantilevers under ambient mass perturbation.

## Simulation Results and Experimental Verification

### Effect of the Coupling Overhang

At first, a comparative study, without and with coupling overhang, has been thus designed and analyzed. It is verified from analytical result that vibration mode

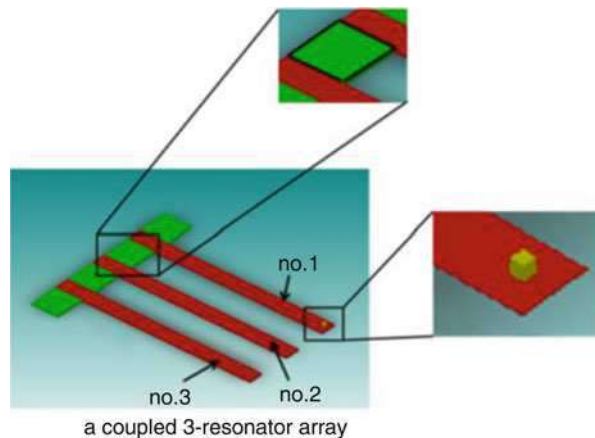
localization happens because of coupling overhang in the cantilever array. It is expected that the number of resonance point is the same as the number of coupled cantilever. Figure 3 shows the structure of three mechanically coupled cantilevers.

Taking off couplings from the model in Fig. 3, cantilevers turn out to be separated from the neighboring one by the size of a coupling overhang (Wang et al. 2012b). In the analytical level, however, all cantilevers are supposed to be far away from each other. So it is necessary to analytically combine these cantilevers via a common base to ensure of a separating distance corresponding to the coupling overhang. As expected before the fact, in the case of without coupling overhang, only one resonance point can be obtained for all cantilevers as shown in Fig. 4. This indicates that all cantilevers vibrate individually so that vibration mode localization cannot be observed. In the case of with coupling overhang, as shown in Fig. 5, three resonance points (211.29, 214.42, and 217.87 kHz) can be obtained for all cantilevers. At all resonance points, the amplitudes of cantilever No. 1 and No. 3 are very different from that of cantilever No. 2. It means that vibration mode localization happens because of a cantilever interacting under vibration via coupling overhang.

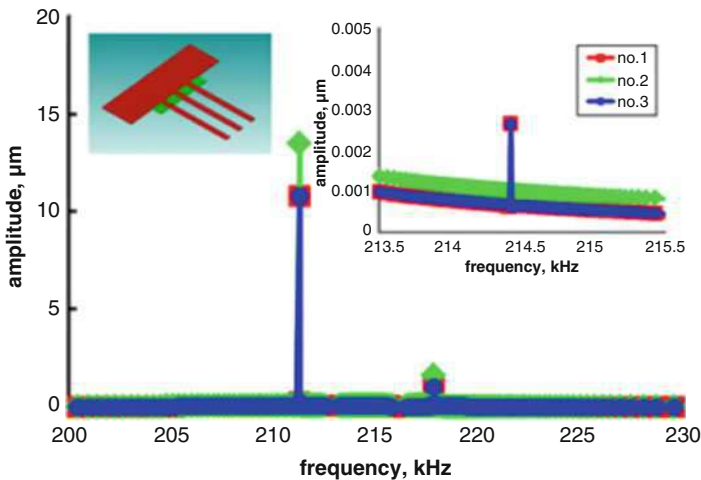
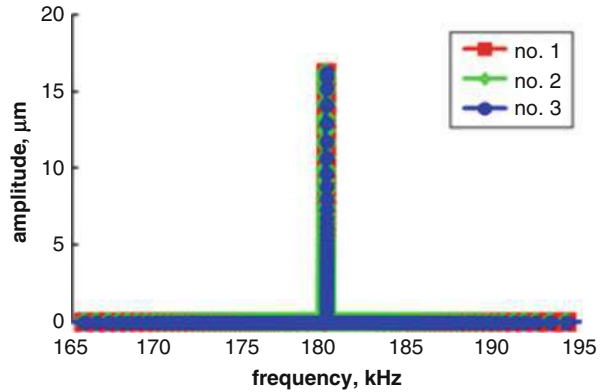
Also, the effect of changing geometrical designs has been researched (Wang et al. 2012c). As shown in Table 1, the effects of five kinds of geometrical designs of the coupling overhang on the amplitude enhancement are summarized. Design No. 4 (4a:4b) undergoes 3.81 times (compared to a single cantilever) of magnitude greater than corresponding shifts in resonant frequency of cantilever No. 14 at mode 15. While for cantilever No. 13 at mode 1, still 2.5 times of magnitude can be obtained. By comparing the analytical results in Table 1, design No. 4 is believed to be desirable and is thus selected in the following study.

Table 2 further summarizes the effect of design No. 4 of the coupling overhang on the amplitude shift enhancement in a perfect array of five or ten identical spring-mass beams. Due to the complexity of 15-cantilever array structure, a simpler five-cantilever array is selected to further investigate its characterization of mode localization. In the case of the five-cantilever array, design No. 4 also undergoes 3.81

**Fig. 3** Schematic of the coupled beam-shaped three-cantilever array with a mass perturbation applied to cantilever No. 1



**Fig. 4** Analytical results of micromechanically coupled beam-shaped three-cantilever array simulated by using CoventorWare™ software, in the case of without coupling overhang








**Fig. 5** Analytical results of micromechanically coupled beam-shaped three-cantilever array simulated by using CoventorWare™ software, in the case of with coupling overhang, where the inserted image shows the amplified image at the resonance point of 214.42 kHz

times of magnitude greater than corresponding shifts in resonant frequency of cantilever No. 4 at mode 5. The analytical results show that higher amplitudes could be obtained for the two outermost cantilevers (No. 1 and 5) due to the symmetrical design.

**Simulation Results of a 15-Cantilever Array**

Considering the advantage of mechanically coupled structure, a large amount of simulation work has been conducted. In the work (Chatani et al. 2011), a 15-cantilever array is designed and simulated using CoventorWare™ software. As shown in Fig. 5, the mass perturbation is induced on the tip of the cantilever No. 14, and the vibration localization at cantilever No. 13, No. 14, and No. 15 were thus

**Table 1** The effect of different geometrical design of coupling overhang on the relative change (%) of amplitude before and after a small mass perturbation

	Design of overhang a:b	Amplification of resonator 13 (mode 1)	Amplification of resonator 14 (mode 15)	Amplification of resonator 15 (mode 1)
Design ①	1:1 	2.54 (44.92/17.70%)	3.67 (65.21/17.70%)	2.44 (43.18/17.70%)
Design ②	4:1 	2.52 (44.69/17.70%)	3.73 (66.08/17.70%)	2.54 (45.02/17.70%)
Design ③	1:4 	2.09 (36.93/17.70%)	3.51 (62.12/17.70%)	1.98 (35.01/17.70%)
Design ④	4:4 	2.58 (45.74/17.70%)	3.81 (67.43/17.70%)	2.34 (41.49/17.70%)
Design ⑤	8:4 	2.38 (42.16/17.70%)	3.81 (67.37/17.70%)	2.12 (37.58/17.70%)

**Table 2** The effect of design No. 4 (a:b) of coupling overhang on the relative change (%) of amplitude before and after a small mass perturbation

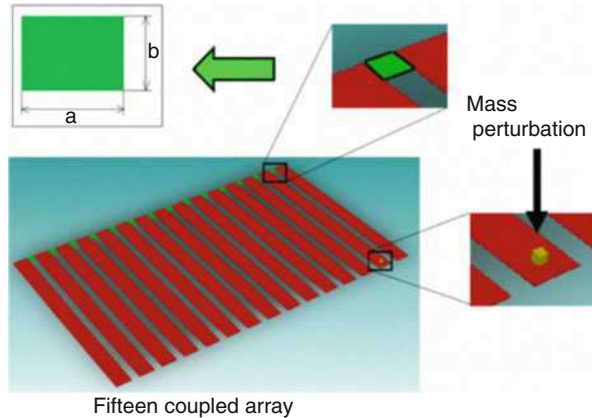
	Design of overhang a:b	Amplification of resonator 3, 8, 13 (mode 1)	Amplification of resonator 4, 9, 14 (mode: resonator 5:4, 10:9, 15:14)	Amplification of resonator 5, 10, 15 (mode 1)
Design ④ 5-beam	4:4	2.42 (42.88/17.70%)	3.81 (67.4/17.70%)	2.37 (42.88/17.70%)
Design ④ 10-beam	4:4	2.36 (41.76/17.70%)	3.85 (68.16/17.70%)	2.49 (44.15/17.70%)
Design ④ 15-beam	4:4	2.58 (45.74/17.70%)	3.81 (67.43/17.70%)	2.34 (41.49/17.70%)

analyzed. As shown in Fig. 6, the eigenstate shifts outnumber eigenfrequency shifts when the system is induced in vibration mode 1 and 15, respectively, before and after a mass perturbation of 10 pg. It is noticed that mode 15 undergoes the greatest change of 67.43%, and both the mode 1 and mode 2 also show relatively great changes.

### Experimental Results of a Five-Cantilever Array

The symmetrical vibration characteristics, resulted from symmetrical design observed in analytical results as shown in Fig. 7, are not very clear in Fig. 8, especially comparing cantilever No. 1 with 5 or No. 2 with 4. It can be seen that the amplitude response (18.457 dBm from a more detailed measurement) of cantilever No. 5 is greatly magnified by vibration mode localization when driven by its resonant frequency of 209.05 kHz. Additionally, all cantilevers show the same resonant frequency of 209.05 kHz and are thus believed to be fabricated geometrically perfect, while cantilever No. 1 and 5 are connected by one coupling overhang rather than two like the others. This is the only different point among the above cantilevers and might account for why cantilever No. 5 displays a great amplitude by vibration mode localization. However, this cannot explain the very low amplitude response (−5.584 dBm from a more detailed measurement) of the symmetrical cantilever (No. 1). Despite that microfabrication errors might cause such

**Fig. 6** The schematic of coupled beam-shaped 15-cantilever array with a mass perturbation of 10 pg applied to cantilever No. 14, and the size definition of coupling overhang



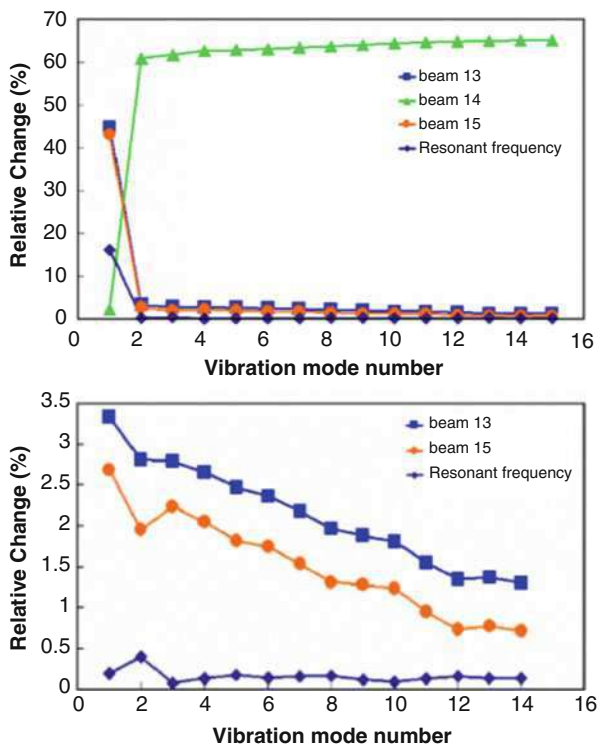
a phenomenon, further study is conducted by using relatively simple system like three-cantilever array to empirically verify the correlation between the analytical and the experimental results (Fig. 9).

### Simulated and Experimental Results of a Three-Cantilever Array

As for the specific research method, a comparative study has been done by using three identical spring-mass beams on both analytical characterizations with small mass perturbations and sensing verification by manipulating a picogram polystyrene microsphere. In the case of analytical characterizations, by applying a picogram-order mass perturbation (10 pg) shown in Fig. 10b on one side cantilever, significant enhancements in amplitude changes were obtained at the second resonance mode from both of the unloaded cantilevers. The amplitude change from the center cantilever is about 7000 times higher than that with no mass perturbation, while the change in amplitude from another side cantilever is about 4000 times higher. In the aspect of experimental characterizations, the enhancement in amplitude change at the second resonance mode is verified by applying two polystyrene microspheres (about 8.8 pg) shown in Fig. 10a as a picogram-order mass perturbation onto one side cantilever. Due to the operational difficulties in quantitatively manipulating polystyrene microspheres, the effects of mass variations on the enhancement in amplitude changes from unloaded cantilevers were further analytically characterized under a range of 0.01~100 pg for three resonance modes, respectively.

Then, in Fig. 11, the vibration characteristics have been analyzed on both frequency and amplitude (Wang et al. 2015a). From the view point of frequency shifts, the first (in blue) and third (in red) resonance modes of 211.29 kHz and 217.87 kHz were decreased to 205.47 kHz and 215.33 kHz, respectively, due to mass perturbation on center cantilever, while the second (in orange) resonance mode of 214.42 kHz is kept unchanged. In the case of mass on side cantilever, however, the first and second resonance modes of 211.29 kHz and 214.42 kHz were decreased to 205.41 kHz and 212.63 kHz, respectively, while the third resonance mode of 217.87 kHz has been shifted slightly to 217.21 kHz.

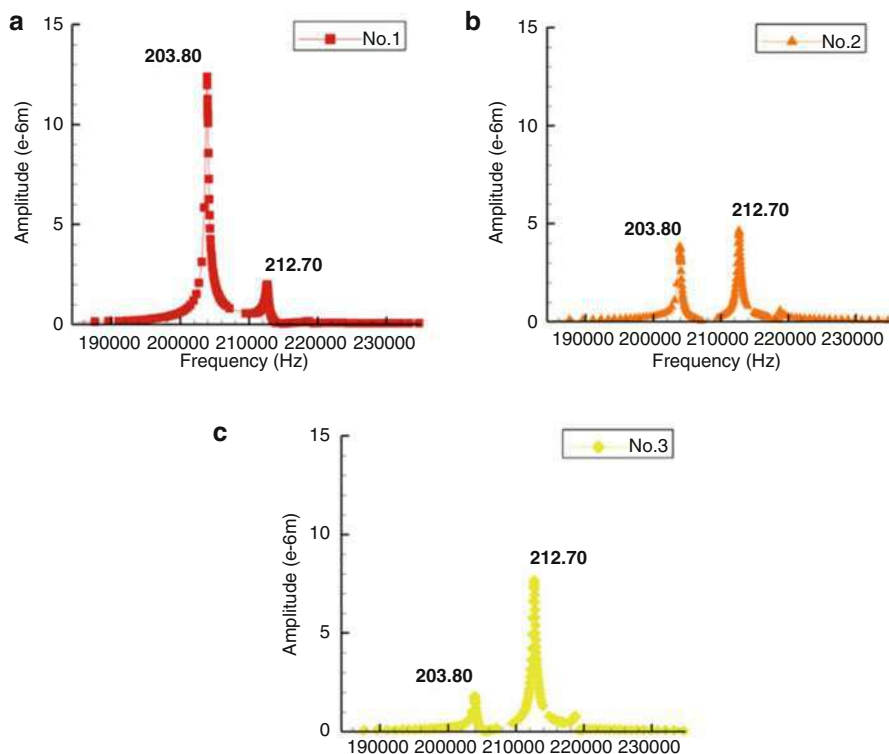
**Fig. 7** The relative change (%) of eigenstate (amplitude) due to the mass perturbation as a function of vibration modes, with a relation to a geometrical design of coupling overhang as 4a:4b. The resonant frequency of a single cantilever as a function of vibration mode is also drawn for comparison



On the other hand, considering the amplitude changes (defined as the ratio of the amplitude change due to mass perturbation to the amplitude value with no mass perturbation), which are summarized in Table 3a, b, an amplitude change of 6.748 is obtained at the third resonance mode of 215.33 kHz from side cantilever for mass perturbation on the center cantilever, while large amplitude changes of 4064 and 7873 were obtained at the second resonance mode of 212.63 kHz from side and center cantilever, respectively, for mass perturbation on another side cantilever. As shown in Table 3a, although a higher amplitude change of 8.550 appeared at the second resonance mode, both of the analytical values with and without mass perturbation are too small to be detected and are thus not desirable (Fig. 12).

Compared with simulated results shown in Fig. 11, three measured amplitude-frequency spectra from cantilevers No. 1, No. 2, as well as No. 3 with mass perturbation have been, respectively, depicted in Fig. 13. Three eigenstate resonance modes of 217.32 kHz (the first resonance mode in blue), 225.88 kHz (the second resonance mode in orange), and 231.82 kHz (the third resonance mode in red) can be correspondingly observed for all three cantilevers, although the measured resonance frequencies are shifted to a lower level compared to those of 223.96, 228.36, and 232.87 kHz with no mass perturbation (Wang et al. 2012b).

The observation results of amplitude changes have been summarized in Table 4. It can be seen that at second resonance mode, the measured amplitude values of side and



**Fig. 8** Amplitude with respect to frequency spectra, where (a) cantilever No. 1 or 5; (b) cantilever No. 2 or 4; and (c) cantilever No. 3

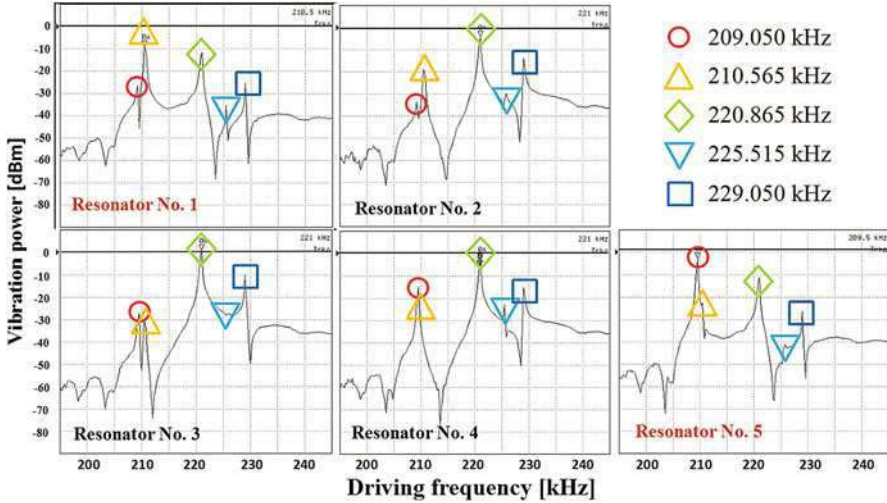
center cantilevers greatly changed from  $-16.414$  dBm and  $-33.248$  dBm to  $-3.846$  dBm and  $-5.786$  dBm, respectively. In other words, by applying two microspheres on one of the side cantilevers, a great enhancement in amplitude change at the second resonance mode can be observed and verified from the center cantilever.

In addition to the first comparative study on both analytical and experimental characterizations with a fixed mass perturbation by applying picogram polystyrene microspheres, the effect of mass changes has been also analytically conducted under a range of  $0.01\sim 100$  pg, which is depicted in Fig. 14.

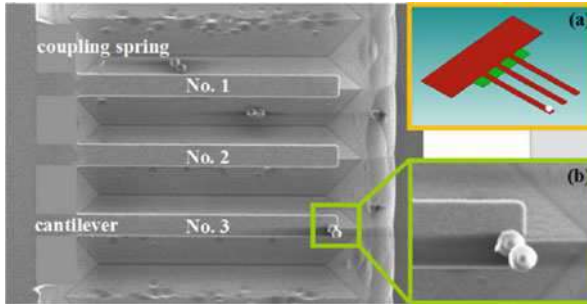
To sum, a comparison of experimental results, simulated results, and analytical results in both cases of no mass perturbation and with mass perturbation is summarized in Table 5 from the viewpoints of eigenstates (resonance modes) and eigenvalues (amplitudes).

### Simulation Conditions

All simulations are conducted under an exciting force of  $0.01 \text{ kg} \times \text{mm/s}^2$  and a damping of  $0.001 \text{ kg/s}$ .



**Fig. 9** The measured amplitude-frequency spectra for all five cantilevers. The amplitude response (18.457 dBm from a more detailed measurement) of cantilever No. 5 is greatly magnified by vibration mode localization when driven by its resonant frequency of 209.05 kHz



**Fig. 10** Micrograph of a coupled three-cantilever array with a fixed mass perturbation applied on the side cantilever (No. 3), where (a) showing the schematic of a perfect array of three identical spring-mass beams; (b) showing two polystyrene micro-spheres as the fixed mass perturbation for experimental characterizations

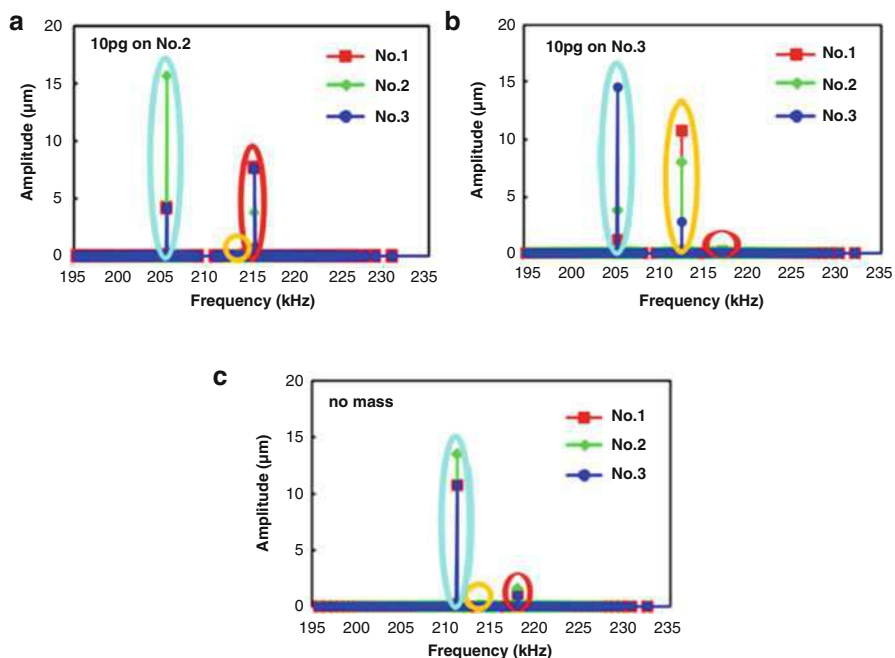
### A Coupled Five-Cantilever Array

The cantilevers are nominally 66.7~133.4  $\mu\text{m}$  long, 5  $\mu\text{m}$  wide, and 0.1~0.5  $\mu\text{m}$  thick, are separated by a distance of 2.5~20  $\mu\text{m}$ , and are made of SOI (silicon on insulator) wafer (Wang et al. 2012c).

### A Coupled Three-Cantilever Array

The geometrical size of the overhang is detailed in Fig. 9a, where the gap between beams is sized 10  $\mu\text{m}$  and the contact length of overhang adjacent to the beam is



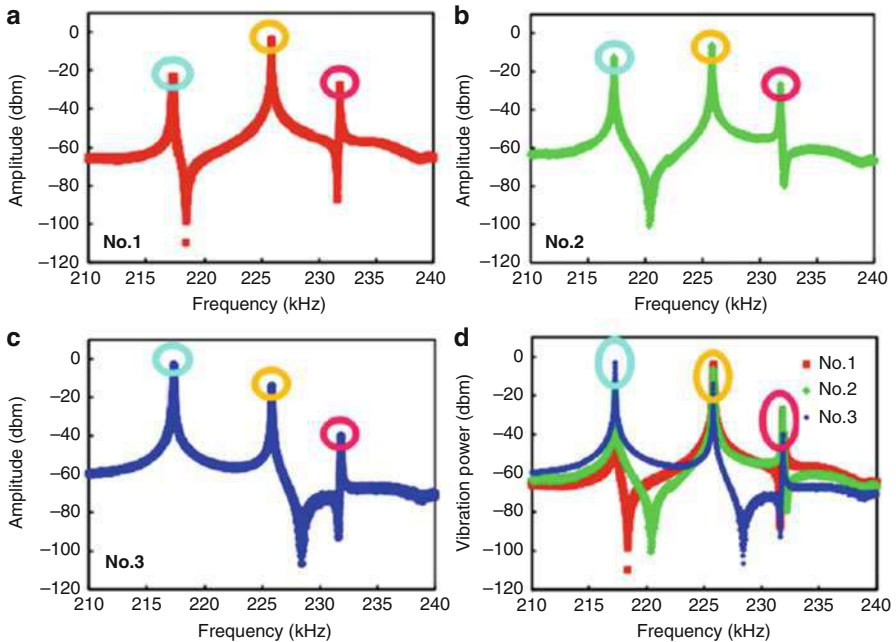
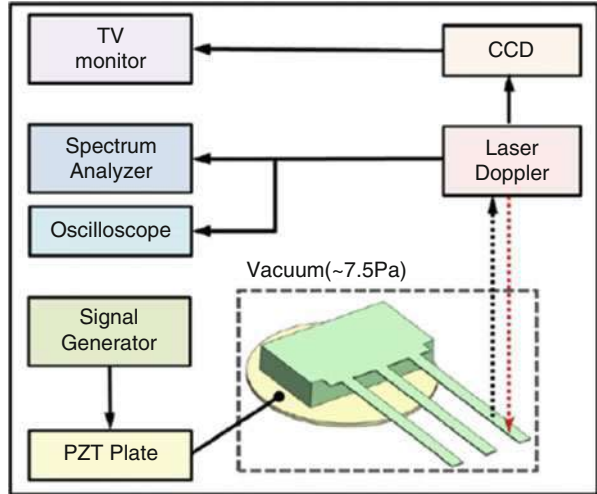


**Fig. 11** Analytical characterizations of a coupled three-cantilever array analyzed by using CoventorWare™ software, in the case of applying a mass perturbation, where (a) with mass on center cantilever (No. 2), (b) with mass on side cantilever (No. 3), and (c) with no mass on any cantilever. Different color circles correspond to the 1st resonance mode (in blue), the second resonance mode (in orange), as well as the third resonance mode (in red), respectively

**Table 3** (a) Comparison of relative changes of magnitudes with and without mass perturbation (applying 10 pg mass on center cantilever of No. 2). (b) Comparison of relative changes of magnitudes with and without mass perturbation (applying 10 pg mass on side cantilever of No. 3)

(a)				
Resonance mode	10 pg on No. 2 (kHz)	No mass (kHz)	$\Delta A/A_0$	
			No. 1 and No. 3	
1st	205.47	211.29	0.616	
2nd	214.42	214.42	8.55	
3rd	215.33	217.87	6.748	
(b)				
Resonance mode	10 pg on No. 3 (kHz)	No mass (kHz)	$\Delta A/A_0$	
			No. 1	No. 2
1st	205.41	211.29	0.891	0.72
2nd	212.63	214.42	4064	7873
3rd	217.21	217.87	0.818	0.983

**Fig. 12** Experimental setup for measuring frequency responses and demonstrating amplitude change enhancement by mode localization with a small mass perturbation



**Fig. 13** Three measured amplitude-frequency spectra in (a–c) correspond to three cantilevers in an array, respectively, where (d) summarizes all three spectra in one figure for comparison

**Table 4** Comparison of experimental results showing magnitude values (vibration power) for all three cantilevers under three different resonance modes with and without mass perturbation (applying 8.8 pg mass on side cantilever of No. 3)

Vibration power (dbm)		1st		2nd		3rd	
Resonance mode		0 pg	8.8 pg	0 pg	8.8 pg	0 pg	8.8 pg
Vibration power (dbm)	No. 1	2.780	-22.244	-16.414	-3.846	-10.367	-28.027
	No. 2	5.908	-12.773	-33.248	-5.786	-9.3154	-27.144
	No. 3	4.349	-3.060	-16.890	-14.208	-12.702	-39.959

sized 9.2  $\mu\text{m}$ . All beams are nominally 66.7  $\mu\text{m}$  long, 5  $\mu\text{m}$  wide, and 0.5~0.6  $\mu\text{m}$  thick, are separated by a distance of 20  $\mu\text{m}$ , and are made of a silicon on insulator (SOI) wafer (Wang et al. 2015a).

## Conclusions and Further Reading

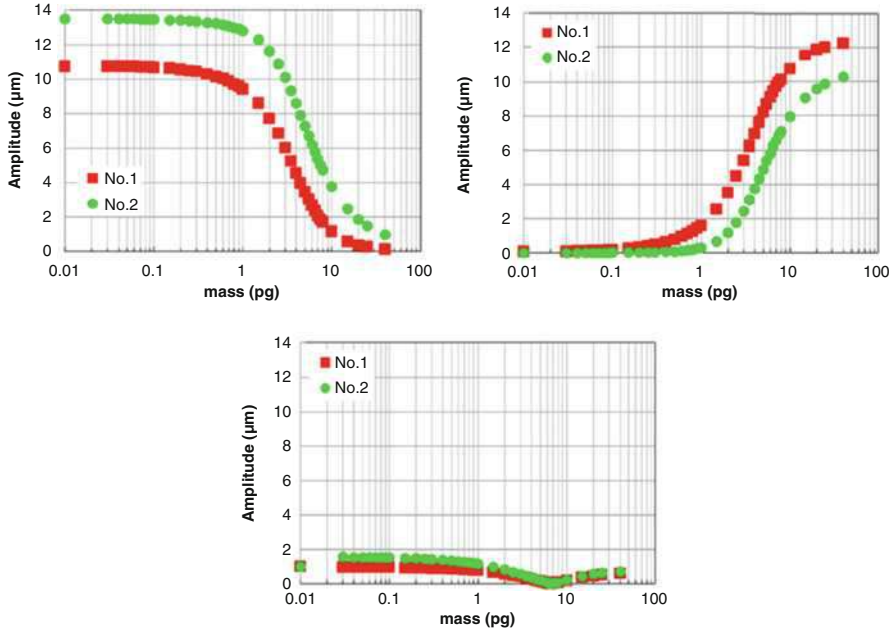
As shown in the research mentioned above, mode localization has a great effect on amplitude enhancement. Cantilever-based resonator array, a main structure for applying mode localization, can greatly contribute to picogram mass sensing. Some typical structures, such as 15-cantilever array, 5-cantilever array, and 3-cantilever array, have been, respectively, theoretically analyzed or experimentally verified. It can be concluded that a weakly coupled N-cantilever array is capable of detecting picogram mass perturbation using mode localization with ultrahigh sensitivity (compared to eigenfrequency shift sensing mechanism).

However, mode localization is limited in linearity vibration, which cannot represent the majority of vibration phenomenon in nature. Therefore, synchronization, which is regarded as an important rule on nonlinearity vibration, has been paid more and more attention to. Accompanying with the development of mode localization, the research on using synchronization as a mechanism on mass sensing has been conducted by large amounts of scholars around the world.

## Approaching with Synchronized Oscillation

### Overview

In recent years, cantilever-based resonators with high sensitivities have been progressively optimized for mass sensing (Ono et al. 2003a, b; Wang et al. 2003, 2004; Ekinici et al. 2004; Lavrik et al. 2004; Gil-Santos et al. 2009), force sensing (Stowe et al. 1997; Mamin and Rugar 2001), and parametric sensing (Napoli et al. 2005; Zhang and Turner 2005; Requa and Turner 2006) over the years. The sensitivity and resolution of an independent (self-supporting) cantilever depends on various factors,



**Fig. 14** Analytical results on amplitude changes as a function of mass variations on side cantilever of No. 3 from 0.01 to 100 pg, where (a) for the first resonance mode, (b) for the second resonance mode, and (c) for the third resonance mode, respectively

i.e., total effective spring constant, resonant frequency, quality factor (energy dissipation), noise level at a given temperature (Albrecht et al. 1991; Cleland and Roukes 2002), and vibration level (Filler 1988).

Synchronization of coupled systems is a common phenomenon in both nature and biological physiology. Its discovery dates back to the time of Christian Huygens, who remarked that two slightly out-of-step pendulum-like clocks become synchronized after they are connected using a thin wooden board (Pikovskiy et al. 2001). Such kind of phenomenon can be further applied to accomplishing high efficiency energy harvester using synchronized switching techniques (Tang and Yang 2011; Becker et al. 2013). Recently, synchronization of laser-induced self-sustained vibrations of radio-frequency micromechanical resonators was demonstrated by applying a small pilot signal either as an inertial drive at the natural frequency of the resonator or by modulating the stiffness of the resonator at double the natural frequency (Zalalutdinov et al. 2003). In the investigation of a two beam resonator system coupled by an overhang, a wide range of frequency entrainment was observed with hysteretic behavior attributed to the strong nonlinearity (Shim et al. 2007). In the study (Wang et al. 2010), phase noise suppression has been demonstrated via two synchronized cantilever system. Theoretical analysis and experimental research for frequency locking and synchronized region were researched by beam-shaped and U-shaped cantilever array (Wang et al. 2010, 2012a, 2012d; Nakajima et al. 2011, 2012;

**Table 5** Comparison of experimental results, simulated results, and analytical results from viewpoints of eigenstates (resonance modes) and eigenvalues (amplitudes) in both cases of no mass perturbation and with mass perturbation

Comparison view points		Experimental results	Simulated results	Theoretical results
Eigenstates (modes) (in kHz)	No mass (Hosaka et al. 2006)	1st: 223.96	1st: 211.29	$u_1^0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ $u_2^0 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ $u_3^0 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$
		2nd: 228.36	2nd: 214.42	
		3rd: 232.87	3rd: 217.87	
	With mass on No. 3	1st: 217.32	1st: 205.41	$u_1 = \begin{bmatrix} +0.5733 \\ +0.5733 \\ +0.5734 \end{bmatrix}$ $u_2 = \begin{bmatrix} -0.7104 \\ -0.0067 \\ +0.7037 \end{bmatrix}$ $u_3 = \begin{bmatrix} +0.4114 \\ -0.8190 \\ 0.4000 \end{bmatrix}$
		2nd: 225.88	2nd: 212.63	
		3rd: 231.82	3rd: 217.21	
Eigenvalues (amplitudes) in 2nd mode	No mass (Hosaka et al. 2006)	-33.248 dbm zero amplitude	A <sub>0</sub> zero amplitude	Expanded eigenvalues not available
	With mass on No. 3	-5.786 dbm	$\Delta A/A_0 = 7873$	Expanded eigenvalues not available

Itoh et al. 2012b), and picogram mass sensitivity of synchronized structure has been demonstrated in the study (Itoh et al. 2012a; Wang et al. 2015b, 2016; Du et al. 2016).

## Frequency Locking in Coupled Cantilevers

The identity of the frequencies that holds within a finite range of the detuning is the hallmark of synchronization and is often called frequency locking.

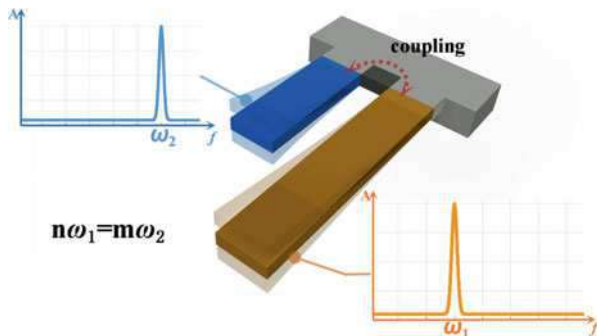
If the resonant frequencies of two geometrically designed cantilevers obey the relation of  $\omega_2 \approx n\omega_1$  ( $n$  is integer), they will synchronize with each other and lock at  $\omega_2 = n\omega_1$  when they are coupled through a mechanical element, even if the frequency  $\omega_1$  slightly shifts to  $\omega'_1 = \omega_1 + \Delta\omega$  due to a small mass perturbation or an external force (mechanical or thermal noise) (Fig. 15).

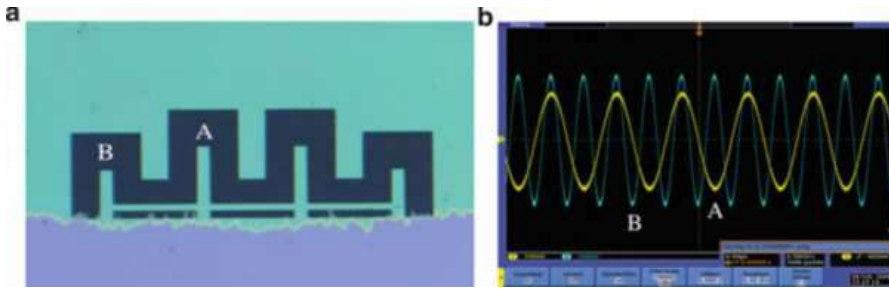
This means that the frequency response signal can be enhanced by a factor of  $n$  if a mechanically coupled system is constructed for two independent resonators. This can be applied to further improve the detectable limit (resolution) of cantilever-based resonators for sensing application.

The classical structure consists of two geometrically designed singly-clamped beam-shaped cantilevers was designed using software. In 2010 (Wang et al. 2010), two cantilevers, with calculated resonant frequencies of 306.690 kHz (detecting) and 153.298 kHz (sensing), respectively, obey the relation of  $\omega_2 \approx 2\omega_1$  and are expected to be coupled through a thin support as a mechanical element.

The vibration of the cantilevers is measured using a double beam laser Doppler system, and the vibration amplitude is detected using a lock-in amplifier. The frequency responses of the fabricated two-cantilever resonant system (19.5- $\mu\text{m}$ -long cantilever B, 30- $\mu\text{m}$ -long cantilever A), micromechanically coupled by a 30- $\mu\text{m}$ -long thin beam, as shown in Fig. 16a, were measured using frequency counters. The measured resonant frequencies of cantilevers B and A are 295.72 kHz and 145.01 kHz, respectively, in Fig. 16b. The empirically measured frequency ratio of 1:2.04 is almost the same as the designed frequency ratio of 1:2.

**Fig. 15** The schematic figure for mutual synchronization





**Fig. 16** (a) Micrograph of the fabricated resonator system; (b) Waveforms of coupled resonator system

### Synchronization Region

#### Methodology for Characterizing the Entrainment Range

Synchronized region is the frequency range which coupled oscillators can realize frequency locking. As drawn in Fig. 17, the vibration power was measured as a function of the above driving frequency with respect to two coupled cantilevers, respectively (Wang et al. 2016). The value of 0 dBm on the vibration power corresponds to the velocity peak of  $1000 \text{ mm s}^{-1}$  at the tip of the cantilever. The vibration power (amplitude response) reaches the maximum value at point C when the driving frequency closes to the resonant frequency of 182.85 kHz.

By summarizing the frequency ratio (frequency of rectangular cantilever to that of U-shaped cantilever) as a function of the driving frequency in Fig. 17, phase-locking is observed, and a plateau with a frequency ratio of 2.000 corresponding to an entrainment range has been characterized from bifurcation point B to bifurcation point D, when the driving signal is close to a half of the resonant frequency of  $\omega_2 = 372.503 \text{ kHz}$ , the measured resonant frequency of rectangular cantilever.

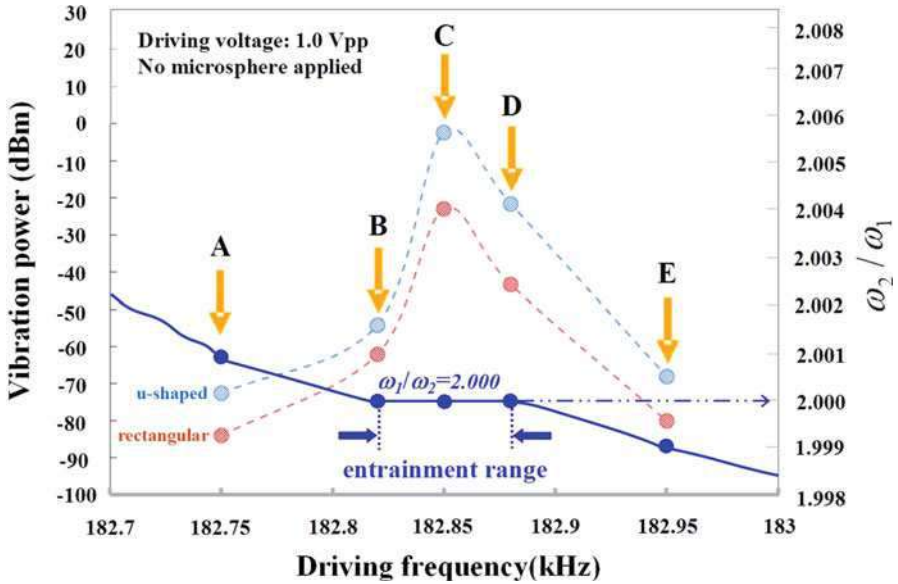
The characterized frequency entrainment range in Fig. 17 essentially corresponds to the obtainable regime of doubled frequencies. Further considering the factors influencing the entrainment range by using a lumped element model of two cantilevers coupled by two symmetrical overhangs as schematically shown in Fig. 18 (Wang et al. 2016), each cantilever is modeled as a damped simple harmonic oscillator, while the effect of the overhang coupling is modeled as a spring connecting the two cantilevers (oscillators). The motion equation, containing a nonlinear spring constant and a term corresponding to the effect of the coupling spring, can be generally written as follows:

$$m_1 \ddot{x}_1 + c_1 \dot{x}_1 + k_{1, \text{linear}} x_1 + k_{1, \text{nonlinear}} x_1^3 + 2k_{\text{coupling}} (x_1 - x_2) = F_1 \cos \omega t \quad (7)$$

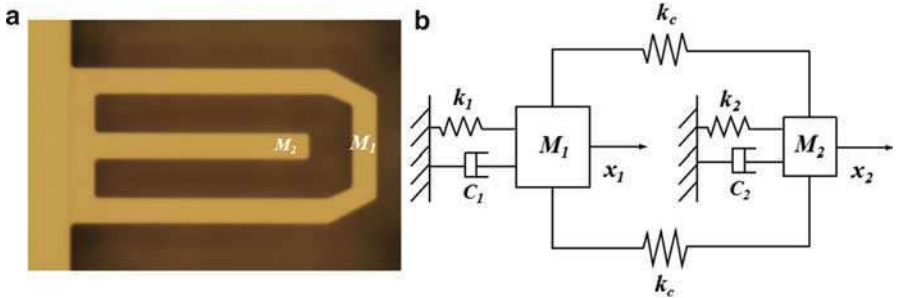
Its general solution as follows, respectively:

$$x_1 = A_1 \cos \{ \omega t + \varphi_1 \} \quad (8)$$

$$x_2 = A_2 \cos \{ \omega t + \varphi_2 \} \quad (9)$$



**Fig. 17** Vibration power and frequency ratio are drawn as a function of the driving frequency with respect to two coupled cantilevers, where point *A* is 182.75 kHz, point *B* is 182.82 kHz, point *C* is 182.85 kHz, point *D* is 182.88 kHz, and point *E* is 182.95 kHz. The flat interval (from bifurcation point *B* to bifurcation point *D*) with a frequency ratio of 2.000 corresponds to an entrainment range, where the driving frequencies were doubled via phase-locking



**Fig. 18** Optical image of the fabricated micro-cantilevers coupled by two symmetrical overhangs (**a**) corresponds to a lumped element model (**b**), where each cantilever is modeled as a damped simple harmonic oscillator while the effect of the overhang coupling is modeled as a spring connecting the two cantilevers (oscillators)

where  $m_n$  ( $n = 1, 2$ , referring to the respective cantilevers),  $x_n$ ,  $c_n$ ,  $k_{n,linear}$ ,  $k_{n,nonlinear}$ ,  $k_{coupling}$ ,  $A_n$ ,  $\varphi_n$ ,  $\omega$ , and  $F_n$  are, respectively, the mass, displacement, damping coefficient, spring constant (linear), spring constant (nonlinear), spring constant (coupling), amplitude, phase, driving frequency, and force acting on the cantilever.



Both the amplitude and the phase are time-varying parameters. If introducing a minimum term of  $\varepsilon$ , defining the expressions in as follows:

$$\frac{c_1}{m_1} = 2\varepsilon\zeta_1, \frac{k_{1,l}}{m_1} = \omega_{0,1}^2, \frac{k_{1,n}}{m_1} = \varepsilon\alpha_1, \frac{k_c}{m_1} = \varepsilon\kappa_1, \frac{F_1}{m_1} = F_1' \tag{10}$$

Substituting above equation to the first three equations, equations for  $dA_2$  and  $dA_1$ , respectively, can be approximately derived as follows:

$$\frac{dA_1}{dt} = -\varepsilon\zeta_1 A_1 + \frac{\varepsilon\kappa_1 A_2}{\omega_d} \sin(\varphi_1 - \varphi_2) - \frac{\varepsilon}{2\omega_d} F_1 \sin \varphi_1 \tag{11}$$

$$\frac{d\varphi_1}{dt} = \frac{\varepsilon\delta_1}{2\omega_d} + \frac{3}{8\omega_d} \varepsilon\alpha_1 A_1^2 + \frac{\varepsilon\kappa_1}{\omega_d} \left\{ 1 - \frac{A_2}{A_1} \cos(\varphi_1 - \varphi_2) \right\} - \frac{\varepsilon}{2\omega_d A_1} F_1 \cos \varphi_1 \tag{12}$$

Based on the above equation, if  $d\varphi_n/dt = 0$  and the difference (entrainment range) between the driving frequency ( $\omega_d$ ) and the eigenstate frequency ( $\omega_{0,1}$ ) is defined as  $\Delta\omega$ , the occurrence condition for phase-locking can be determined as follows:

$$|\Delta\omega| = |\omega_{0,1} - \omega_d| \leq \frac{1}{m_1(\omega_{0,1} + \omega_d)A_1} \left\{ F_1 - \frac{3}{4}k_{1,n}A_1^3 - 2k_c(A_1 - A_2) \right\} \tag{13}$$

It can be noted from the above equation, increasing the cantilever mass leads to decreasing the entrainment range, while increasing the force  $F_l$  acting on the cantilever and adjusting the coupling overhang lead to expanding the entrainment range, as bordered by the bifurcation points of B and D in Fig. 17.

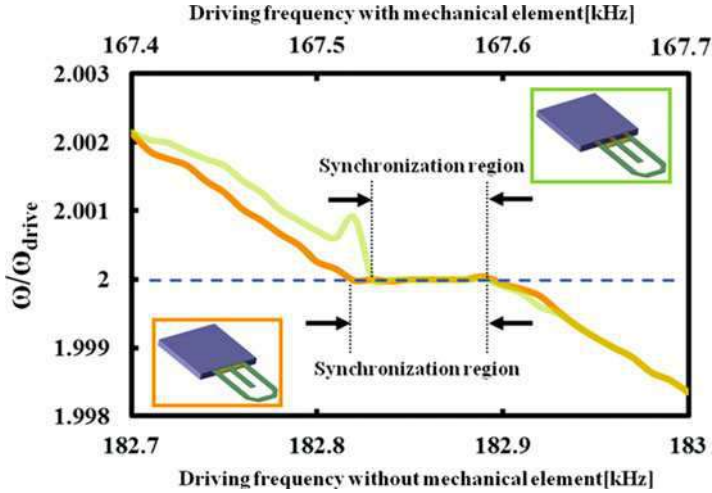
### Effect of Coupling Overhang

Coupling overhang can realize the transfer of force between coupled oscillators, which is an important part in synchronized system.

To investigate the application potential of the micromechanically coupled resonator system as high-resolution MEMS/NEMS mass or force sensors, the structural (geometrical) designs for accomplishing synchronization in two ‘‘U-shaped cantilever-based oscillator’’ systems, with and without micromechanical elements(coupling overhang), were carried out, and synchronized oscillations were then demonstrated and characterized (Wang et al. 2012a).

Structural design without mechanical element: The ‘‘U-shaped cantilever-based oscillator’’ system consists of a U-shaped cantilever (500 nm thick) functioning as a sensing element and a beam-shaped cantilever (45 mm long  $\times$  5 mm wide  $\times$  00 nm thick) functioning as a detecting element. The two cantilevers are coupled through two coupling overhangs as defined in Fig. 19. The two cantilevers are coupled through two coupling overhangs and connected to the support by three mechanical elements. The effect of three mechanical elements on synchronization region could thus be studied. Additionally, the position of the coupling overhang also shifts 5 mm.

Figure 19 shows the frequency ratio of the self-oscillating cantilever (beam shaped) to the signal-driven cantilever (U shaped) as a function of the change of



**Fig. 19** Frequency ratio of the self-oscillating cantilever (beam shaped) to the signal-driven cantilever (U shaped) applied a driving frequency (1.0 V<sub>pp</sub>). Flat interval with a frequency ratio of 2.000 corresponds to a stable synchronized oscillation region

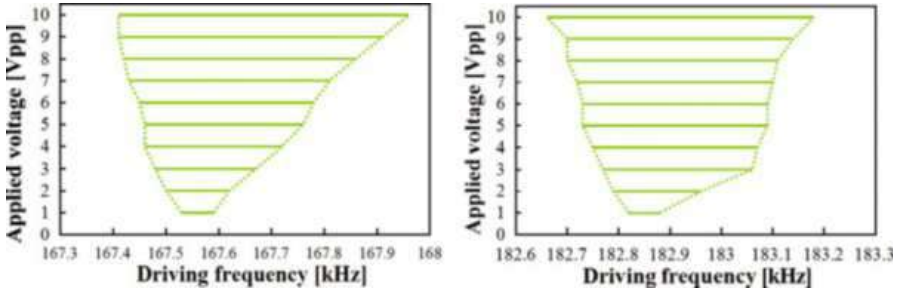
driving frequency. From the plot, a plateau with a frequency ratio of 2.000 corresponding to a synchronization region is observed for both systems.

For the system with three mechanical elements, the lengths of both cantilevers were increased by 5 mm. When a lower voltage of 1 V<sub>pp</sub> is applied, the amplitude is increased, but the support loss, resulting from the mechanical elements, is also increased. Superimposing these two aspects, the synchronization region is observed to be narrowed by approximately 10 Hz as shown in Fig. 20.

The observed plateau, also known as the synchronized region, essentially corresponds to the obtainable region of doubled frequencies. So it is very desirable to widen the plateau via structural optimization and nonlinearity so as to accomplish a relatively stable and broad operating range. In light of the above results, enhancing the nonlinearity without increasing support loss seems to be necessary to widen the frequency entrainment range.

### Effect of Driving Voltage

In this study (Nakajima et al. 2012), a plateau, with a frequency ratio of 2.000 corresponding to a synchronization bandwidth (double region), was observed by spectrum analyzer (10 Hz/step with average of 10 scans), and this bandwidth can be expanded from 70 Hz to over 500 Hz when the driving AC voltage was increased from 1.0 to 10.0 V<sub>pp</sub> with the driving frequency around the subharmonic resonant frequency. Usually, the diagram which shows the relation between synchronization bandwidth of synchronized frequency responses and external force is called “Arnold’s tongue,” and the synchronized region is often represented by Arnold’s tongue, which is the region of frequency entrainment in the parameter space.



**Fig. 20** Synchronization bandwidth with respect to driving voltage (induced power) at subharmonic driving. Frequency-power sweep with subharmonic of  $\omega_1 = \omega_2/2$ . A signal generator drives the U-shaped cantilever, and the response of the beam-shaped cantilever is measured with a spectrum analyzer. The coupling is purely mechanical as a result of two coupling overhang. U-shaped oscillator system with and without mechanical element

In case of U-shaped oscillator system with mechanical element, the synchronization bandwidth can be also expanded from 70 Hz to about 550 Hz when driving AC voltage is increased from 1.0 to 10.0Vpp. A more quantitative comparison of synchronization bandwidths is further shown in Fig. 20. The largest difference occurs at 3Vpp. This expansion is believed to be related to nonlinear characteristic of the coupled oscillator system.

## Phase Noise

### Beam-Shaped Cantilever System

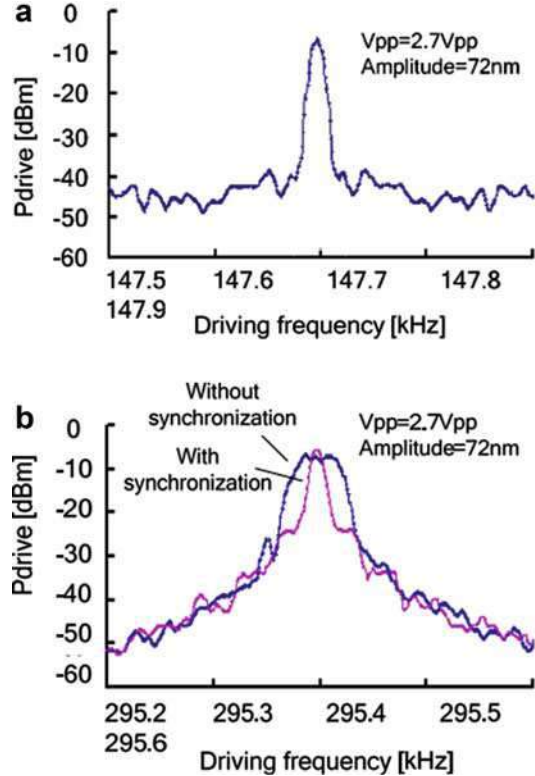
In the next place, the effect of synchronization to the phase noise was evaluated (Wang et al. 2010). The cantilever B (high-frequency cantilever) is self-oscillated at 295.5 kHz as well as above experiment, and vibration spectra were measured with and without synchronization to low-frequency cantilever A. Without additional AC driving frequency, synchronization didn't occur, but with applying AC driving voltage at 147.7 kHz, two cantilevers were synchronized. Figure 21a shows the spectrum of the cantilever A with low frequency in synchronization, and Fig. 21b shows the spectra of cantilever B with and without synchronization.

As shown in Fig. 21, occurrence of the synchronization can cause a narrow resonant peak of the self-oscillating cantilever B. This means that the self-oscillation cantilever without synchronization had a large phase noise, but this phase noise can be decreased when it synchronizes to the other cantilever (resonator) with a low phase noise.

### U-Shaped Beam Cantilever System

The phase noise of the coupled cantilevers with or without mechanical element has also been investigated (Wang et al. 2012a). In system without mechanical element, the effect of synchronization on the phase noise is also evaluated. The beam-shaped cantilever is self-oscillated at 365.713 kHz and 372.503 kHz, respectively,

**Fig. 21** (a) Spectrum of cantilever A with low frequency in synchronization; (b) Spectra of the self-oscillating cantilever B with higher frequency with and without synchronization



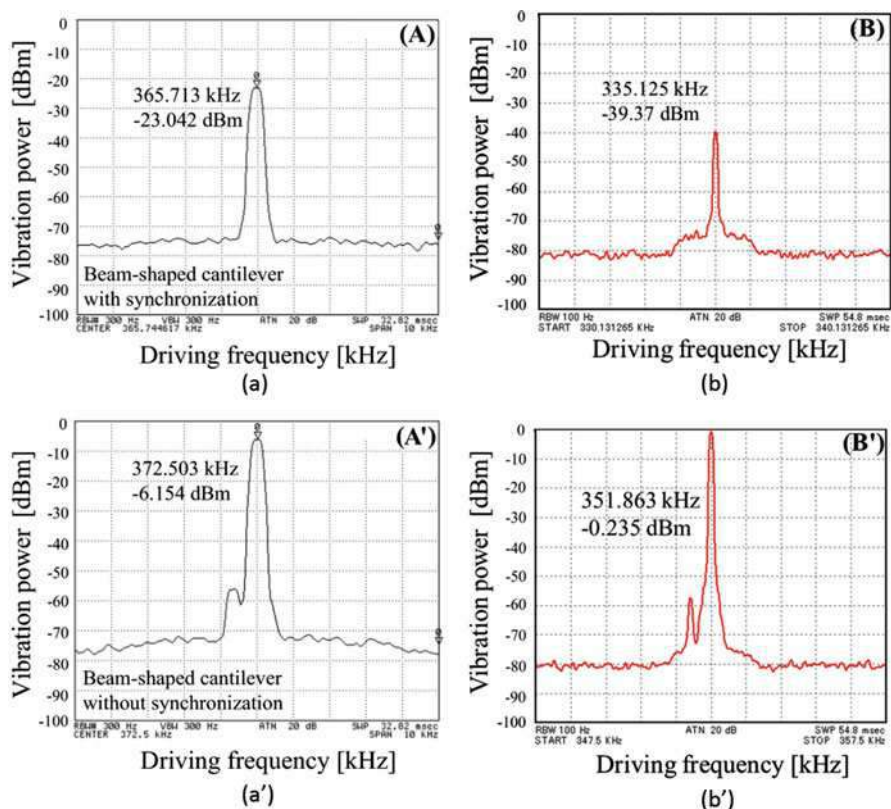
and vibration spectra were measured with and without synchronization to a low-frequency U-shaped cantilever. Without additional AC driving frequency, synchronization did not occur, but by applying AC driving voltage at 182.850 kHz, two cantilevers were synchronized.

Figure 22a, a' shows the 1 kHz span spectra of the beam-shaped cantilever with and without synchronization, and almost no apparent difference in phase noise can be observed in this work.

Since the synchronized frequency of 365.713 kHz is lower than the natural frequency of 372.503 kHz, the magnitude of synchronized vibration is also smaller than that of resonant vibration without synchronization. This can be further improved via structural design of beam-shaped cantilever to achieve similar synchronized and natural frequencies.

In system with mechanical element, a similar result can also be obtained in Fig. 22b, b' in system with mechanical elements.

Actually, the occurrence of the synchronized oscillation can narrow down the resonant peak of the self-oscillating beam-shaped cantilever, if the synchronized frequency is almost identical to the natural frequency, as reported in the related study. In this case, an almost identical magnitude can also be obtained. This implies that the larger phase noise of the self-oscillation cantilever without synchronization can be



**Fig. 22** (a) Vibration spectrum of the beam-shaped cantilever with synchronization (a') Spectrum of the beam-shaped cantilever in resonant vibration (without synchronization) (b) Vibration spectrum of the beam-shaped cantilever with synchronization (b') Spectrum of the beam-shaped cantilever in resonant vibration (without synchronization)

decreased to some degree when it is synchronized to the other cantilever (coupled oscillator) with a lower phase noise.

## Picogram Mass Sensitivity

### Beam-Shaped Cantilever System

In the above chapters, we demonstrated the synchronized cantilevers can realize frequency enhancement and phase noise suppression, which show the potential value for mass sensing via synchronization.

The work (Wang et al. 2015b) is devoted to research the possibility to achieve mass sensing of synchronized system. Geometrical design and micromanufacturing, the simplified cantilever system, consists a long cantilever ( $68.2 \mu\text{m}$  long  $\times$   $5 \mu\text{m}$  wide  $\times$   $600 \text{ nm}$  thick) and another short cantilever ( $48.7 \mu\text{m}$  long  $\times$   $5 \mu\text{m}$

wide  $\times$  600 nm thick). In light of simulation results, fabrication difficulty, as well as characterization simplicity, two coupling frequencies of 150.11 kHz (for loading and sensing) and 300.18 kHz (for unloading and detecting), respectively, were selected for experimental demonstration.

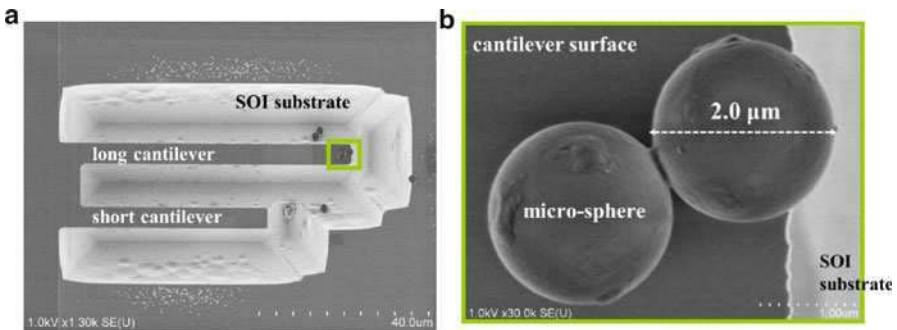
Two polystyrene microspheres were applied onto the tip of low-frequency cantilever to demonstrate small mass sensing in synchronized regime as shown in Fig. 23. The microsphere is geometrically measured as 2.0  $\mu\text{m}$  in diameter. The mass is thus calculated as 8.792 picogram with a chemical density of 1.05  $\text{g}/\text{cm}^3$  from Duke Scientific Corporation. It should be noted that the calculated 8.792 picogram is an estimated mass including measurement error of geometry and density.

### Characterization of Resonant Frequencies

In the case of no mass perturbation, the resonant frequencies of the fabricated low-frequency cantilever and high-frequency cantilever are characterized as 177.38 kHz (analytically as 150.11 kHz) in Fig. 24b and 351.66 kHz (analytically as 300.18 kHz) in Fig. 24a, respectively. The calculated frequency ratio of 1:1.983 is very close to the physical relationship of  $\omega_2 \approx 2\omega_1$ .

In the case of applying a small mass perturbation, the resonant frequencies are characterized as 171.25 kHz for low-frequency cantilever in Fig. 24a and 351.56 kHz for high-frequency cantilever in Fig. 24b, respectively. The calculated frequency ratio of 1:2.053 is also very close to the physical relationship of  $\omega_2 \approx 2\omega_1$ .

Under a measuring span of 5 kHz, a quality factor of 8079 is obtained for high-frequency cantilever, almost 3.2 times higher than that for low-frequency one. The support loss, mainly related to the geometrical design of the structure, should account for the relatively lower value of low-frequency cantilever. However, with a small mass perturbation, the quality factor of low-frequency cantilever decreased from 2524 to 1933. The decrease in Q-factor is mainly due to surface and



**Fig. 23** Micrograph of a micromechanically coupled cantilever system fabricated, characterized, and demonstrated in this work. Two polystyrene microspheres, highlighted by green rectangular marks in (a) and scaled up in (b), were applied onto the tip of the low-frequency cantilever for demonstrating small mass sensing by synchronized oscillation. The microspheres were manipulated by using micro-tweezers

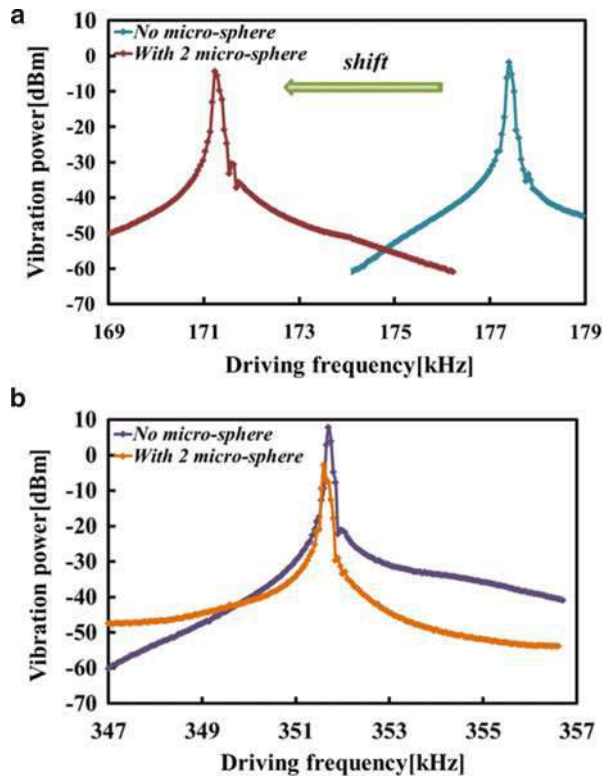
viscoelastic loss from contact to polystyrene microspheres, and the support loss should not change much with the applied small mass.

### Sensing Verification with Mass Perturbation

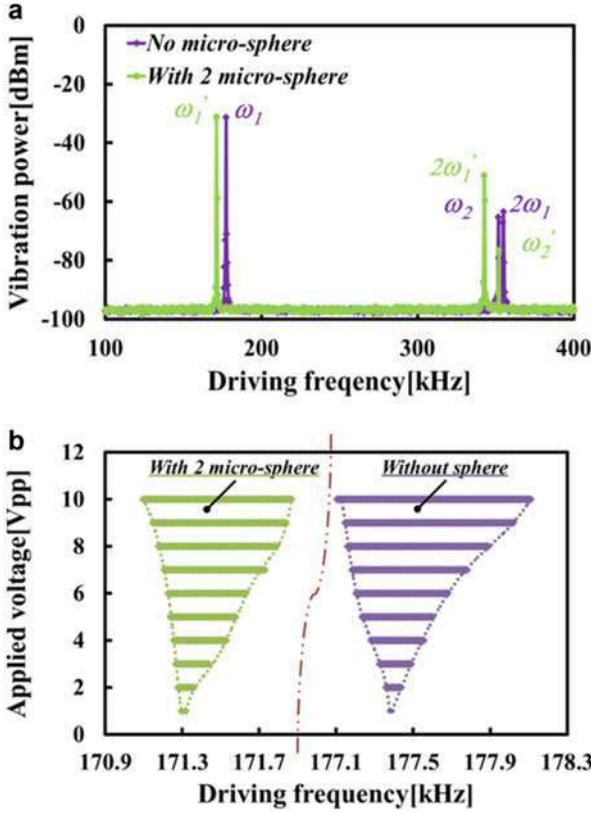
Under an applied voltage of 5.0 V<sub>pp</sub>, the vibration power (amplitude response) is measured as a function of the above driving frequency at the short cantilever tip. When the driving frequency closes to the resonant one, the vibration power reaches the maximum value; the vibration power with respect to frequency spectra is shown in Fig. 25a.

If  $\Delta m$  is the mass change, and  $k$  ( $=Ewt^3/4l^3$ ) is the spring constant, simply making use of the following well-known equation for a single beam-shaped cantilever, a mass of 8.609 picogram can be derived for the applied polystyrene microspheres based on synchronized oscillation, corresponding to the calculated value of 8.792 picogram from geometrical diameter and chemical density. For the calculation of  $k$ , the dimensions after microfabrication of 48.1  $\mu\text{m}$  in length,  $l$ , 4.63  $\mu\text{m}$  in width  $w$ , and 600 nm in thickness  $t$ , as well as a Young's modulus  $E$  of 188 GPa, are applied.

**Fig. 24** (a) The resonant frequency of low-frequency cantilever is shifted from  $\omega_1 = 177.38$  kHz to  $\omega_1 = 171.25$  kHz due to small mass perturbation (b) Almost no frequency shift is observed for high-frequency cantilever when two-microspheres were applied to low-frequency cantilever







**Fig. 25** (a) The frequencies of  $2\omega_1$  or  $2\omega_1'$  are present in the response of the short cantilever when excitation frequencies of  $\omega_1$  or  $\omega_1'$  are the resonant frequency of the long cantilever and half the resonant frequency of the short cantilever. Frequency peaks in purple are from no mass perturbation, while those in green are from small mass perturbation. The frequency shift due to small mass perturbation is also multiplied from 6.13 kHz to 12.26 kHz by a factor of 2 due to synchronized oscillation. Another two peaks, close to synchronized peaks of  $2\omega_1$  or  $2\omega_1'$ , were characterized as primary peaks of  $\omega_2$  or  $\omega_2'$  (b) Synchronized regions of  $(2\omega_1)$  versus  $(2\omega_1')$ , comprised of synchronized regime or entrainment plateau, were expressed with respect to applied voltage and driven frequency around primary resonance of low-frequency cantilever. Purple region is for no mass perturbation, while green one is for small mass perturbation. The region shift due to small mass perturbation is clearly observed from left side to right side

$$\Delta m = \frac{k}{4\pi^2} \left( \frac{1}{(\omega_2')^2} - \frac{1}{(\omega_2)^2} \right) \tag{14}$$

Similar to the preliminary demonstration in the related study with no mass perturbation, phase-locking is also observed after applying a small mass perturbation, when the driving signal is close to a half of the self-oscillating frequency of  $\omega_2 = 351.56$  kHz, the measured primary frequency of short (unloading and



detecting) cantilever. The frequency ratio of  $\omega_{\text{syn}}/\omega_1 = 2\omega_1/\omega_1 = 2.000 = \omega_{\text{syn}}'/\omega_1' = 2\omega_1'/\omega_1' = 2.000$  proves that the above small mass sensing has been accomplished in synchronized regime.

### U-Shaped Cantilever System

The U-shaped cantilever, functioning as a sensing element, is designed to sense the small mass perturbation with a lower resonant frequency of  $\omega_1$ , while the rectangular cantilever, functioning as a detecting element, is designed to detect the magnified frequency shift with a higher resonant frequency of  $\omega_2$  (Wang et al. 2016).

There are several beneficial points in the corresponding design. Firstly, relatively independent detecting cantilever is not influenced by unexpected factors caused by small mass perturbation (like analyte mass applying or gas molecules absorbing) on sensing cantilever. Secondly, U-shaped cantilever increases surface sensing area for small mass perturbation so as to increase frequency shift. Thirdly, the center-positioned rectangular cantilever is bilaterally phase-locked via two coupling overhangs from the outside U-shaped cantilever.

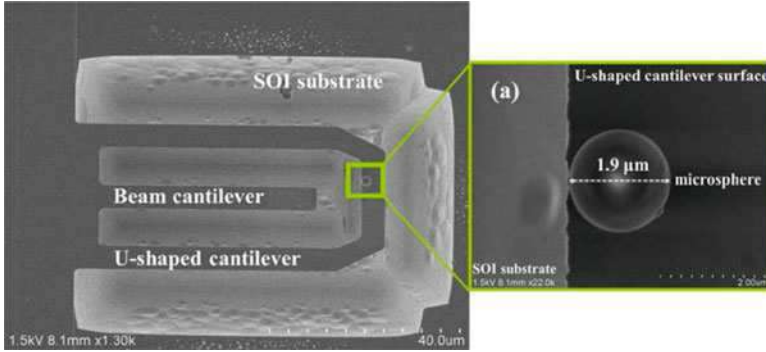
### Application of a Polystyrene Microsphere as Mass Perturbation

A polystyrene microsphere is applied onto the tip of the sensing in the case of no mass perturbation, the resonant frequencies of U-shaped cantilever and rectangular cantilever are measured as 182.506 kHz (analytically as 182.56 kHz) and 372.503 kHz, respectively. The calculated frequency ratio of 1:2.041 is very close to the theoretical relationship of  $\omega_2 \approx 2\omega_1$  as described in Fig. 26.

Under a measuring span of 100 Hz, a quality factor of 41,591 is obtained for rectangular cantilever, almost 1.7 times higher than that for U-shaped cantilever. The support loss, mainly related to the geometrical design of the structure, should account for the relatively lower value of U-shaped (sensing) cantilever. However, with a small mass perturbation, the quality factor (measuring span of 100 Hz) of U-shaped (sensing) cantilever decreased from 23,868 to 20,957. The slight drop in quality factor is mainly due to the slight deformation of the lorentzian response. The surface and viscoelastic loss from contact to the polystyrene microsphere as well as the support loss should not change much with a small absorbed mass.

### Picogram Mass Sensing in the Entrainment Range

Under certain applied voltage, a plateau with a frequency ratio of 2.0, obtained from a large amount of experimental measurements, corresponding to an entrainment range can be characterized from two bifurcation points (boundaries). Both the left and the right bifurcation points of each flat interval can be determined when the frequency ratio higher than or lower than 2.0. In light of this consideration, two entrainment regions comprised of entrainment ranges due to phase-locking were experimentally measured and comparatively shown in Fig. 27 for no mass perturbation (in purple color) and with mass perturbation (in green color), respectively. Although the entrainment region shifts from higher frequency side to lower one after applying a small mass perturbation, in both cases however, enhancing the vibration amplitude (increasing the applied voltage) results in the widening of the entrainment range.



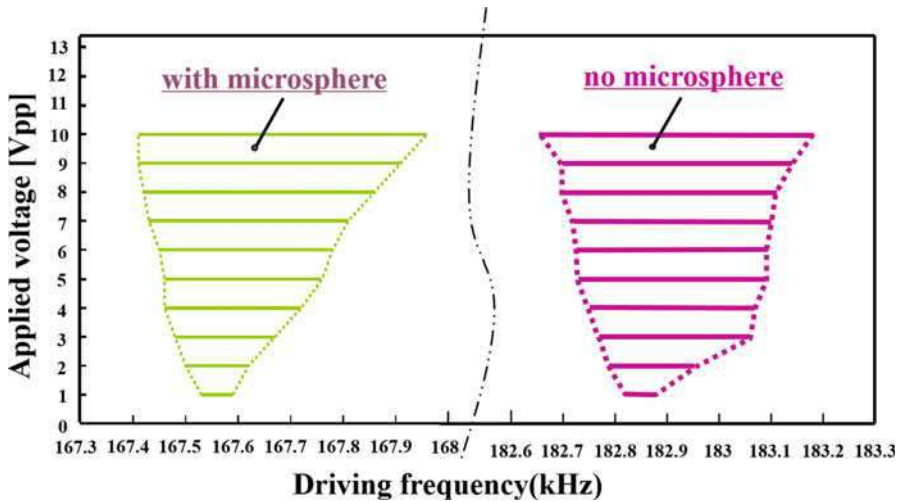
**Fig. 26** SEM micrograph of the fabricated mass-sensing device, where a polystyrene microsphere, highlighted in Fig. 14a, is applied onto the tip of the U-shaped (sensing) cantilever for demonstrating the picogram-order mass sensing by phase-locking

In the case of 1 Vpp driving voltage, the resonant frequency of sensing (U-shaped) cantilever (left part) decreased from  $\omega_1 = 182.85$  kHz to  $\omega_1' = 180.29$  kHz showing a frequency shift of 2560 Hz, while the resonant frequency of detecting (rectangular) one (right part) decreased from  $\omega_2 = 365.713$  kHz to  $\omega_2' = 360.580$  kHz showing a doubled frequency shift of 5133 Hz. For clarification, all the related frequency values are summarized in Tables 6 and 7. If  $\Delta m$  is the mass change, and  $k (=Ewt^3/4l^3)$  is the spring constant, simply making use of a well-known equation expressed as follows for a single rectangular cantilever, a mass of 3.732 picogram can be derived for the applied polystyrene microsphere based on the frequency entrainment oscillation, corresponding to the calculated value of 3.771 picogram from diameter and density. For the calculation of  $k$ , 4.2  $\mu\text{m}$  in width  $w$ , 500 nm in thickness  $t$ , 39.6  $\mu\text{m}$  in length  $l$ , and a Young's modulus  $E$  of 188 GPa are applied.

## Conclusions and Further Reading

As shown in the research mentioned above, frequency enhancement is available by using two synchronized oscillators, which consists of two geometrically different cantilevers, are coupled by coupling overhangs. The frequency response signal can thus be doubled with a decreased phase noise when the low-frequency cantilever is synchronized with the high-frequency one. Mass sensing in picogram order is demonstrated by applying a small mass perturbation onto the tip of low-frequency cantilever, and the theoretical analysis and experimental measurements of synchronized regions are discussed in this research. This phenomenon demonstrated the possibility of increasing the mass resolution via synchronization in the coupled micro-cantilevers which could be expected to be applicable to ultimate sensing of mass, force, field, and any other external perturbations.

However, high-frequency ratio of two couple oscillators are difficult to be obvious in experiment; the theoretical analysis and experimental measurements of



**Fig. 27** Entrainment regions of  $(2\omega_1)$  versus  $(2\omega_1')$ , comprised of entrainment regime or entrainment plateau due to phase-locking, were expressed with respect to applied voltage and driven frequency around primary resonance of low-frequency cantilever. The purple region is for no mass perturbation, while the green one is for small mass perturbation. The region shift due to small mass perturbation is clearly observed from the left side to the right side

**Table 6** Comparison of related frequencies with no mass perturbation

No microsphere	$\omega_1$	$\omega_2$	$\omega_2/\omega_1$
<b>Measured without phase-locking (KHz)</b>	182.506	372.503	2.041
<b>Measured with phase-locking (KHz)</b>	182.850	365.713	2.000

**Table 7** Comparison of related frequencies with small mass perturbation

No microsphere	$\omega_1'$	$\omega_2'$	$\omega_2/\omega_1'$
<b>Measured without phase-locking (KHz)</b>	180.300	371.790	2.062
<b>Measured with phase-locking (KHz)</b>	180.290	360.580	2.000

phase noise suppression and mass detection range in geometrically different cantilever array should be further researched in the future plan.

## Summary

This chapter presents an overview of picogram-order mass sensors via cantilever-based micro- or nanostructures, mainly concentrated on applying two physical phenomenon which are mode localization and synchronization from perspective of linear and nonlinear vibration to enhance the sensitivity and resolution of mass sensors. The topics include the theory of mode-localized coupled n-resonator

array; the simulation results and experimental verification of mode-localized coupled 3-, 5-, and 15-resonator array; the experimental verification of mass sensing via coupled 3-resonator array of mode localization; the theory and experiment of frequency locking and synchronized region in coupled cantilevers; the suppression of phase noise via synchronization; and finally the improving picogram mass sensitivity via synchronization in coupled micro-cantilevers. These theory and experiment result demonstrated the possibility of increasing the mass sensitivity and resolution via mode localization and synchronization in the coupled micro-cantilevers which could be expected to be applicable to ultimate sensing of mass, force, field, and any other external perturbations.

**Acknowledgments** This research was granted by the Japan Society for the Promotion of Science (JSPS) through the “Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program),” initiated by the Council for Science and Technology Policy (CSTP).

This work is partially supported by the National Natural Science Foundation of China (Grant No. 51675229). Part of this work is also financially supported by Scientific Research Foundation for Leading Professor Program of Jilin University (Grant No. 419080500171 & No. 419080500264). And this work is partially supported by Graduate Innovation of Jilin University with Project 2017072.

---

## References

- Albrecht TR, Grütter P, Horne D, Rugar D (1991) Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity. *J Appl Phys* 69(2):668–678
- Anderson PW (1958) Absence of diffusion in certain random lattices. *Phys Rev* 109:1492
- Becker P, Hymon E, Folkmer B, Manoli Y (2013) High efficiency piezoelectric energy harvester with synchronized switching interface circuit. *Sensors and Actuators A: Physical* 202:155–161
- Buks E, Roukes ML (2002) Electrically tunable collective response in a coupled micromechanical array. *J Microelectromech Syst* 11:802–807
- Chatani K, Wang DF, Ikehara T, Maeda R (2011) Amplitude enhancement using vibration mode localization with a single micro-mechanically coupled beam-shaped resonator array. In: Design, test, integration and packaging of MEMS/MOEMS (DTIP), symposium on IEEE, pp 339–343
- Chatani K, Wang DF, Ikehara T (2012) Vibration mode localization in coupled beam-shaped resonator array. In: 2012 7th IEEE international conference on. IEEE, pp 69–72
- Cleland AN, Roukes ML (2002) Noise processes in nanomechanical resonators. *J Appl Phys* 92(5):2758–2769
- Davis ZJ, Boisen A (2005) Aluminum nanocantilevers for high sensitivity mass sensors. *Appl Phys Lett* 87:013102
- Du X, Mao M, Wang DF, Zhang Z, Yang X, Wang X, Liu X, Ikehara T, Maeda R (2016) Synchronized oscillation in micro-mechanically coupled oscillator system: part III-optimization of polymer-based structures applicable to gas molecule sensing, presented at the 18th symposium on design, test, integration & packaging of MEMS/MOEMS (IEEE DTIP 2016)
- Ekinci KL, Huang XM, Roukes ML (2004a) Ultrasensitive nanoelectromechanical mass detection. *Appl Phys Lett* 84:4469–4471
- Ekinci KL, Yang YT, Roukes ML (2004b) Ultimate limits to inertial mass sensing based upon nanoelectromechanical systems. *J Appl Phys* 95(5):2682–2689
- Filler RL (1988) The acceleration sensitivity of quartz crystal oscillators: a review. *IEEE Trans Ultrason Ferroelectr Freq Control* 35(3):297–305

- Gil-Santos E, Ramos D, Jana A, Calleja M, Raman A, Tamayo J (2009) Mass sensing based on deterministic and stochastic responses of elastically coupled nanocantilevers. *Nano Lett* 9:4122–4127
- Hodges CH (1982) Confinement of vibration by structural irregularity. *J Sound Vib* 82:411
- Hosaka S, Chiyoma T, Ikeuchi A, Okano H, Sone H, Izumi T (2006) Possibility of a femtogram mass biosensor using a self-sensing cantilever. *Curr Appl Phys* 6:384–388
- Ilic B, Czaplewski D, Zalalutdinov M, Craighead HG, Neuzil P, Campagnolo C, Batt C (2001) Single cell detection with micromechanical oscillators. *J Vac Sci Technol B* 19:2825–2828
- Itoh T, Wang DF, Ikehara T, Nakajima M, Maeda R (2012a) Effect of nonlinear vibration on double region of synchronized frequency responses in mechanically coupled beam-shaped oscillator system. In: *Nano/micro engineered and molecular systems (IEEE-NEMS 2012)*, pp 136–139
- Itoh T, Wang DF, Ikehara T, Nakajima M, Maeda R (2012b) Effect of small mass perturbation on synchronized vibration in mechanically-coupled U-shaped oscillator system, *IUMRS-ICEM 2012*
- Lavrik NV, Sepaniak MJ, Datskos PG (2004) Cantilever transducers as a platform for chemical and biological sensors. *Rev Sci Instrum* 75(7):2229–2253
- Mamin HJ, Rugar D (2001) Sub-attonewton force detection at millikelvin temperatures. *Appl Phys Lett* 79(20):3358–3360
- Nakajima M, Wang DF, Ikehara T, Maeda R (2011) Synchronized oscillation in micro mechanically coupled opposite C-shaped cantilever-based oscillator system. In: *Solid-state sensors, actuators and microsystems (IEEE Transducers 2011)*, pp 1492–1495
- Nakajima M, Wang DF, Ikehara T, Itoh T, Maeda R (2012) Synchronized oscillation in micro-mechanically coupled oscillator system: part I – synchronization bandwidth, presented at the 14th symposium on design, test, integration & packaging of MEMS/MOEMS (IEEE DTIP 2012)
- Napoli M, Zhang WH, Turner K, Bamieh B (2005) Characterization of electrostatically coupled microcantilevers. *J Microelectromech Syst* 14:295–304
- Ono T, Wang DF, Esashi M (2003a) Mass sensing with resonating ultrathin double beams. In: *Proceedings of IEEE sensors, Oct 2003*, pp 825–829
- Ono T, Wang DF, Esashi M (2003b) Time dependence of energy dissipation in resonating silicon cantilevers in ultrahigh vacuum. *Appl Phys Lett* 83(10):1950–1952
- Pierre C, Dowell EH (1987) Localization of vibrations by structural irregularity. *J Sound Vib* 114:549
- Pikovsky A, Rosenblum M, Kurths J (2001) *Synchronization: a universal concept in nonlinear sciences*. Cambridge University Press, Cambridge, UK
- Requa MV, Turner KL (2006) Electromechanically driven and sensed parametric resonance in silicon microcantilevers. *Appl Phys Lett* 88(26):263508
- Sato M, Hubbard BE, Sievers AJ, Ilic B, Czaplewski DA, Craighead HG (2003) Observation of locked intrinsic localized vibrational modes in a micromechanical oscillator array. *Phys Rev Lett* 90:044102
- Shim S-B, Imboden M, Mohanty P (2007) Synchronized oscillation in coupled nanomechanical oscillators. *Science* 316(5821):95–99
- Spletzer M, Raman A, Wu AQ, Xu X, Reifenberger R (2006) Ultrasensitive mass sensing using mode localization in coupled cantilever. *Appl Phys Lett* 88:254102
- Spletzer M, Raman A, Sumali H, Sullivan JP (2008) Highly sensitive mass detection and identification using vibration localization in coupled microcantilever arrays. *Appl Phys Lett* 92:114102
- Stowe TD, Yasumura K, Kenny TW, Botkin D, Wago K, Rugar D (1997) Attonewton force detection using ultrathin silicon cantilevers. *Appl Phys Lett* 71(2):288–290
- Tang L, Yang Y (2011) Analysis of synchronized charge extraction for piezoelectric energy harvesting. *Smart Mater Struct* 20(8):085022
- Thiruvengathanathan P, Yan J, Woodhouse J, Aziz A, Seshia AA (2010) Ultrasensitive mode-localized mass sensor with electrically tunable parametric sensitivity. *Appl Phys Lett* 96:081913

- Thiruvengathan P, Woodhouse J, Yan J, Seshia AA (2011) Limits to mode-localized sensing using micro- and nanomechanical resonator arrays. *J Appl Phys* 109:104903
- Wang DF, Ono T, Esashi M (2003) Crystallographic influence on nanomechanics of (100)-oriented silicon resonators. *Appl Phys Lett* 83(15):3189–3191
- Wang DF, Ono T, Esashi M (2004) Thermal treatments and gas adsorption influences on nanomechanics of ultra-thin silicon resonators for ultimate sensing. *Nanotechnology* 15(12):1851–1854
- Wang DF, Feng J, Ono T, Esashi M, Ye X (2010) Micromechanically coupled resonated system for synchronized oscillation with improved phase noise. In: *Proceedings of the IEEE 23rd international conference on MEMS*, Jan 2010, pp 703–706
- Wang DF, Ikehara T, Nakajima M, Maeda R (2012a) Characterization of micromechanically coupled U-shaped cantilever-based oscillators with and without micromechanical elements for synchronised oscillation-based applications. *Micro Nano Lett* 7(2):188–192
- Wang DF, Chatani K, Ikehara T (2012b) Mode localization analysis and characterization in a 5-beam array of coupled nearly identical micromechanical resonators for ultra-sensitive mass detection and analyte identification. *Microsyst Technol* 18:1923–1929
- Wang DF, Chatani K, Ikehara T (2012c) Observation of localised vibration modes in a micro-mechanical oscillator trio with coupling overhang for highly sensitive mass sensing. *Micro Nano Lett* 7:713–716
- Wang DF, Itoh T, Ikehara T, Maeda R (2012d) Doubling flexural frequency response using synchronized oscillation in a micro-mechanically coupled oscillator system. *Micro Nano Lett* 7(8):717–720
- Wang DF, Li X, Yang X (2015a) Enhancing amplitude changes by mode localization in trio cantilevers with mass perturbation. *J Micromech Microeng* 25:095017
- Wang DF, Du X, Wang X, Ikehara T, Maeda R (2015b) Multiplication in frequencies by synchronized and super-harmonic oscillations: sensing verification with picogram order micro-spheres. *IEEE Sensors J* 15(8):4464–4471
- Wang DF, Du X, Wang X, Ikehara T, Maeda R (2016) Improving picogram mass sensitivity via frequency doubling in coupled silicon micro-cantilevers. *J Micromech Microeng* 26:015006
- Zalalutdinov M et al (2003) Frequency entrainment for micromechanical oscillator. *Appl Phys Lett* 83(16):3281–3283
- Zhang W, Turner KL (2005) Application of parametric resonance amplification in a single-crystal silicon microoscillator based mass sensor. *Sens Actuators* 122:23–30



# Micromachined Humidity Sensors

Jianqiu Huang and Qing-An Huang

## Contents

Introduction .....	788
Definition of Humidity .....	788
Absolute Humidity .....	789
Relative Humidity .....	789
Dew Point/Frost Point .....	790
Parts per Million .....	790
Classification of Humidity Sensors .....	790
Resistive Humidity Sensors .....	791
Piezoresistive Humidity Sensors .....	792
Piezoresistance Basics .....	793
Membrane Structure .....	794
Cantilever Structure .....	797
Capacitive Humidity Sensors .....	801
Interdigitated Structure .....	801
Sandwich Structure .....	804
Sandwich-Interdigitated Structure .....	805
Self-Heating Technology .....	807
Package of Humidity Sensors .....	810
Conclusions .....	813
References .....	813

## Abstract

Humidity sensors play a significant role in agriculture, industry, household, medicine, and so on. Traditional mechanical humidity sensors have been replaced by electrical humidity sensors since the revolution of electronic industry. A lot of

---

J. Huang (✉) · Q.-A. Huang

Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, Jiangsu, China

e-mail: [hjq@seu.edu.cn](mailto:hjq@seu.edu.cn); [hqa@seu.edu.cn](mailto:hqa@seu.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies, [https://doi.org/10.1007/978-981-10-5945-2\\_24](https://doi.org/10.1007/978-981-10-5945-2_24)

787

efforts have been made to perfect the design and fabrication of the electrical humidity sensors.

This chapter will discuss some researches of the electrical humidity sensors with the aid of micromachining technologies. Here, a humidity sensor is not an instrument. Instead, it is an electronic device which is fabricated on a wafer just like a MOS FET. However, unlike transistors, design and fabrication of micromachined humidity sensors involve the knowledge of electrics, chemistry, mechanics, and so on.

We aim to give a general view on the micromachined humidity sensors. Several design and fabrication technologies will be reviewed in this chapter.

---

**Keywords**

Microfabrication · Micromachining · CMOS MEMS · Nanomaterials · Humidity sensors · Resistive sensors · Piezoresistive sensors · Capacitive sensors · Self-heating · Package

---

## Introduction

With the emergence of microelectromechanical systems (MEMS), various micromachined humidity sensors have been developed (Farahani et al. 2014; Islam et al. 2015; Su et al. 2006; Steffens et al. 2014). Micromachined humidity sensors offer numerous advantages such as outstanding properties (Borini et al. 2013; Zhang et al. 2016), small-sizes (Bakhoun and Cheng 2012; Baltes 1993; Molina-Lopez et al. 2013), batch fabrication (Baltes 1993), and compatibility with digital electronics (Dai 2007; Eder et al. 2014).

The goal of this chapter is to review some basic knowledge and generic technology of the micromachined humidity sensors. At the beginning, the [definition of humidity](#) is stated. After that, the [classification of humidity sensors](#) is introduced. Although there are various kinds of humidity sensors, three widely used micromachined humidity sensors are discussed, including [resistive humidity sensors](#), [piezoresistive humidity sensors](#) and [capacitive humidity sensors](#). Then the [self-heating technology](#) is presented. At last, the package technology of the micromachined humidity sensor is proposed (section “[Package of Humidity Sensors](#)”).

---

## Definition of Humidity

As a mixture of gases, air consists of nitrogen, oxygen, carbon dioxide, water vapor, and other trace gases. “Humidity” is the measurement of the amount of water vapor in air. According to Dalton’s law of partial pressures, the total pressure of a mixed gas is equal to the sum of the partial pressure of different components (E+E ELEKTRONIK n.d.):



$$\begin{aligned}
 P_t &= P_{N_2} + P_{O_2} + P_{CO_2} + P_w + \dots \\
 \text{or} \\
 P_t &= P_{\text{dry}} + P_w
 \end{aligned}
 \tag{1}$$

where  $P_t$  is the total pressure of air.  $P_{O_2}$ ,  $P_{CO_2}$ ,  $P_w$ , and  $P_{\text{dry}}$  represent the partial pressure of  $O_2$ ,  $N_2$ , water vapor, and dry air, respectively.  $P_w$ , which is usually specified by Pa or mbar, is a key parameter to define humidity related parameters.  $P_w$  is only decided by the content of water vapor in air and the pressure of air.

Based on different measurement techniques, several parameters are used to describe the ‘‘humidity,’’ including absolute humidity, relative humidity, dew point, mixing ratio, and parts per million (E+E ELEKTRONIK n.d.; Oyj 2013, n.d.).

### Absolute Humidity

Absolute humidity is the mass of water vapor in a unit volume of air, which is usually specified in  $\text{g}/\text{m}^3$ .

$$a = \frac{m_w}{V} \tag{2}$$

where  $a$  represents the absolute humidity,  $m_w$  is the mass of water vapor, and  $V$  is the volume of air. The absolute humidity can be easily calculated from the water vapor partial pressure  $P_w$ :

$$a = C \frac{P_w}{T} \tag{3}$$

where  $C$  is a constant equal to 216.679 g,  $P_w$  is the partial pressure of water vapor, and  $T$  is the temperature in K.

### Relative Humidity

Evaporation and condensation occurs at the same time in air. In a closed volume, as evaporation and condensation reach equilibrium over water, the partial pressure of water vapor ( $P_w$ ) reaches the maximal value, which is called the saturation vapor pressure ( $P_{ws}$ ).  $P_{ws}$  increases with the temperature. Water vapor remains in gas state when  $P_w < P_{ws}$  while it is condensed into liquid when  $P_w > P_{ws}$ . The most important parameter related to  $P_{ws}$  is the temperature of air.

Relative humidity (RH) is defined as the ratio of water vapor partial pressure ( $P_w$ ) to the saturation pressure ( $P_{ws}$ ) at a certain temperature. The RH is usually expressed as a percentage.

$$\text{RH} = \frac{P_w}{P_{ws}} \times 100\% \tag{4}$$

Relative humidity is a temperature dependent parameter.

### Dew Point/Frost Point

As discussed above,  $P_{ws}$  increases with the temperature. In other words, water vapor in air can be condensed into liquid by decreasing the temperature. The temperature at which the condensation just happens and  $P_w = P_{ws}$  is defined as the dew point ( $T_d$ , if  $>0^\circ\text{C}$ ) or frost point ( $T_f$ , if  $<0^\circ\text{C}$ ).  $T_{d/f}$  is a function of water vapor partial pressure but not a temperature-dependent parameter.

### Mixing Ratio

Mixing ratio is the ratio of the mass of water vapor and dry air, which is usually specified in g/kg.

$$X = \frac{m_w}{m_{\text{dry}}} \quad (5)$$

where  $X$  represents the mixing ratio.  $m_x$  and  $m_{\text{dry}}$  are the mass of water vapor and other gases in air, respectively. The factor which affects mixing ratio is the water vapor content in air.

### Parts per Million

Parts per million by volume ( $\text{PPM}_v$ ) is defined as the volume of water vapor per total volume of dry gas. The parameter can be calculated from the water vapor partial pressure  $P_w$ :

$$\text{PPM}_v = \frac{P_w}{(P_t - P_w)} \times 10^6 \quad (6)$$

where  $P_w$  is the vapor partial pressure of water and  $P_t$  is the total pressure of air.

Parts per million by mass ( $\text{PPM}_m$ ), which is defined as the mass of water vapor per total mass of the dry air, can be obtained according to the data of  $\text{PPM}_v$ .

$$\text{PPM}_m = \frac{M_w P_w}{M_{\text{dry}} (P_t - P_w)} \times 10^6 \quad (7)$$

where  $M_w$  is the mole weight of water and  $M_{\text{dry}}$  is the average mole weight of other gases in air.

---

## Classification of Humidity Sensors

As humidity has different definitions, humidity sensors can be classified base on the "humidity" terms. Among all of the humidity sensors, relative humidity sensors and absolute humidity sensors are normally applied and actually most of the

micromachined humidity sensors are relative humidity sensors (Farahani et al. 2014). The humidity sensors can be classified according to the sensitive material as well. However, because of the variety of materials, different classifications were given by researchers (Chen and Lu 2005; Traversa 1995; Yamazoe and Shimizu 1986). Within all the humidity sensitive materials, ceramic and organic polymers are commonly used. The ceramic materials used for humidity sensors including  $\text{Al}_2\text{O}_3$  (Kim et al. 2009a),  $\text{ZnO}$  (Erol et al. 2011),  $\text{TiO}_2$  (Gu et al. 2011),  $\text{BaTiO}_3$  (He et al. 2010),  $\text{BaZrO}_3$  (Chen et al. 2010),  $\text{WO}_3$  (Dai et al. 2007), and so on. The polymers used for humidity sensors including cellulose acetate butyrate (CAB) (Duc  r   et al. 2005), polyimide (Lee et al. 2011), polyelectrolytes (Lee et al. 2005), polypyrro (Zhang et al. 2008), polyaniline (Nohria et al. 2006), and so on.

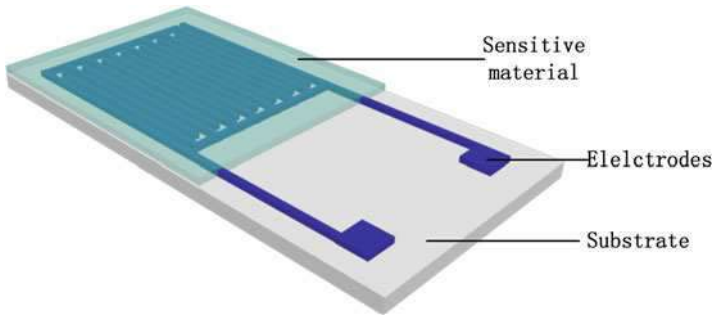
As to micromachined humidity sensors, it is convenient to divide the sensors according to the operating principle. According to the classification provided by Ralph Fenner et al., the micromachined humidity sensors consist of four types: capacitive, mass-sensitive, optical, and resistive (Fenner and Zdankiewicz 2001). The mass-sensitive humidity sensors are composed of bulk acoustic wave (BAW) humidity sensors and surface acoustic wave (SAW) humidity sensors. The optical humidity sensors have two operating models: absorptive and reflective. As described by another paper reviewed by Z. M. Rittersma et al., the micromachined humidity sensors consist of capacitive humidity sensors, resistive humidity sensors, hygrometric humidity sensors, gravimetric (mass-sensitive) humidity sensor, and optical humidity sensors (Rittersma 2002). Moreover, as presented by Chia-Yen Lee et al., humidity sensors are classified into optical humidity sensors, gravimetric humidity sensors, capacitive humidity sensors, resistive humidity sensors, piezoresistive humidity sensors, and magnetoelastic humidity sensors (Lee and Lee 2005). Although various kinds of humidity sensors were reported, three of them are commercially available and generally utilized. They are resistive humidity sensors, piezoresistive humidity sensors, and capacitive sensors (Comparison of Relative Humidity Sensing Technologies 2004).

---

## Resistive Humidity Sensors

Resistive humidity sensors are one of the earliest developed electrical humidity sensors (Lai and Hidy 1968). In resistive humidity sensors, the resistance or impedance of the sensitive material between two electrodes is affected by the humidity. As shown in Fig. 1, interdigitated electrodes fabricated on an insulated substrate are commonly used in most of the resistive humidity sensors. The performances of these sensors mainly depend on the sensitive material deposited on the electrodes.

Various materials are utilized as resistive sensitive materials. These materials are mainly divided into three types: ceramics, polymers, and electrolytes (Lee and Lee 2005). Normally, the resistance of the sensors decreases with the humidity because water molecules promote the flow of ions in sensitive materials (Zhang et al. 2009). However, in some other resistive humidity sensors the resistance increases with the humidity. For example, multiwall carbon nanotubes can be dispersed into the



**Fig. 1** Schematic of a resistive humidity sensor

polyimide to form a resistive material (Tang et al. 2011). As water molecules diffuse into the material, the polyimide swells and increases the distance between the nanotubes. As a result, the resistance of the sensitive material rises.

Recently, various two-dimensional materials have been developed and some of them exhibit excellent properties in the resistive sensitive applications. Wen-Hao Chen et al. reported the application of ultrasonic exfoliated black phosphorus nanosheets as the sensitive material of a humidity sensor (Chen et al. 2016). The sensor shows large sensitivity from 40–70%RH as illustrated in Fig. 2. As the nanosheet layer is about 77 nm, the response time of this sensor is very short. Moreover, Chongyang Zhu et al. treated the black phosphorus crystal with household kitchen blender and formed black phosphorus quantum dots (BPQDs) (Zhu et al. 2016). BPQDs show different properties from the nanosheets. As the humidity changes from 10% to 90%, the resistance of the BPQDs film varies by  $\sim 4$  orders of magnitude.

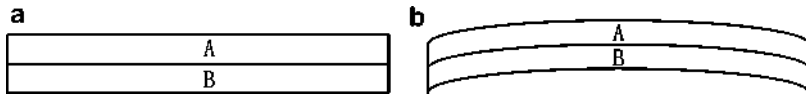
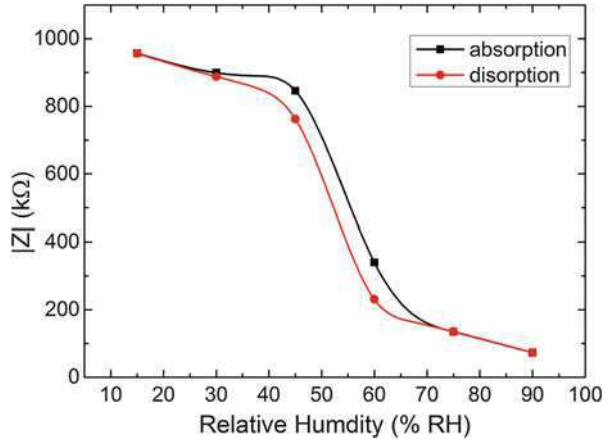
---

## Piezoresistive Humidity Sensors

Resistive-based humidity sensors can be divided into two classes, those which measure the change in conductivity and those measure the change in mechanical deformation. The first kind is the resistive humidity sensors as introduced above and the second one is the piezoresistive humidity sensors described in this section. The performances of resistive humidity sensors are directly affected by the electrical characteristics of sensitive materials. It usually leads to nonlinear properties of resistive sensors at low relative humidity (Packirisamy et al. 2005). In piezoresistive humidity sensors, sensitive materials are separated into humidity-sensitive materials and mechanical-sensitive materials. As a result, the performances of piezoresistive humidity sensors are improved (Huang et al. 2015).

As shown in Fig. 3 the piezoresistive humidity sensors utilize the expansion of the humidity-sensitive layer (A) due to the uptake of water molecules. The expansion leads to a stress change in the mechanical layer (B). The piezoresistor

**Fig. 2** Humidity measurement results of a black phosphorus humidity sensor



**Fig. 3** Operating principle of the piezoresistive humidity sensor: (a) before and (b) after uptake of water molecules

(mechanical-sensitive materials) in the mechanical layer (B) transforms the stress change into an electrical signal.

### Piezoresistance Basics

The piezoresistive effect dates back to the experiment reported by William Thomson (Lord Kelvin) in 1856 (Thomson 1856). In the experiment, both a copper and an iron wire with the same sizes were stretched together. Although deformations of the two wires were the same, changes of their resistance were different. It implied the resistivity was related to the mechanical load as well. Tomlinson tested the piezoresistive effect in different metals and reported the response curves under varied orientations of mechanical loads (Tomlinson 1883).

In 1962, Bridgman started to use a tensor formulation to characterize the stress induced change in resistance (Dean and Douglas 1962). The piezoresistive effect can be described as follows:

$$\frac{\Delta\rho}{\rho} = \pi\sigma \tag{8}$$

where  $\rho$  is the unstressed resistivity,  $\sigma$  is the mechanical stress, and  $\pi$  is the piezoresistive coefficient.

Silicon is widely used in piezoresistive devices because its piezoresistive coefficient is much larger than that of metals (Smith 1954). In single crystal silicon the piezoresistive effect can be described in the form of tensor (Allen 2005; Smith 1954):

$$\frac{1}{\rho} \begin{Bmatrix} \Delta\rho_{xx} \\ \Delta\rho_{yy} \\ \Delta\rho_{zz} \\ \Delta\rho_{xy} \\ \Delta\rho_{yz} \\ \Delta\rho_{zx} \end{Bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{11} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{44} \end{bmatrix} \begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{xy} \\ \sigma_{yz} \\ \sigma_{xz} \end{Bmatrix} \quad (9)$$

where  $\rho_{ij}$ ,  $\pi_{ij}$ ,  $\sigma_{ij}$  are the components of resistivity, piezoresistive coefficient, and stress, respectively (Fig. 4).

Also, the piezoresistive effect can be expressed as (Smith 1954):

$$\frac{\Delta\rho}{\rho} = \pi_L \sigma_L + \pi_T \sigma_T \quad (10)$$

where  $\pi_L$  and  $\pi_T$  are longitudinal and transverse piezoresistive coefficient, respectively.

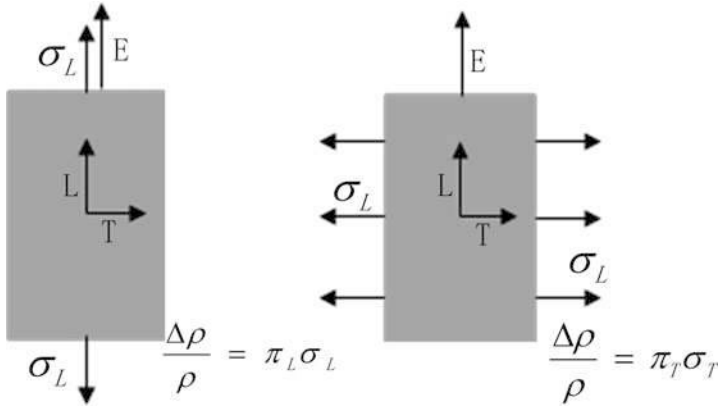
$$\begin{aligned} \pi_L &= \pi_{11} + 2(\pi_{44} + \pi_{12} - \pi_{11})(l_1^2 m_1^2 + l_1^2 n_1^2 + m_1^2 n_1^2) \\ \pi_T &= \pi_{12} - 2(\pi_{44} + \pi_{12} - \pi_{11})(l_1^2 l_2^2 + m_1^2 m_2^2 + n_1^2 n_2^2) \end{aligned} \quad (11)$$

where  $l$ ,  $m$ , and  $n$  are the direction cosines of the direction associated with  $\pi_L$  or  $\pi_T$ , with respect to the crystallographic axes.

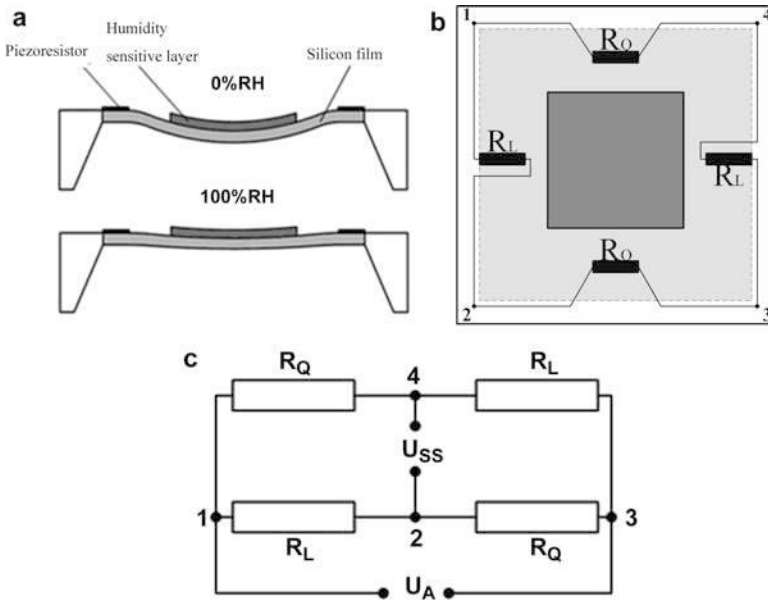
## Membrane Structure

A typical piezoresistive humidity sensor with a square membrane structure is shown in Fig. 5. In the structure a humidity sensitive layer is deposited on a silicon membrane. Only the central of the membrane is covered by the humidity sensitive material to optimize the performances of the sensor. Four piezoresistors formed by doped silicon are located near each edge of the membrane. As shown in Fig. 5b, two of the piezoresistors are parallel to the edge (RL) and the other two are perpendicular to the edge (RQ). The four resistors are connected to a full Wheatstone bridge which transfers variations of resistance into voltage outputs.

A circular or a rectangular membrane can be used as a piezoresistive sensitive structure as well. However, according to the FEM analysis, the shape of the membrane has little effect on the sensitivity of the sensor (Buchhold et al. 1998). The sensitivity mainly depends on the thickness of both the humidity sensitive layer and the silicon membrane. The sensitivity increases with the thickness of the sensitive layer and decreases with the silicon layer. A thicker sensitive layer will

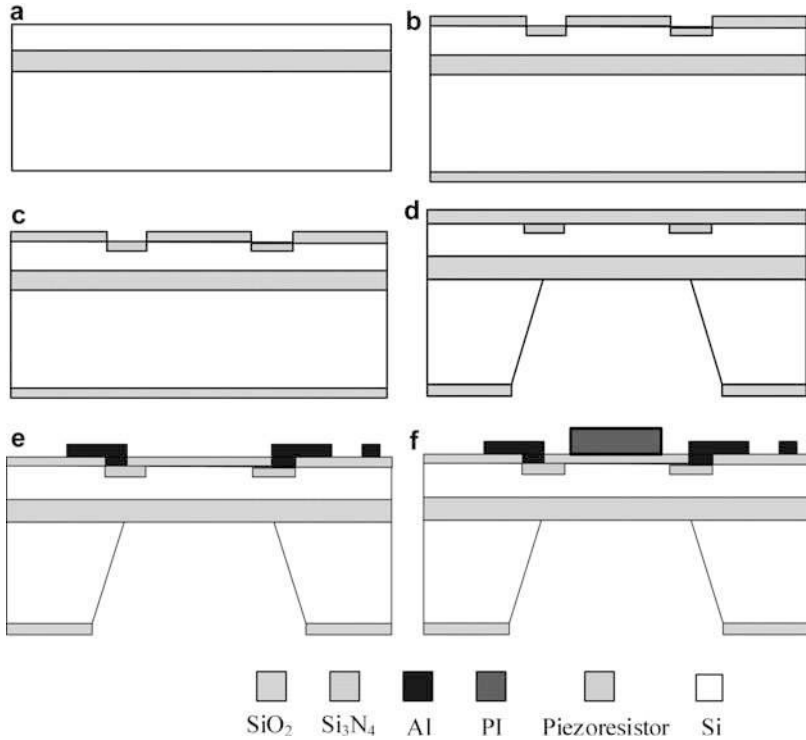


**Fig. 4** Longitudinal and transverse piezoresistive effect



**Fig. 5** Piezoresistive humidity sensor with membrane structure: (a) side view; (b) top view; (c) the equivalent circuit of the sensor

enhance the stress in the membrane but induce a longer response time of the sensor. A thinner silicon membrane is preferred to enlarge the sensitivity of the sensor but it is limited by the fabrication technology and the reliability of the structure. The sensitivity is also related to the coverage rate, the pre-stress, and the stress coefficient of the sensitive layer. The coverage rate is defined as the ratio between the thickness of the sensitive layer to that of the silicon membrane. For maximum sensitivity, the



**Fig. 6** Fabrication process of a membrane type piezoresistive humidity sensor

coverage rate is about 0.8. The pre-stress in the sensitive layer is the main cause of the offset voltage. It also induces the degradation of the sensitivity of the sensor. The stress coefficient  $L$  is determined by the material properties of the sensitive material:

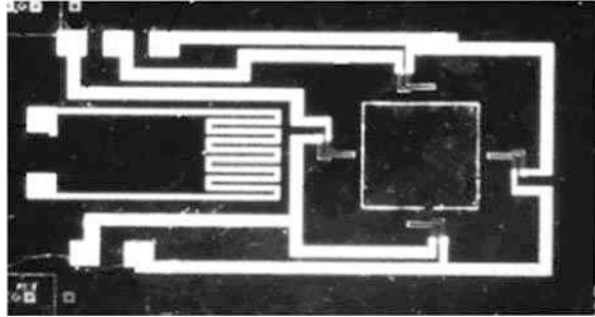
$$L = \frac{E_2}{(1 - \nu_2)} \alpha_\varphi \tag{12}$$

where  $E$  and  $\nu$  are respectively the Young’s modulus and the Poisson’s ratio of the sensitive material.  $\alpha$  is the humidity expansion coefficient of the sensitive material. A material with a large stress coefficient should be used to form the sensitive layer.

The membrane type piezoresistive humidity sensor can be fabricated by a bulk micromachining process as shown in Fig. 6. The fabrication process starts with an N-type SOI wafer. First, a thermal  $\text{SiO}_2$  is grown and patterned. Then piezoresistors are formed by implanting the silicon directly with boron. Second, a LPCVD  $\text{Si}_3\text{N}_4$  film is deposited on both sides of the wafer with  $\text{SiO}_2$  as a buffer layer. Third, the  $\text{Si}_3\text{N}_4/\text{SiO}_2$  film is patterned on the backside of the wafer. The wafer is wet-etched with the protection of  $\text{SiO}_2/\text{Si}_3\text{N}_4$  and stopped at the buried  $\text{SiO}_2$  layer. After the wet-etching process,  $\text{Si}_3\text{N}_4$  is removed on both sides by wet etching. Fourth, the  $\text{SiO}_2$  on the front side is patterned for the contact area. Then an aluminum layer is sputter-deposited



**Fig. 7** Photograph of a membrane type piezoresistive humidity sensor



and patterned as wires and electrodes. At last, polyimide is deposited on the front side and patterned as the sensitive material.

A photograph of the fabricated piezoresistive humidity sensor is shown in Fig. 7. As reported by Xiaoyan Cai et al. (2009), the sensitivity of the sensor is 0.236 mV/% RH and the hysteresis of the sensor is less than 5%RH. The temperature drift of the sensor is about 0.06 mV/°C.

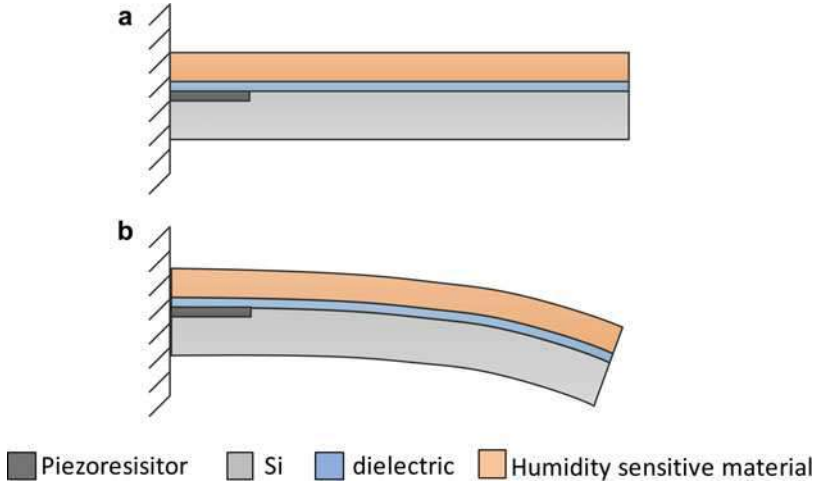
## Cantilever Structure

The operating principle of a cantilever type piezoresistive humidity sensor is depicted in Fig. 8. Usually, four cantilevers are designed with a piezoresistor on each beam. The piezoresistors are placed near the fix end of the beam as shown in Fig. 9. Similar to the membrane structure, two of the piezoresistors are parallel to the fix end (RL) and the other two are perpendicular to the fix end (RQ). They are connected into a full Wheatstone bridge.

According to the stress analysis of the composite film, the sensitivity of the sensor mainly depends on the thickness of each layer and the distance between the piezoresistors and the neutral plane of the beam. The beam with a thicker sensitive layer shows higher sensitivity but slower response time.

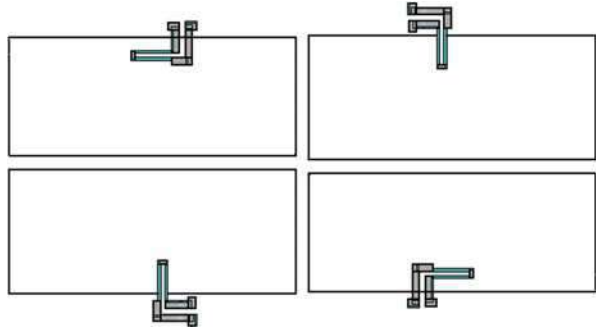
For a typical cantilever type piezoresistive humidity sensor, a finite element model can be used to optimize the size of the sensor. It implies that the shape of the cantilever will affect the sensitivity of the sensor. The sensitivity decreases with the length of the cantilever but increases with the width of the cantilever as shown in Fig. 10. The thickness of the silicon layer is also related to the sensitivity as shown in Fig. 11.

Usually, a bulk micromachining process is used to fabricate a piezoresistive humidity sensor (Jackson et al. 2008). However, the machining precision is always limited by the dual side alignment and bulk wet etching process. Figure 12 illustrates a surface machined fabrication process of the cantilever type piezoresistive humidity sensor (Huang et al. 2015). Fabrication starts with an N type (100) wafer. A two-step etching process is performed as shown in Fig. 12a. First, an array of grooves is fabricated by anisotropic dry etching. Next, a cavity is fabricated by isotropic dry



**Fig. 8** Operating principle of a cantilever type piezoresistive humidity sensor

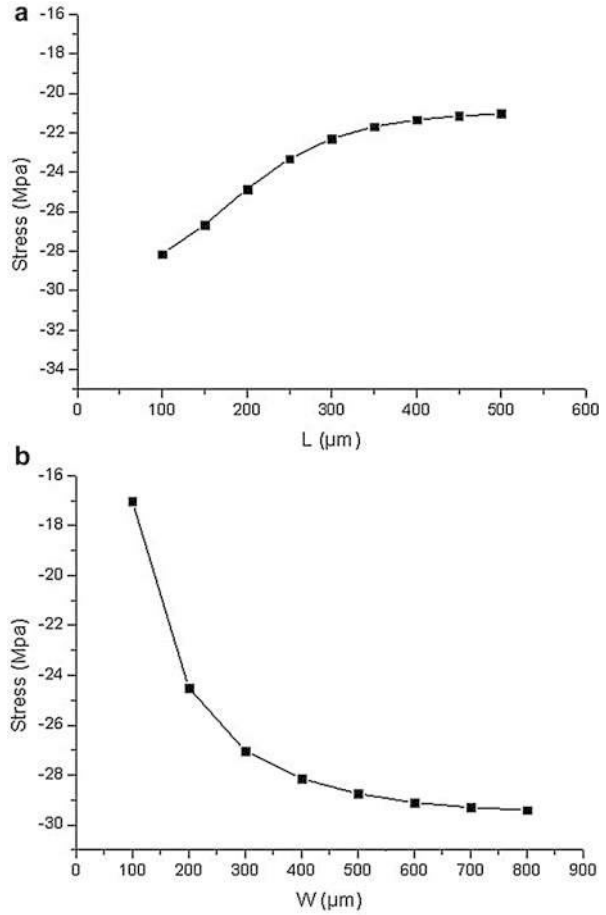
**Fig. 9** Piezoresistors in the cantilever type piezoresistive humidity sensor



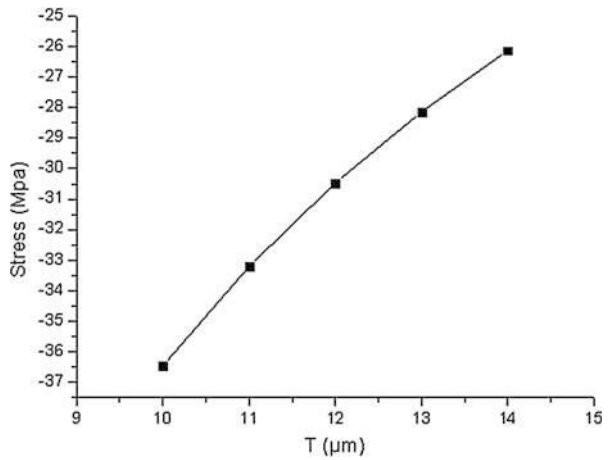
etching below the grooves. Then the cavity is sealed by an epitaxial growth process. As follows, a typical CMOS process is used to form piezoresistors, dielectric and metal wires, as shown in Fig. 12c. In the CMOS process, ion-implanted p-type resistors are fabricated as the piezoresistors. After the CMOS process, a polyimide layer is spun on the wafer and patterned to form the humidity-sensitive material. Lastly, an anisotropic dry etching process is used to release the micro-cantilevers (Fig. 12d).

Figure 13 shows the details of the cantilevers. Considering the sizes of the whole devices and the limitation of fabrication process, the optimum length and width of the cantilever are  $100\ \mu\text{m}$  and  $400\ \mu\text{m}$  and the thickness of the silicon layer is  $10\ \mu\text{m}$ . It can be seen that a network of ripples distributed on the cantilever. The ripples resulted from the grooves fabricated at the first step, as shown in Fig. 12a. Moreover, the shapes and distributions of the grooves can be designed to achieve a stress concentrated structure.

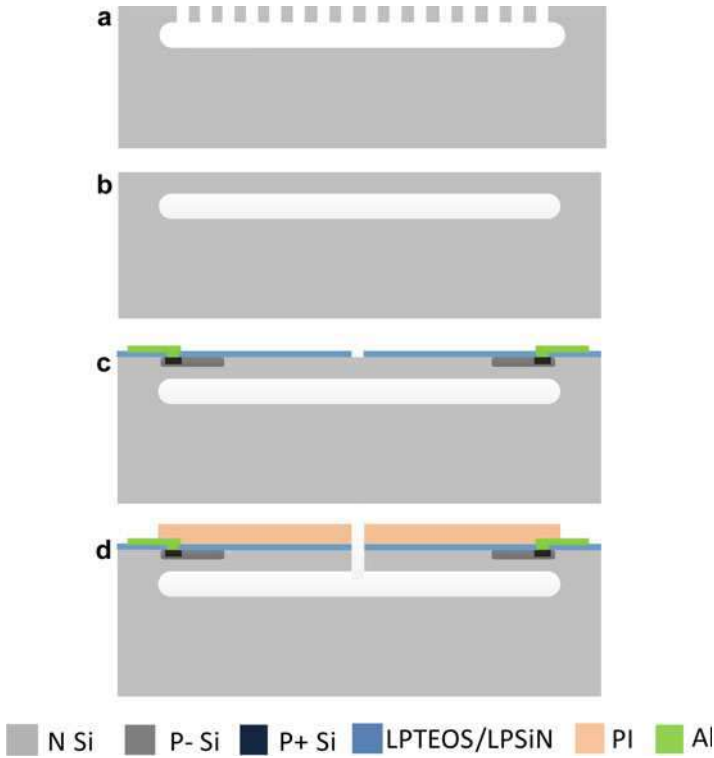
**Fig. 10** Stress analysis for beams with different length and width



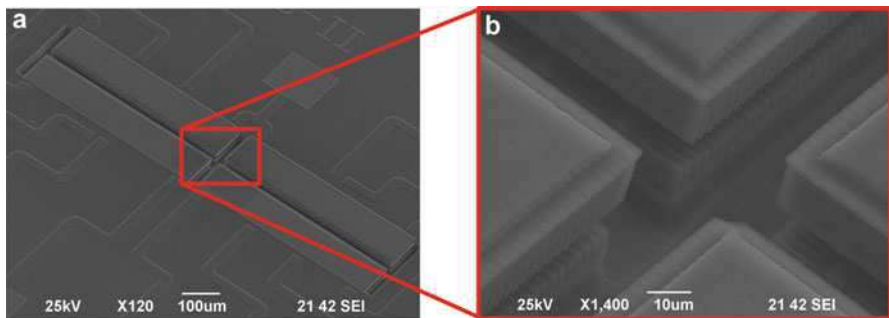
**Fig. 11** Stress analysis for beams with different thickness of the silicon layer



According to the test results, the sensitivity of the sensor is 7 mV/%RH and the linearity of the sensor is 1.9% at 20 °C. Both the sensitivity and linearity are not sensitive to the temperature. However, the hysteresis of the humidity sensor



**Fig. 12** Surface machining fabrication process of the cantilever type piezoresistive humidity sensor



**Fig. 13** SEM view of the cantilevers in the piezoresistive humidity sensor

decreases from 3.2% to 1.9% RH as the temperature increases from 10 to 40 °C. The recovery time of the sensor is 85 s at a room temperature (25 °C).

## Capacitive Humidity Sensors

In capacitive humidity sensors, the dielectric is a sensitive material. As illustrated in Fig. 14 the dielectric constant of the sensitive material  $\epsilon_s$  changes with the humidity (Schubert and Nevin 1985):

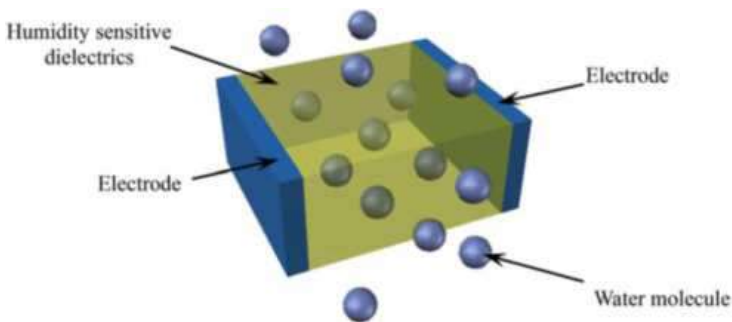
$$\epsilon_s = \left[ \gamma \left( \epsilon_w^{1/3} - \epsilon_p^{1/3} \right) + \epsilon_p^{1/3} \right]^3 \quad (13)$$

where  $\epsilon_w$  and  $\epsilon_p$  are the dielectric constant of water and the sensitive material without water, respectively.  $\gamma$  is the fractional volume of water molecules in the sensitive material. Basically, there are two kinds of capacitive structure: one is the interdigitated structure and the other is the sandwich structure (Comparison of Relative Humidity Sensing Technologies, n.d.). Based on these two structures, several improved structures were developed to improve the properties of capacitive humidity sensors.

### Interdigitated Structure

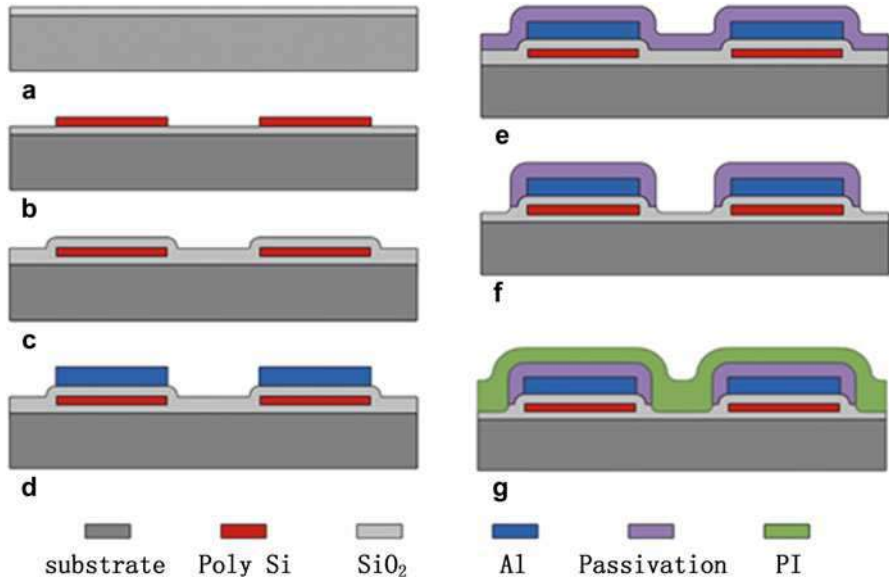
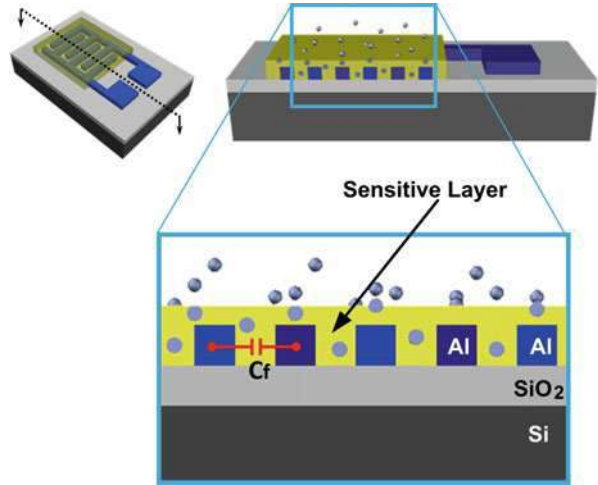
The interdigitated structure (Fig. 15) is widely used because of its simple fabrication processes (Gu et al. 2004; Nizhnik et al. 2012). In gas sensors, the interdigitated structure is usually applied because one side of this structure can be open to the ambient conditions.

The interdigitated structure can be easily fabricated by a CMOS compatible process (Zhao et al. 2011a). As an example in Fig. 16, first, a thermal SiO<sub>2</sub> layer is grown on the silicon wafer as the insulator. Second, a polysilicon layer is LPCVD deposited and patterned to form the heaters. (The self-heating technology will be



**Fig. 14** Operating principle of a capacitive humidity sensor

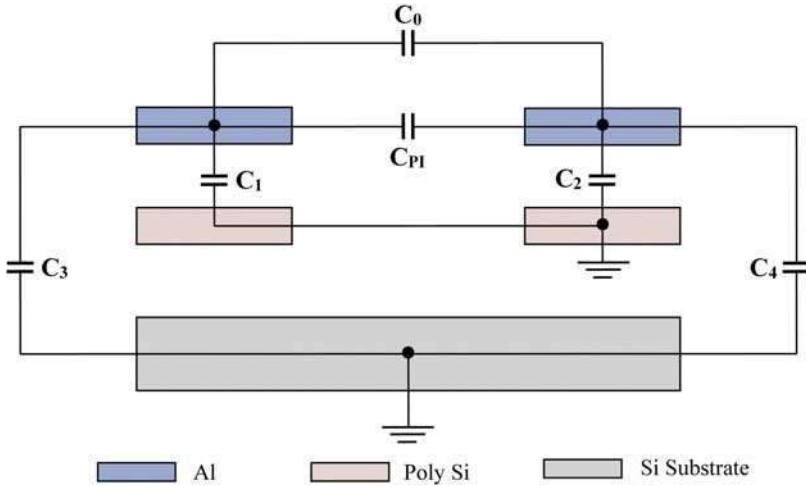
**Fig. 15** Schematic of an interdigitated capacitive humidity sensor



**Fig. 16** Fabrication process of the interdigitated capacitive humidity sensor

discussed in section “Self-Heating Technology.”) Third, another  $\text{SiO}_2$  layer is LPCVD deposited and patterned as the insulator. Fourth, an Al layer is sputtered and patterned to form the electrodes. Fifth, a passivation layer is deposited on the wafer and patterned to protect the electrodes. At last, a polyimide film is spin onto the wafer to form the sensitive layer.

The sensitive capacitance of an interdigitated device is generally small. As a result, stray capacitances or parasitic capacitances should be eliminated to improve



**Fig. 17** Equivalent circuit of the interdigitated humidity sensor

the measurement accuracy of the sensors. Figure 17 shows the equivalent circuit of the interdigitated humidity sensor fabricated by the process as illustrated in Fig. 16.  $C_{Pr1}$  is the sensitive capacitance with polyimide as the dielectric,  $C_0$  is the sensitive capacitance with air as the dielectric,  $C_1$  and  $C_2$  are the parasitic capacitances between the interdigitated electrodes and poly-silicon layer, and  $C_3$  and  $C_4$  are the parasitic capacitances between the interdigitated electrodes and the silicon substrate.

In some measurement instruments, there are three terminals. One of the terminals is ground. A three-terminal test method can be used to eliminate parasitic capacitances (Schroder 2006). The interdigitated electrodes are connected to the device terminals and other conductors nearby, such as the polysilicon heater and the substrate, should be connected to the ground terminal (the third terminal), so that parasitic capacitances will not contribute to the test results.

In a theoretical model the variation of the sensitive capacitance of an interdigitated structure can be given by:

$$\Delta C_s = (n - 1)n - 1 \cdot \epsilon_0 \cdot \Delta \epsilon_s \cdot L \left[ \frac{m \cdot K(k')}{2 \cdot K(k)} + \frac{h_1}{g} \right] \tag{14}$$

where  $n$  is the total number of electrode fingers,  $\epsilon_s$  is the relative dielectric constant of the sensitive material,  $L$  is the length of electrodes,  $h_1$  is the thickness of the sensitive layer,  $g$  is the gap between each finger,  $m$  is the coefficient base on electrode geometry,  $K(.)$  is the complete elliptic integral of the first kind,  $k$  is the modulus, and  $k'$  is the complementary modulus of  $k$ .

According to the model, the sensitivity of the interdigitated capacitive humidity sensors depends on the sizes of both the electrodes and the sensitive layer. On the one hand, finer and longer fingers means higher sensitivity, but it is limited by the

chip sizes and the fabrication accuracy. On the other hand, the sensitivity can be enhanced by a thick sensitive layer but it leads to a larger hysteresis and longer response time.

The sensitivity of the interdigitated capacitive humidity sensors are related to the sensitive material as well. In recent years, some two-dimensional materials were developed to work as the sensitive material. For example, graphene oxide which is an important derivative of graphene exhibits excellent humidity sensitive properties because of its oxygen functional groups. Zhao et al. compared the properties between graphene oxide and polyimide (Zhao et al. 2011b). With a same electrodes structure, the humidity sensitivity of the graphene oxide is one order higher than that of polyimide. As a two-dimensional material, water molecules can easily transfer into the material. The response time of a humidity sensor with graphene oxide as the sensitive material is about five times faster than that of a humidity sensor with polyimide as the sensitive material.

## Sandwich Structure

The sandwich structure shows much higher sensitivity than that of the interdigitated structure. The variation of the sensitive capacitance of a sandwich structure can be given by:

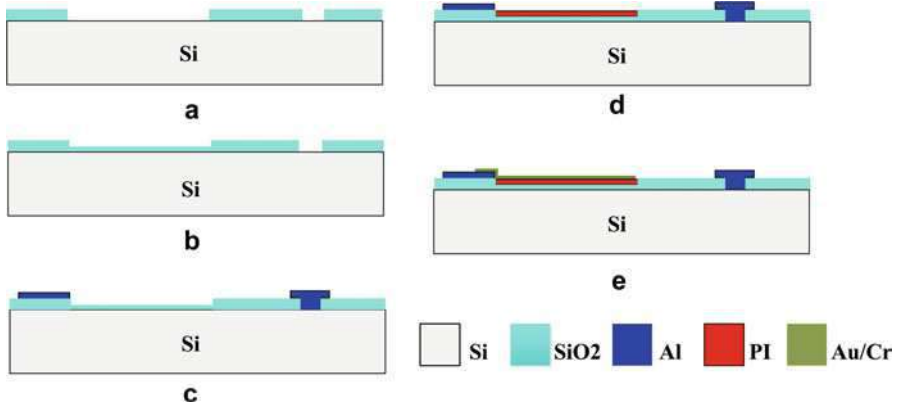
$$\Delta C = \frac{\varepsilon_0 \Delta \varepsilon_s A}{h} \quad (15)$$

where  $\varepsilon_s$  is the relative dielectric constant of the sensitive material,  $A$  is the area of the capacitor, and  $h$  is the thickness of the sensitive layer. It is notable that both the sensitive capacitance and the sensitivity can be enhanced by decreasing the thickness of the sensitive material, so a thinner sensitive layer is preferred. However, a thinner sensitive layer means a larger leakage current, especially in a high humidity environment. As a result, the thickness of the sensitive material should be carefully designed.

Figure 18 shows an example of the fabrication process of a sandwich capacitive humidity sensor. The fabrication starts from a heavily doped silicon wafer. First, a thick  $\text{SiO}_2$  layer is thermal grown on the wafer and patterned to define the sensitive area. Second, a thin  $\text{SiO}_2$  layer is thermal grown and patterned as the insulator layer. Third, an Al film is sputtering deposited and patterned as pads and lead wires. Fourth, a PI layer is spun onto the wafer and patterned as the sensitive material. At last, a top electrode is deposited and patterned.

In this structure, the silicon substrate is the bottom electrode of the sandwich capacitor and a multilayer dielectric instead a single sensitive layer is used. A  $\text{SiO}_2$  layer is inserted into the dielectric to reduce the leakage current. However, the  $\text{SiO}_2$  layer should be as thin as possible to avoid the degradation of the sensitivity. Moreover, outside the sensitive area, a thick  $\text{SiO}_2$  layer is applied to reduce the parasitic capacitance between lead wires and the substrate.





**Fig. 18** Fabrication process of a sandwich capacitive humidity sensor

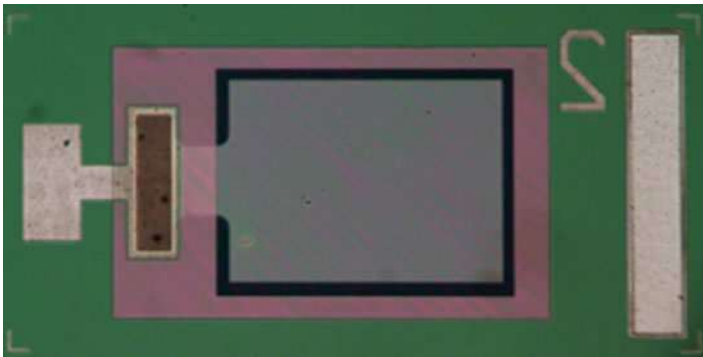
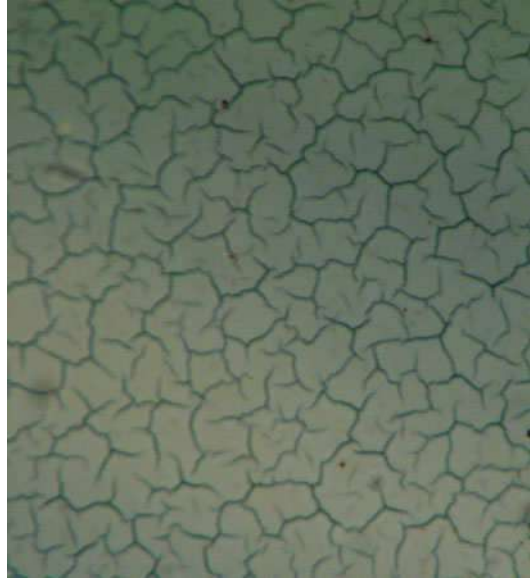
In a sandwich structure, the sensitive material is covered by a top electrode. A grid or a porous electrode should be applied to accelerate the diffusion of water molecules. There are several ways to fabricate the top electrode. In early researches the top electrode was patterned by a lithography process to form an array of windows (Kim et al. 2009b). Then some metal films with cracks were applied as the porous electrode (Fig. 19). After that a glancing angle deposition process was developed to form the porous electrode (Lee et al. 2013). Recently some nanomaterials such as nanowires or nanotubes were applied to improve the properties of the top electrode (Robbie and Brett 1997).

Figure 20 shows the photograph of a sandwich capacitive humidity sensor with a glancing angle deposited Au/Cr electrode. With the aid of a porous electrode and optimized dielectric thickness, the sensitive capacitance of this sensor is about 160 pF and the sensitivity is 0.187 pF/%RH, the response time of the sensor is 2.7 s.

## Sandwich-Interdigitated Structure

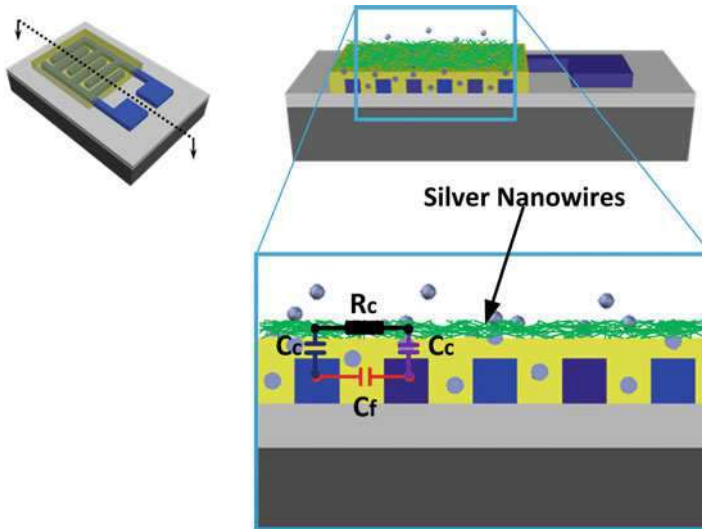
The sensitive capacitance and the sensitivity of an interdigitated capacitive structure are usually small (Rivadeneira et al. 2014). Although finer geometry of the electrodes means higher sensitivity, small finger sizes are limited by the lithography process. As a result, much effort has been made to improve the structure of capacitive interdigitated sensors. As shown in Fig. 21, a capacitive coupling method is applied to improve the performances of an interdigitated capacitive humidity sensor (Huang et al. 2016). The sensitive capacitor of this sensor is a sandwich-interdigitated structure.

A coupling electrode is employed in the improved structure to enhance the sensitive capacitance as well as the sensitivity. The coupling electrode is made of silver nanowire networks so that a grid structure is obtained to allow the penetration of water molecules. Unlike a conventional interdigitated structure, a thinner sensitive

**Fig. 19** Cr film with cracks**Fig. 20** Photograph of a sandwich capacitive humidity sensor with glancing angle deposited electrode

layer is preferred to increase both the coupling capacitance and the sensitivity of the sensor. Considering the limitation of the leakage current a  $0.1\text{--}0.2\ \mu\text{m}$  thick sensitive layer is used which is 10–20 times thinner than that in a conventional interdigitated structure. Although in the improved structure, an additional nanowire layer is deposited on the top, the total thickness of the nanowire layer and the sensitive layer is still thin enough to realize a fast response. Therefore, the hysteresis and the response/recovery time are reduced.

Compared to a conventional sandwich structure, the top electrode (nanowire networks) in the sandwich-interdigitated structure is not connected outside and it



**Fig. 21** Schematic of a sandwich-interdigitated capacitive humidity sensor

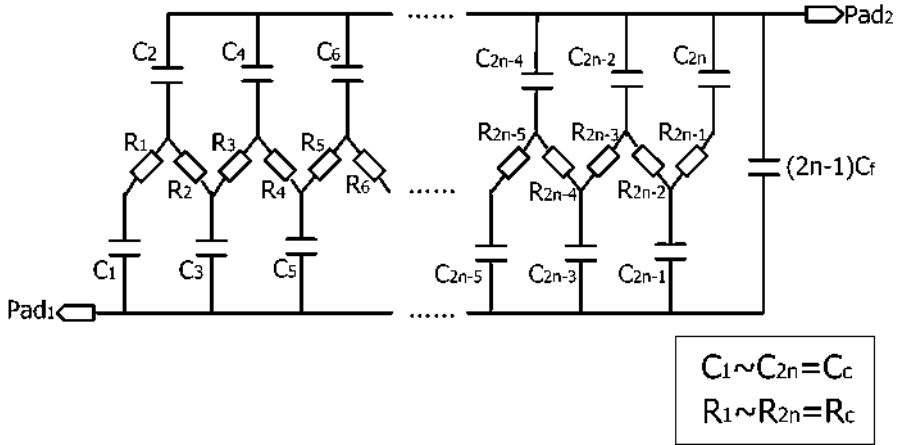
is just a floating electrode which is more robust, avoiding step coverage problems over different materials.

Figure 22 shows the equivalent circuit of the capacitive coupling structure. Where  $C_f$  is the capacitance between adjacent fingers,  $R_c$  is the parasitic resistance of the coupling electrode between adjacent fingers and  $C_c$  is the coupling capacitance between each finger and the coupling electrode. The parasitic capacitance between the interdigitated electrode and the substrate are not included in this model because it can be eliminated by connecting the substrate to the ground (Dean et al. 2005). With a thinner sensitive layer,  $C_c$  is usually much larger than  $C_f$  so that the sensitive capacitance is dominated by the coupling capacitance.

The fabrication process of the a sandwich-interdigitated structure is described in Fig. 23. The first few steps to fabricate the interdigitated electrodes are the same as that in Fig. 16. After that a thin humidity sensitive film is spun onto the wafer and patterned (Fig. 23b). At last, silver nanowires are dispersed onto the sensitive film to form the coupling electrode as shown in Fig. 23c. The scanning electron microscope (SEM) image of the humidity sensor before and after the post process is shown in Figs. 24 and 25.

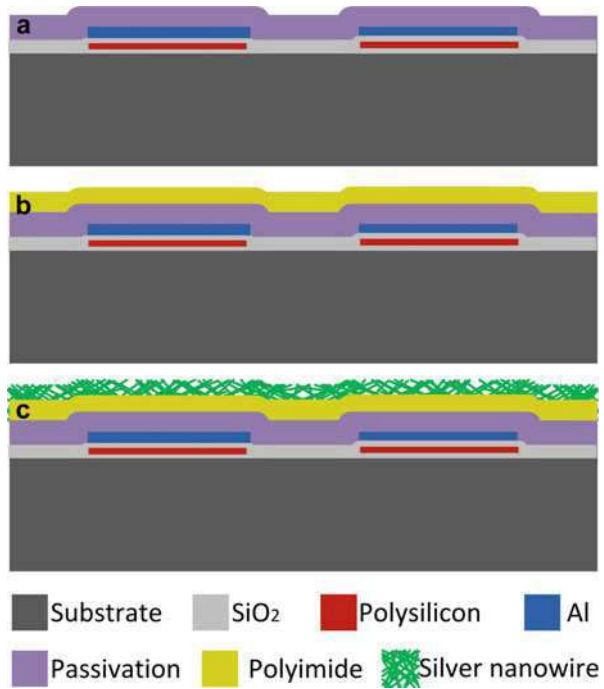
## Self-Heating Technology

Self-heating technology is widely used to improve the performances of humidity sensors (Dai et al. 2007; Zhao et al. 2011a). It is well known that heat will make water molecules move faster and then accelerate the diffusion of vapor.



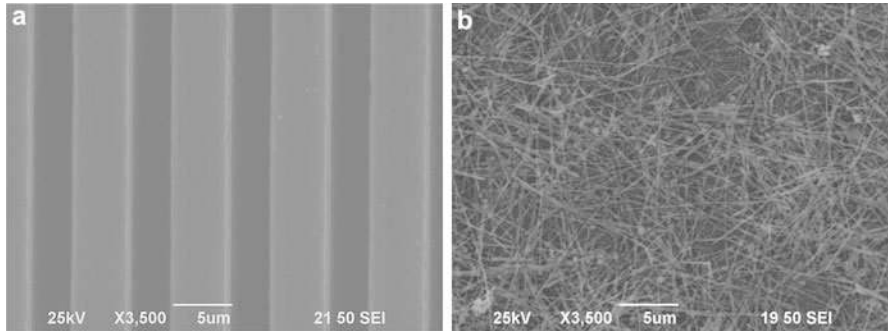
**Fig. 22** Equivalent circuit of the sandwich-interdigitated capacitive sensor

**Fig. 23** Fabrication process of the sandwich-interdigitated capacitive humidity sensor



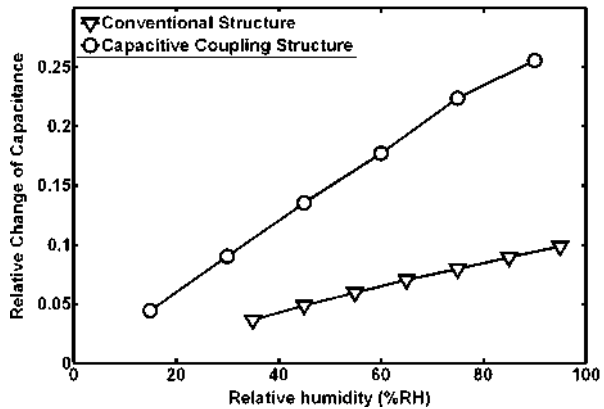
As shown in Fig. 26, a self-heating resistor is designed in a interdigitated capacitive humidity sensor (Zhao et al. 2011a).

As the voltage applied on the heater is 5 V, the heating power is 139 mW. The distribution of temperature on a heated chip can be tested by an infrared



**Fig. 24** SEM view of the sensor: (a) before post process; (b) after post process

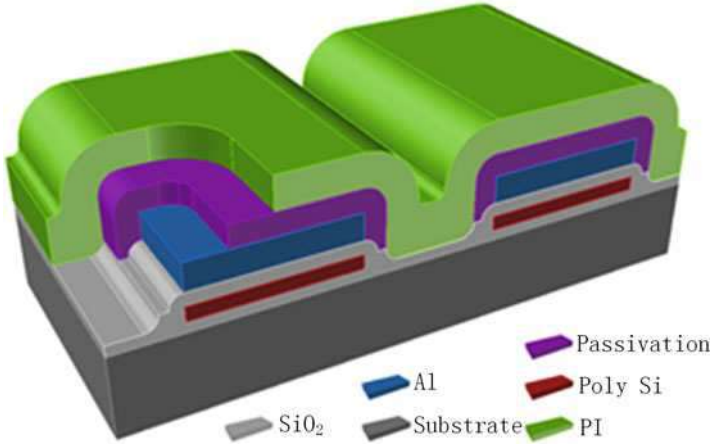
**Fig. 25** Comparison of the sensitivity between sandwich-interdigitated structure and conventional interdigitated structure



thermography. As shown in Fig. 27, the temperature distributes uniformly in the sensitive area and the variation of the temperature is about 30 °C with a heating voltage of 5 V.

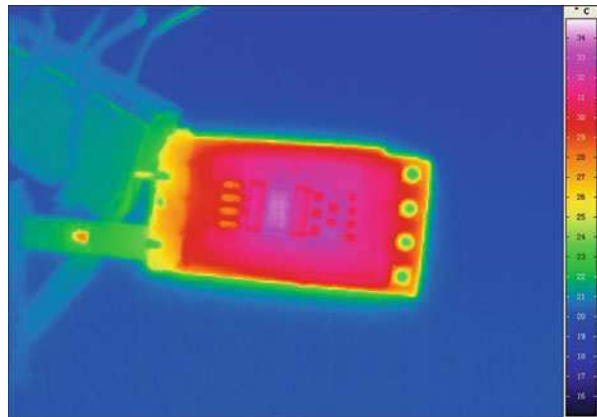
Figure 28 shows the heating process of the chip. The temperature at five points (P1–P5) on the chip is recorded. In the beginning, the temperature rises rapidly with the time. After 5 s it rises by about 10 °C. Then it rises gradually and tend to be stable.

The self-heating process improves both the hysteresis and response speed of the sensor. As reported by Zhao et al. (2011a), without heating, the hysteresis of the sensor is about 5%RH and after heating, the hysteresis of the sensor is less than 1% RH. The response time is defined as the time for the sensitive capacitance to rise to 90% of its final steady-state value and the recovery time is defined as the time for the sensitive capacitance to fall to 10% of its initial value. The response time of the sensor decreases from 38 s to 32 s and the recovery time decreases from 40 s to 25 s with the aid of heating.



**Fig. 26** Schematic of an interdigitated humidity sensor with a self-heating resistor

**Fig. 27** Temperature distribution on the humidity sensor with a heating voltage of 5 V

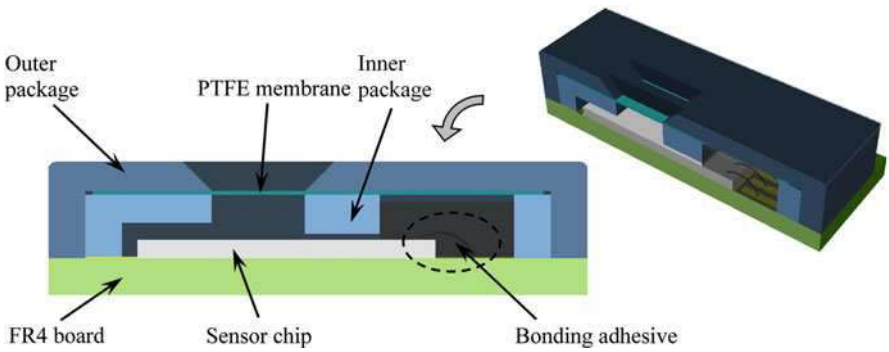
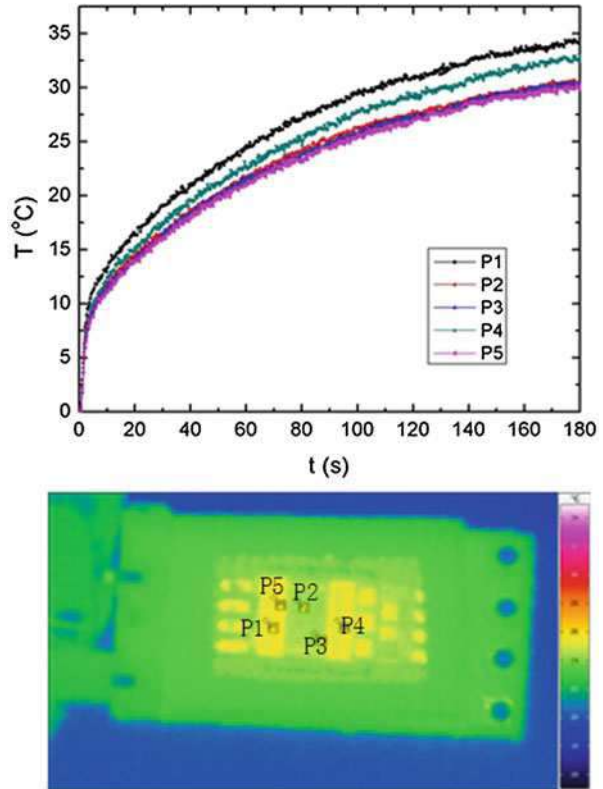


### Package of Humidity Sensors

In order to protect the devices, humidity sensors should be packaged. Unlike conventional integrated circuit (IC) packaging technology, package of sensors should meet the requirements of sensing applications.

There are several kinds of package for humidity sensors, including transistor outline (TO) package, single in-line package (SIP), leadless chip carrier (LCC) package, dual fat no-lead (DFN) package, flip chip package (Dean et al. 2005), and so on. Regardless of the type of package, for humidity sensing applications, breath windows should be designed to guarantee the interactions between the humidity sensors and the air outside.

**Fig. 28** Heating process of the humidity sensor

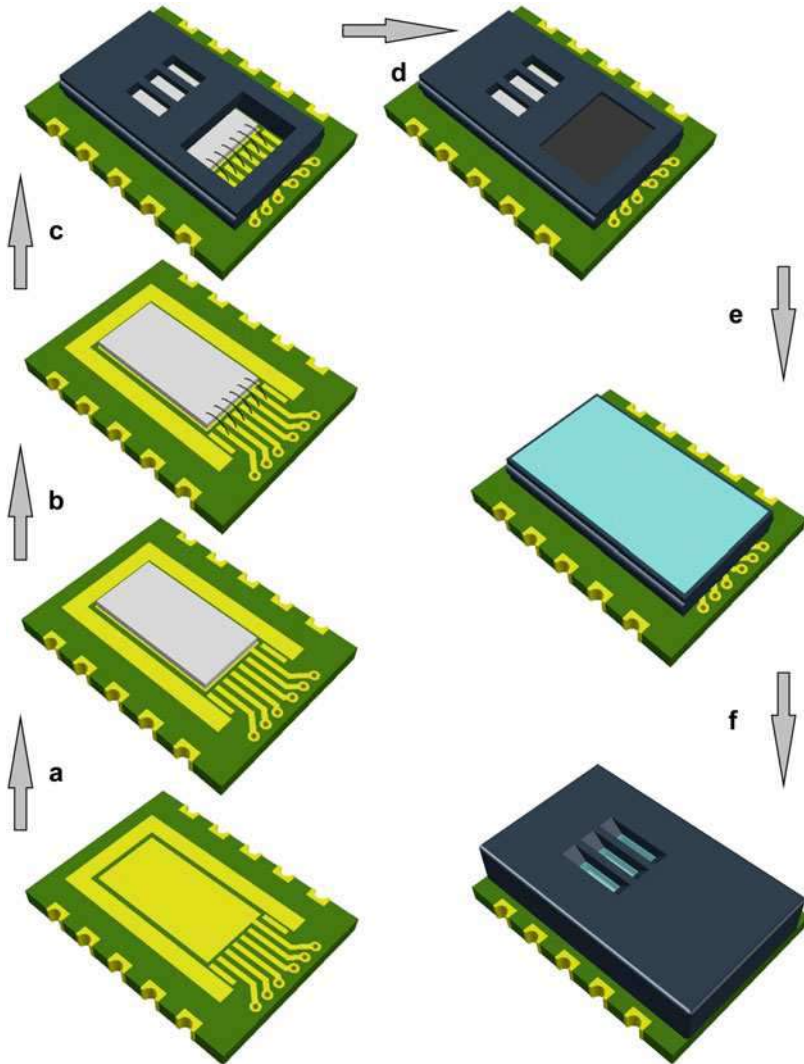


**Fig. 29** An LCC package for humidity sensor

As shown in Fig. 29, an LCC package is designed (Zhao et al. 2011a). The substrate of this package is a FR4 board. Both an inner package and an outer package are employed to protect the sensor inside and several breath windows are opened so that humidity can transfer into the package. An LCP material is applied to fabricate the

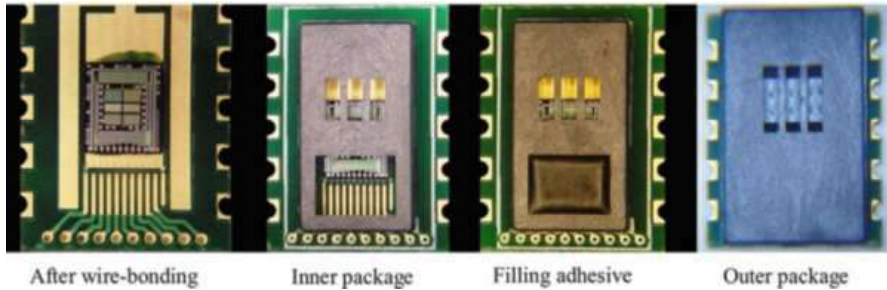
package because of its high temperature stability. A PTFE membrane is sandwiched between the inner and outer package. The PTFE membrane keeps the dust and water drops outside but water vapor can flow into the package.

The packaging process is illustrated in Fig. 30. After the fabrication of the sensor chip, the chip is attached onto the FR4 board by the adhesive material. Then Au wires are bonded between the chip and the substrate. Next, the inner package is adhered onto the substrate and the wires are protected by the bonding adhesive. After that, a PTFE membrane is fixed onto the inner package. At last, the outer package is mounted outside. Figure 31 shows the photographs of the packaged humidity sensor.



**Fig. 30** The package process of the humidity sensor





**Fig. 31** Photographs of the packaged humidity sensor

The properties of the humidity sensor is affected by the package. Although the package has little effect on the static properties, it increases the response time of the sensor. As reported by Zhao et al. (2011a) The sensitivity of the humidity sensor changes from 1.46 fF/%RH to 1.36 fF/%RH after package. However, the response time of the humidity sensor increases from 25 s to 38 s after package. To reduce the negative effect of the package, a self-heating technology can be used as described in the previous section.

---

## Conclusions

This chapter focuses on the basic technology of micromachined humidity sensors. Humidity, as an environment parameter, is defined in different ways. However, most of the micromachined humidity sensors are relative humidity sensors. The micromachined humidity sensors are normally clarified by sensitive materials or operating principle. The most popular micromachined humidity sensors including resistive humidity sensors, piezoresistive humidity sensors, and capacitive humidity sensors are described in detail. Beside the sensitive material and sensitive structure, various technologies are developed to improve the performances of the humidity sensor or facilitate their applications. In this chapter, self-heating technology and package technology are discussed. Because of the multiplicity of the micromachined humidity sensors, this chapter does not cover the nuts-and-bolts of the related technologies. More details can be found in references (Chen and Lu 2005; Farahani et al. 2014; Fenner and Zdankiewicz 2001; Rittersma 2002; Traversa 1995; Yamazoe and Shimizu 1986).

---

## References

- Allen JJ (2005) Micro electro mechanical system design. CRC Press, Boca Raton
- Bakhomou EG, Cheng MHM (2012) Miniature moisture sensor based on ultracapacitor technology. *IEEE Trans Compon Packag Manuf Technol* 2(7):1151–1157. <https://doi.org/10.1109/TCPMT.2012.2185496>
- Baltes H (1993) CMOS as sensor technology. *Sensors Actuators A Phys* 37-38(C):51–56. [https://doi.org/10.1016/0924-4247\(93\)80011-5](https://doi.org/10.1016/0924-4247(93)80011-5)

- Borini S, White R, Wei D, Astley M, Haque S, Spigone E, Harris N, Kivioja J, Ryhänen T (2013) Ultrafast Graphene oxide humidity sensors. *ACS Nano* 7(12):11166–11173
- Buchhold R, Nakladal A, Gerlach G, Neumann P (1998) Design studies on piezoresistive humidity sensors. *Sensors Actuators B Chem* 53(1–2):1–7. [https://doi.org/10.1016/S0925-4005\(98\)00297-4](https://doi.org/10.1016/S0925-4005(98)00297-4)
- Cai XY, Tao T, Qin M (2009) Study of the piezoresistive humidity sensor. *Chin J Electron Devices*
- Chen WH, Huang JQ, Zhu CY, Huang QA (2016) A black phosphorus humidity sensor with high sensitivity and fast response. Paper presented at the SENSORS, 2016 IEEE, Orlando
- Chen Z, Lu C (2005) Humidity sensors: a review of materials and mechanisms. *Sens Lett* 3(4):274–295. <https://doi.org/10.1166/sl.2005.045>
- Chen X, Rieth L, Miller MS, Solzbacher F (2010) Comparison of Y-doped Bazo3 thin films for high temperature humidity sensors by RF sputtering and pulsed laser deposition. *Sensors Actuators B Chem* 148(1):173–180. <https://doi.org/10.1016/j.snb.2010.03.074>
- Comparison of Relative Humidity Sensing Technologies (2004) [http://pasternack.ucdavis.edu/files/6213/7271/8210/hyd151\\_read13.pdf](http://pasternack.ucdavis.edu/files/6213/7271/8210/hyd151_read13.pdf). Accessed 6 July 2017
- Dai CL (2007) A capacitive humidity sensor integrated with micro heater and ring oscillator circuit fabricated by CMOS-MEMS technique. *Sensors Actuators B Chem* 122(2):375–380. <https://doi.org/10.1016/j.snb.2006.05.042>
- Dai CL, Liu MC, Chen FS, Wu CC, Chang MW (2007) A nanowire WO3 humidity sensor integrated with micro-heater and inverting amplifier circuit on chip manufactured using CMOS-MEMS technique. *Sensors Actuators B Chem* 123(2):896–901. <https://doi.org/10.1016/j.snb.2006.10.055>
- Dean M, Douglas RD (1962) Semiconductor and conventional strain gages. Academic, New York
- Dean R, Pack J, Sanders N, Reiner P (2005) Micromachined LCP for packaging MEMS sensors. In: *IECON proceedings (industrial electronics conference)*, Raleigh, pp 2363–2367. <https://doi.org/10.1109/IECON.2005.1569273>
- Ducéré V, Bernès A, Lacabanne C (2005) A capacitive humidity sensor using cross-linked cellulose acetate butyrate. *Sensors Actuators B Chem* 106(1 Special issue):331–334. <https://doi.org/10.1016/j.snb.2004.08.028>
- E+E ELEKTRONIK (n.d.) Measure humidity – basics. [http://downloads.epluse.com/fileadmin/data/miscellaneous/principles\\_of\\_humidity\\_measurement.pdf](http://downloads.epluse.com/fileadmin/data/miscellaneous/principles_of_humidity_measurement.pdf). Accessed 6 July 2017
- Eder C, Valente V, Donaldson N, Demosthenous A (2014) A CMOS smart temperature and humidity sensor with combined readout. *Sensors (Switzerland)* 14(9):17192–17211. <https://doi.org/10.3390/s140917192>
- Erol A, Okur S, Yağmurcakardeş N, Arıkan MÇ (2011) Humidity-sensing properties of a ZnO nanowire film as measured with a QCM. *Sensors Actuators B Chem* 152(1):115–120. <https://doi.org/10.1016/j.snb.2010.09.005>
- Farahani H, Wagiran R, Hamidon MN (2014) Humidity sensors principle, mechanism, and fabrication technologies: a comprehensive review, vol 14. <https://doi.org/10.3390/s140507881>
- Fenner R, Zdankiewicz E (2001) Micromachined water vapor sensors: a review of sensing technologies. *IEEE Sensors J* 1(4):309–317. <https://doi.org/10.1109/7361.983470>
- Gu L, Huang Q-A, Qin M (2004) A novel capacitive-type humidity sensor using CMOS fabrication technology. *Sensors Actuators B Chem* 99(2–3):491–498. <https://doi.org/10.1016/j.snb.2003.12.060>
- Gu L, Zheng K, Zhou Y, Li J, Mo X, Patzke GR, Chen G (2011) Humidity sensors based on ZnO/TiO2 core/shell nanorod arrays with enhanced sensitivity. *Sensors Actuators B Chem* 159(1):1–7. <https://doi.org/10.1016/j.snb.2010.12.024>
- He Y, Zhang T, Zheng W, Wang R, Liu X, Xia Y, Zhao J (2010) Humidity sensing properties of BaTiO3 nanofiber prepared via electrospinning. *Sensors Actuators B Chem* 146(1):98–102. <https://doi.org/10.1016/j.snb.2010.02.030>
- Huang JQ, Li B, Chen W (2016) A CMOS MEMS humidity sensor enhanced by a capacitive coupling structure. *Micromachines* 7(5). <https://doi.org/10.3390/mi7050074>

- Huang JQ, Li F, Zhao M, Wang K (2015) A surface micromachined CMOS MEMS humidity sensor. *Micromachines* 6(10):1569–1576. <https://doi.org/10.3390/mi6101440>
- Islam T, Nimal AT, Mittal U, Sharma MU (2015) A micro interdigitated thin film metal oxide capacitive sensor for measuring moisture in the range of 175–625 ppm. *Sensors Actuators B Chem* 221:357–364. <https://doi.org/10.1016/j.snb.2015.06.101>
- Jackson T, Mansfield K, Saafi M, Colman T, Romine P (2008) Measuring soil temperature and moisture using wireless MEMS sensors. *Measurement* 41(4):381–390. <https://doi.org/10.1016/j.measurement.2007.02.009>
- Kim Y, Jung B, Lee H, Kim H, Lee K, Park H (2009a) Capacitive humidity sensor design based on anodic aluminum oxide. *Sensors Actuators B Chem* 141(2):441–446. <https://doi.org/10.1016/j.snb.2009.07.007>
- Kim JS, Lee MJ, Kang MS, Yoo KP, Kwon KH, Singh VR, Min NK (2009b) Fabrication of high-speed polyimide-based humidity sensor using anisotropic and isotropic etching with ICP. *Thin Solid Films* 517(14):3879–3882. <https://doi.org/10.1016/j.tsf.2009.01.129>
- Lai JR, Hidy GM (1968) Microsensor for measuring humidity. *Rev Sci Instrum* 39(8):1197–1203. <https://doi.org/10.1063/1.1683616>
- Lee MJ, Hong HP, Kwon KH, Park CW, Min NK (2013) Fast-speed, high-sensitivity polyimide humidity sensors with superhydrophilic carbon nanotube network electrodes. *Sensors Actuators B Chem* 185:97–104. <https://doi.org/10.1016/j.snb.2013.04.117>
- Lee C-Y, Lee G-B (2005) Humidity sensors: a review. *Sens Lett* 3(1):1–15. <https://doi.org/10.1166/sl.2005.001>
- Lee H, Lee S, Jung S, Lee J (2011) Nano-grass polyimide-based humidity sensors. *Sensors Actuators B Chem* 154(1):2–8. <https://doi.org/10.1016/j.snb.2009.11.054>
- Lee CW, Park HS, Kim JG, Choi BK, Joo SW, Gong MS (2005) Polymeric humidity sensor using organic/inorganic hybrid polyelectrolytes. *Sensors Actuators B Chem* 109(2):315–322. <https://doi.org/10.1016/j.snb.2004.12.063>
- Molina-Lopez F, Briand D, De Rooij NF (2013) Decreasing the size of printed comb electrodes by the introduction of a dielectric interlayer for capacitive gas sensors on polymeric foil: modeling and fabrication. *Sensors Actuators B Chem* 189:89–96. <https://doi.org/10.1016/j.snb.2013.01.008>
- Nizhnik O, Higuchi K, Maenaka K (2012) Self-calibrated humidity sensor in CMOS without post-processing. *Sensors* 12(1):226–232. <https://doi.org/10.3390/s120100226>
- Nohria R, Khillan RK, Su Y, Dikshit R, Lvov Y, Varshramyan K (2006) Humidity sensor based on ultrathin polyaniline film deposited using layer-by-layer nano-assembly. *Sensors Actuators B Chem* 114(1):218–222. <https://doi.org/10.1016/j.snb.2005.04.034>
- Vaisala Oyj (2013) Calculation formulas for humidity. [http://www.vaisala.fi/Vaisala%20Documents/Application%20notes/Humidity\\_Conversion\\_Formulas\\_B210973EN-F.pdf](http://www.vaisala.fi/Vaisala%20Documents/Application%20notes/Humidity_Conversion_Formulas_B210973EN-F.pdf). Accessed 6 July 2017
- Vaisala Oyj (n.d.) Humidity theory, terms & definitions. [http://www.vaisala.com/Vaisala%20Documents/Common%20Files/Humidity101\\_HumidityTheoryTermsDefinitions.pdf?\\_ga=1.214434967.1482519469.1446505017](http://www.vaisala.com/Vaisala%20Documents/Common%20Files/Humidity101_HumidityTheoryTermsDefinitions.pdf?_ga=1.214434967.1482519469.1446505017). Accessed 6 July 2017
- Packirisamy M, Stiharu I, Li X, Rinaldi G (2005) A polyimide based resistive humidity sensor. *Sens Rev* 25(4):271–276. <https://doi.org/10.1108/02602280510620123>
- Rittersma ZM (2002) Recent achievements in miniaturised humidity sensors – a review of transduction techniques. *Sensors Actuators A Phys* 96(2–3):196–210. [https://doi.org/10.1016/S0924-4247\(01\)00788-9](https://doi.org/10.1016/S0924-4247(01)00788-9)
- Rivadeneira A, Fernández-Salmerón J, Agudo M, López-Villanueva JA, Capitan-Vallvey LF, Palma AJ (2014) Design and characterization of a low thermal drift capacitive humidity sensor by inkjet-printing. *Sensors Actuators B Chem* 195:123–131. <https://doi.org/10.1016/j.snb.2013.12.117>
- Robbie K, Brett MJ (1997) Sculptured thin films and glancing angle deposition: growth mechanics and applications. *J Vac Sci Technol A* 15(3):1460–1465
- Schroder DK (2006) *Semiconductor material and device characterization*. Wiley, Hoboken

- Schubert PJ, Nevin JH (1985) A polyimide-based capacitive humidity sensor. *IEEE Trans Electron Devices* 32(7):1220–1223. <https://doi.org/10.1109/T-ED.1985.22104>
- Smith CS (1954) Piezoresistance effect in germanium and silicon. *Phys Rev* 94(1):42–49. <https://doi.org/10.1103/PhysRev.94.42>
- Steffens C, Manzoli A, Leite FL, Fatibello O, Herrmann PSP (2014) Atomic force microscope microcantilevers used as sensors for monitoring humidity. *Microelectron Eng* 113:80–85. <https://doi.org/10.1016/j.mee.2013.07.015>
- Su PG, Ho CJ, Sun YL, Chen IC (2006) A micromachined resistive-type humidity sensor with a composite material as sensitive film. *Sensors Actuators B Chem* 113(2):837–842. <https://doi.org/10.1016/j.snb.2005.03.109>
- Tang QY, Chan YC, Zhang K (2011) Fast response resistive humidity sensitivity of polyimide/multiwall carbon nanotube composite films. *Sensors Actuators B Chem* 152(1):99–106. <https://doi.org/10.1016/j.snb.2010.09.016>
- Thomson W (1856) On the electro-dynamic qualities of metals: – effects of magnetization on the electric conductivity of nickel and of iron. In: *Proceedings of the Royal Society of London*. JSTOR, pp 546–550
- Tomlinson H (1883) The influence of stress and strain on the action of physical forces. *Philos Trans R Soc Lond* 174:1–172
- Traversa E (1995) Ceramic sensors for humidity detection: the state-of-the-art and future developments. *Sensors Actuators B Chem* 23(2–3):135–156. [https://doi.org/10.1016/0925-4005\(94\)01268-M](https://doi.org/10.1016/0925-4005(94)01268-M)
- Yamazoe N, Shimizu Y (1986) Humidity sensors: principles and applications. *Sensors Actuators* 10(3–4):379–398. [https://doi.org/10.1016/0250-6874\(86\)80055-5](https://doi.org/10.1016/0250-6874(86)80055-5)
- Zhang D, Chang H, Li P, Liu R, Xue Q (2016) Fabrication and characterization of an ultrasensitive humidity sensor based on metal oxide/graphene hybrid nanocomposite. *Sensors Actuators B Chem* 225:233–240. <https://doi.org/10.1016/j.snb.2015.11.024>
- Zhang T, He Y, Wang R, Geng W, Wang L, Niu L, Li X (2008) Analysis of dc and ac properties of humidity sensor based on polypyrrole materials. *Sensors Actuators B Chem* 131(2):687–691. <https://doi.org/10.1016/j.snb.2007.12.059>
- Zhang M, Hu C, Liu H, Xiong Y, Zhang Z (2009) A rapid-response humidity sensor based on BaNbO<sub>3</sub> nanocrystals. *Sensors Actuators B Chem* 136(1):128–132. <https://doi.org/10.1016/j.snb.2008.09.021>
- Zhao CL, Qin M, Huang QA (2011a) A fully packaged CMOS interdigital capacitive humidity sensor with polysilicon heaters. *IEEE Sensors J* 11(11):2986–2992. <https://doi.org/10.1109/JSEN.2011.2154325>
- Zhao C-L, Qin M, Li W-H, Huang Q-A (2011b) Enhanced performance of a CMOS interdigital capacitive humidity sensor by graphene oxide. Paper presented at the solid-state sensors, actuators and microsystems conference (TRANSDUCERS), 2011 16th international, Beijing, 5–9 June 2011
- Zhu C, Xu F, Zhang L, Li M, Chen J, Xu S, Huang G, Chen W, Sun L (2016) Ultrafast preparation of black phosphorus quantum dots for efficient humidity sensing. *Chem Eur J* 22(22):7357–7362. <https://doi.org/10.1002/chem.201600719>



# Piezoelectric Micro/Nano Mechanical Devices for Frequency Control and Chemical Sensing

Wei Pang, Menglun Zhang, and Ji Liang

## Contents

Introduction .....	818
Frequency Control Applications .....	818
BAW Devices .....	818
BAW Filters .....	820
BAW Oscillators .....	821
Lamb Wave Resonator .....	823
Piezoelectric Micro-/Nano-Mechanical Sensors .....	831
Sensing Mechanism .....	831
Biological Macromolecule Sensing .....	833
Multimode Sensing System .....	834
Piezoelectric Lab on a Chip .....	836
Sensor Signal Enhancement Strategy in Liquid .....	837
VOC Gas Sensing .....	838
Impact of Nanotechnology .....	842
References .....	843

## Abstract

This chapter reports on the state of the art of piezoelectric micro-/nano-mechanical devices in frequency control and sensing applications. Recent studies on bulk acoustic wave (BAW) devices are introduced, including investigation of high-coupling materials and filter and oscillator designs. A novel class of frequency devices based on Lamb waves is also reviewed. Micro- and nano-mechanical sensors for various sensing applications and integrated module are outlined.

---

W. Pang (✉) · M. Zhang · J. Liang  
State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University,  
Tianjin, China  
e-mail: [weipang@tju.edu.cn](mailto:weipang@tju.edu.cn)

---

**Keywords**

Bulk acoustic wave · FBAR · SMR · LWR · Piezoelectric micro-mechanical sensor

---

**Introduction**

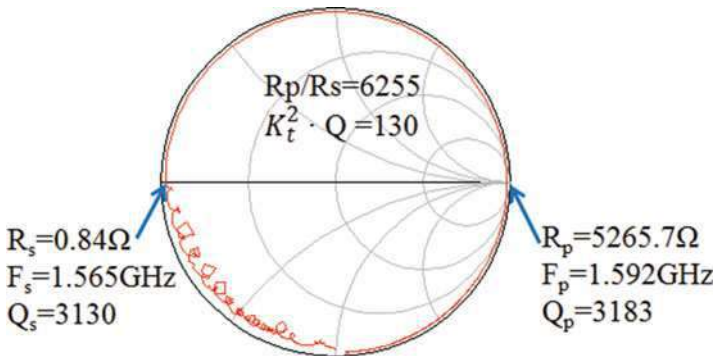
Piezoelectric micro- and nano-mechanical devices are one of the “shining stars” in academic fields and the industrial world. Starting with quartz resonators, these devices have evolved from generation to generation, and up to now, together with the surface acoustic wave (SAW) devices and thin-film piezoelectric devices, they still dominate timing and filter circuits in most communication systems. The number of filters in a cell phone has escalated from less than ten a decade ago to more than 60 today and is predicted to rise to over 100 in the future (Warder and Link 2015). This explosion in the filter number is attributed to the increasing demand for larger amounts of data transmission and more rapid data delivery. With an increasing number of devices connecting to the cloud and the emergence of carrier aggregation (CA) and multiple-input/multiple-output (MIMO) technologies, the world capacity for mobile filters is soon expected to exceed 30 billion. Therefore, piezoelectric micro-devices will continue to play a significant role in the future.

The Internet of Things (IoT) is the networked interconnection of objects from the sophisticated to the material world that relies on a trillion or more sensors as data inputs of our environment. Micro- and nanoscale sensors based on the piezoelectric transduction principle are well-known for their high-resolution mass sensitivity, and they are expected to play a paramount role in successful development of the IoT in the future. These devices can be massively produced with ultra-small size, low power consumption, and high sensitivity. Such features present opportunities for novel label-free detectors and notably high levels of multiplexing and throughput that are not easily matched by other transduction methods, such as quartz crystal microbalance (QCM) or SPR. The size and design flexibility of these devices also offer possibilities for incorporating microfluidics and miniaturized lab-on-a-chip formats. With high mass resolutions, piezoelectric micro- and nano-mechanical sensors have been studied extensively for many applications, such as DNA hybridization, immune-sensor development, and gas detection, among others.

---

**Frequency Control Applications****BAW Devices**

The rapid growth in smartphone adoption has resulted in a boost for filters or duplexers. By virtue of an explosion of the fourth-generation long-term evolution of wireless technology (4G) and a keen desire for the forthcoming fifth-generation (5G), the number of filters required in smartphones is skyrocketing. Moreover, with an increasing need for the Wi-Fi band and other newly created bands, the golden age



**Fig. 1** Measured electrical response of a high-performance FBAR with a high value of  $R_p/R_s$

for filter market is approaching (Warder and Link 2015). Because the available spectrum for communication is overcrowded, the upcoming bands must take over higher frequencies. At frequencies near 2 GHz and higher, BAW filters are superior to SAW devices, and thus they are expected to become the mainstream technology in future filter designs. In addition, by featuring high performance at high frequencies, these devices are useful for radio-frequency (RF) oscillator designs.

In general, BAW devices are classified into two categories. One category uses a sacrificed air cavity as the acoustic mirror, known as the film bulk acoustic wave resonator (FBAR). The other category uses alternating layers of differentiated acoustic velocity as acoustic reflectors, i.e., the solidly mounted resonator (SMR). Each topology has its own advantages in filter design. This section covers both types of resonators.

### Toward a Higher Coupling Coefficient

The performance of resonators is assessed using the ratio of  $R_p$  (resistance at parallel resonance) to  $R_s$  (resistance at series resonance). Filters consisting of resonators with large  $R_p/R_s$  tend to possess low insertion loss (IL). Using a surface micromachining AlN MEMS process, the state-of-the-art technology can produce an FBAR with a remarkably high  $R_p/R_s$  value of over 6000 (Liang et al. 2014), as shown in Fig. 1. The coupling coefficient ( $k_t^2$ ) is another key parameter for resonators because it dominates the bandwidth of a filter. C-axis wurtzite polycrystalline aluminum nitride is the commercial material used in BAWs, and devices based on this substance yielded an average  $k_t^2$  of 6–7%. However, with the challenging need for wider bandwidth, piezoelectric materials with a higher coupling coefficient are desired.

Theoretically, single-crystal AlN or undoped single-crystal AlGaN has higher longitudinal piezoelectric coupling coefficients than polycrystalline films, and therefore resonators based on these materials have superior mechanical coupling efficiency. Researchers from Akoustis Technologies demonstrated a single-crystal AlN FBAR with a  $k_t^2$  of 14% that greatly outperforms the current polycrystalline device (Shealy et al. 2016b) and they also investigated FBARs based on single-crystal



AlGaN piezoelectric films using metal-organic chemical vapor deposition (MOCVD) (Shealy et al. 2016a). Although the device demonstrated a  $k_r^2$  of only 4.44%, it opened the door for fabrication of high-performance AlGaN FBARs.

Doping of other materials in AlN also contributes to coupling enhancement. The literature (Akiyama et al. 2009) has reported  $\text{Sc}_{0.43}\text{Al}_{0.57}\text{N}$  alloy thin films prepared by dual reactive co-sputtering. The doped film exhibits a tremendously large piezoelectric coefficient  $d_{33}$  of  $27.6 \text{ pCn}^{-1}$ , which is at least 500% larger than that of AlN. The intrinsic effect is verified by quantum mechanical calculations (Tasnadi et al. 2010), which enables high-coupling-resonator fabrication. Moreira et al. studied the performance of FBAR incorporating c-textured  $\text{Al}_{(1-x)}\text{Sc}_x\text{N}$  thin films with Sc concentrations ranging from 0 to 0.15 (Moreira et al. 2011). Within this range concentration, the measured  $k_r^2$  is found to increase by as much as 100%, and a 2.15 GHz FBAR using an  $\text{Al}_{0.85}\text{Sc}_{0.15}\text{N}$  film with a  $k_r^2$  of 12% was presented. Matloub et al. reported a 2.5 GHz FBAR using  $\text{Al}_{0.88}\text{Sc}_{0.12}\text{N}$  films, which yields a  $k_r^2$  of 7.3% and a quality factor of 650 (Matloub et al. 2011). Umedal et al. developed BAW devices using  $\text{Al}_{0.65}\text{Sc}_{0.35}\text{N}$  as piezoelectric films, demonstrating a  $k_r^2$  as high as 15.5% (Umeda et al. 2013).

Moreover, co-doping of AlN using Mg-Zr or Mg-Hf was also found to be effective. Researchers from Taiyo Yuden demonstrated FBARs based on  $(\text{Mg}, \text{Zr})_{0.13}\text{Al}_{0.87}\text{N}$  and  $(\text{Mg}, \text{Hf})_{0.13}\text{Al}_{0.87}\text{N}$  films with  $k_r^2$  values of 8.5% and 10.0%, respectively (Yokoyama et al. 2015).

## BAW Filters

With the rapid development of wireless communication, the available spectrum is becoming overcrowded, and therefore, new frequency bands are continuously created for data transmission. Filter designs based on BAW technologies must keep pace with these emerging bands. Zhang et al. reported a RF bandpass filter in the C-band based on FBAR with an insertion loss of less than 2.0 dB at 5.2 GHz and an attenuation greater than 50 dB (Zhang et al. 2013). The filter is designed with a conventional ladder configuration. To realize high rejection, a high capacitance value of the series branch should be guaranteed. Instead of a large-area resonator, two resonators with the same size are used in the series branch in case of spurious modes. An additional capacitor is also added between two shunt resonators to tune the zeroes of the filter. Qi et al. presented a high-performance TDSCDMA FBAR bandpass filter with ladder configuration (Qi et al. 2012). The filter demonstrated a low passband insertion loss of 1.7 dB and high stopband rejection greater than 35 dB. Researchers from Qorvo developed a full band 41 filter with high Wi-Fi rejection using BAW technology (Kreuzer et al. 2015). Band 41 is one of the most important LTE bands in various areas, and this design fulfilled the spectrum demand. Additionally, this band shows excellent rejection in the close spectrum for Wi-Fi, which perfectly meets the strict coexistence requirements.

Currently, AlN-based BAW filters have satisfied most of the commercial needs of wireless communication systems. However, the special demand for filters with large



bandwidth, i.e., up to 10%, still poses a challenge to the conventional AlN-based BAW filters because the available bandwidth of the BAW filter is usually limited by the coupling coefficient. Two corresponding approaches are used to widen the bandwidth of the BAW filter. One approach uses an alternative piezoelectric material with an inherent coupling coefficient higher than that of AlN, as mentioned previously. The other approach is development of a novel filter topology. In the literature (Yang et al. 2016), a modified lattice configuration is reported to achieve a wideband filter response using AlN-based BAW resonators. This novel topology differs from the conventional one in that two auxiliary inductors have been added parallel to each other to both ends of the balanced inlet and outlet. The advantage of this improved topology is apparent when it is adopted in multistage configuration because the two parallel inductors at the facing ports can be combined into one equivalent inductor, which enables a more compact integration of the filter chip. The prototype filter was measured with a relative bandwidth of 12.3% and a central frequency at 3.25 GHz, showing superiority to the typical ladder and lattice topologies. Figure 2 presents a photo of the filter and its electrical performance.

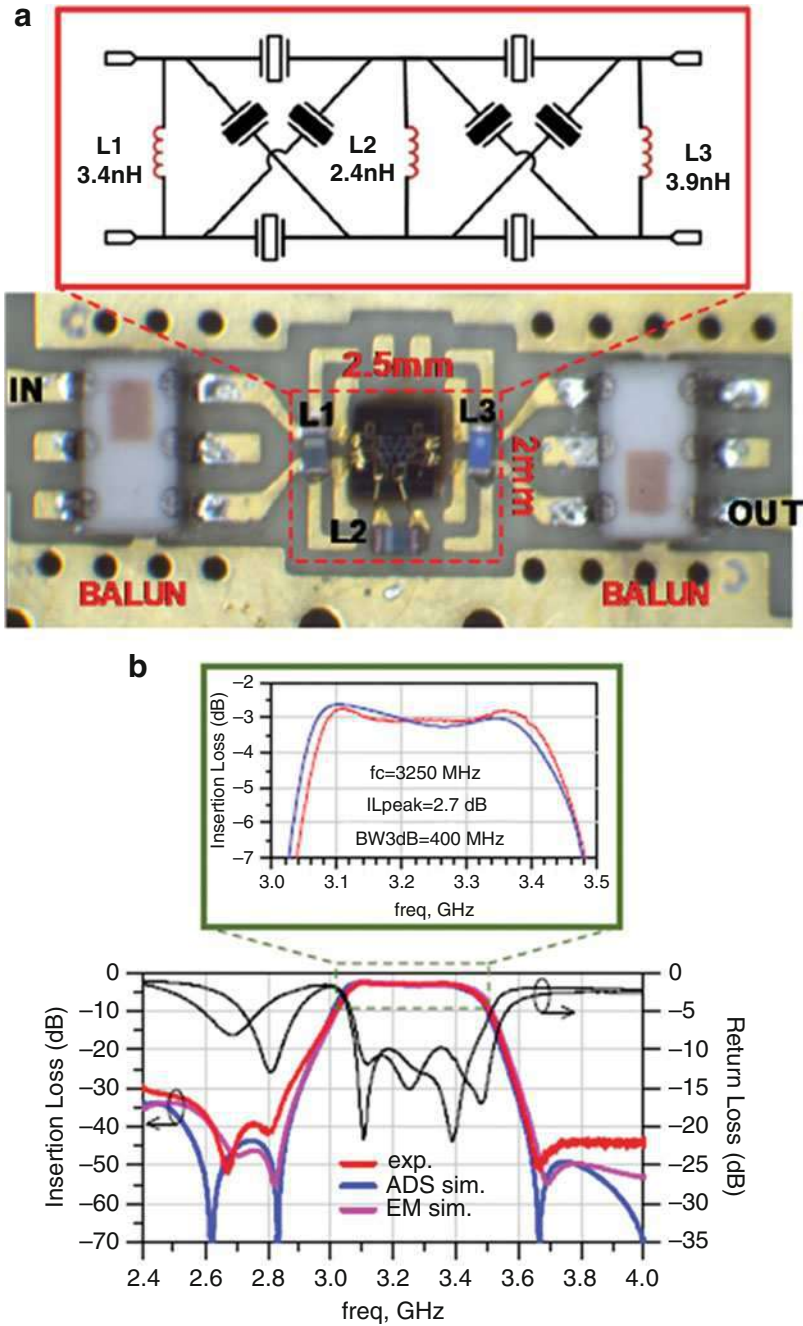
Other meaningful work includes multi-frequency filters and differential filters design. Sun et al. reported a dual-passband filter composed of two bandpass FBAR filters with shared input and output ports (Sun et al. 2015). The measured insertion loss is approximately 2.0 dB for both the lower and the upper passbands, and the filter rejections outside the passbands are greater than 30 dB.

Zhang et al. presented a new architecture of the BAW filter with single ended to balanced mode conversion (Zhang and Pang 2011). As shown in Fig. 3a, the single ended to balanced BAW filter includes a single-ended ladder section at the filter input, a fully balanced lattice section at the output, and a balun in the middle to couple the two sections electrically. By taking full advantage of the characteristics of the ladder and lattice structures, a single to balanced filter with low insertion loss, steep near-band skirts, and deep out-of-band attenuations can be achieved. A photo of the filter and its electrical performance are shown in Fig. 3b, c, respectively.

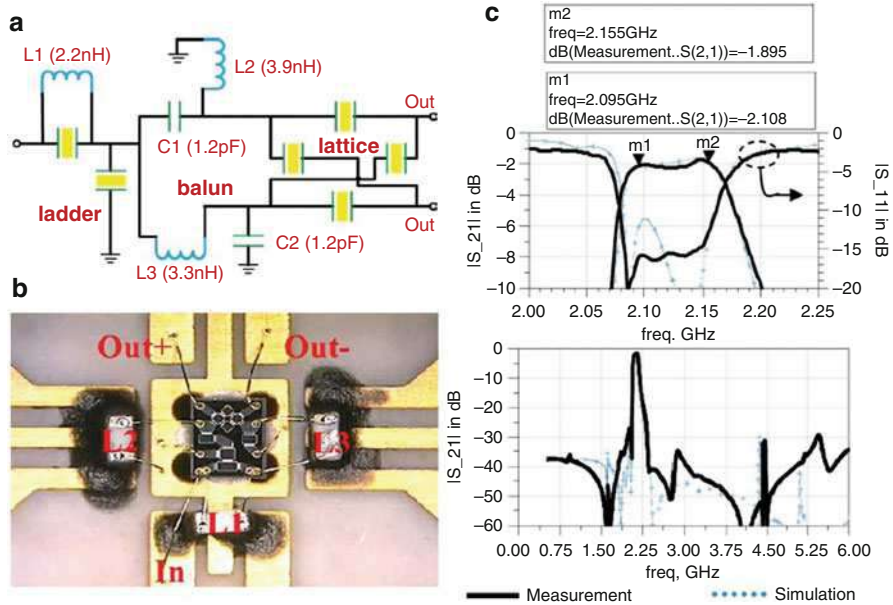
## BAW Oscillators

BAW oscillators feature high-frequency and great complementary metal-oxide-semiconductor (CMOS) compatibility and are thus suitable for application in compact circuits. Recent research on BAW oscillators has primarily focused on temperature compensation and low power consumption.

Pang et al. proposed a passively temperature-compensated FBAR oscillator (Pang et al. 2008). The FBAR incorporates a compensation material that has a positive temperature coefficient of Young's modulus. The 604-MHz oscillator consumes 5.3 mW from a 3.3 V supply and achieves excellent phase noise performances of  $-102$ ,  $-130$ , and  $-149$  dBc/Hz at 1, 10, and 100 kHz carrier offsets, respectively. The oscillator's temperature-dependent frequency drift is less than 80 ppm over a temperature range of  $-35$  °C to  $+85$  °C, which is well suited for high-speed serial



**Fig. 2** (a) Design of the topology and board assembly of the filter with the proposed architecture. (b) The frequency response of the fabricated filter versus simulation results



**Fig. 3** (a) Illustration of the filter topology. (b) Completed single to balanced BAW filter comprising the chip and auxiliary SMT inductors. (c) The test result of the passband and wide band transmission parameters of the single to balanced filter

data applications. The literature (Rai et al. 2010) demonstrated a temperature-compensated 1.5 GHz CMOS/FBAR oscillator with less than  $\pm 10$  ppm drift over 100 °C temperature range and low power consumption (515  $\mu$ A with a 2 V supply).

## Lamb Wave Resonator

### Classification

SAW and BAW technologies have dominated the wireless market for decades, but they have drawbacks for unique applications. For example, the intermediate frequency (IF) bandpass filters (BPF) that are widely adopted in radio-frequency (RF) front ends are highly important. SAW filters are applied to cover this spectrum, but they suffer from bulky volumes and incompatibility for microelectronic integration. Although BAW filters are highly successful for RF MEMS devices, the products made using the current manufacturing process are less reliable for IF band applications due to the impediment posed by the film thickness requirement, which dominates the resonance frequency and is usually on the order of tens of microns. Over the past decade, the Lamb wave resonator, which was investigated as an acoustic sensor years ago, has revived interest in frequency control applications by

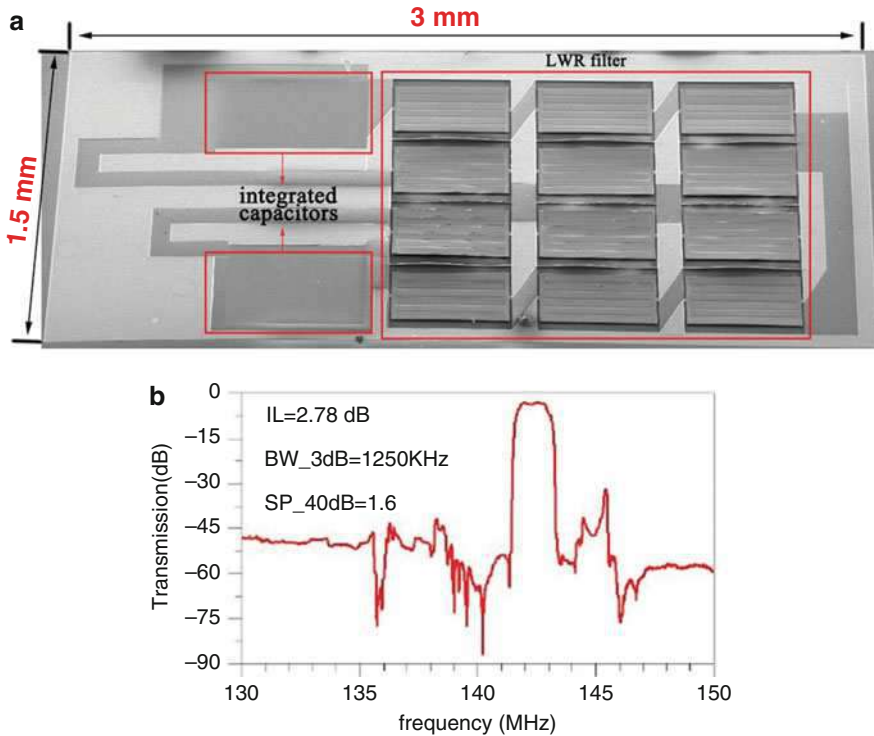
researchers and shows great potential for applications in the IF band. Due to the continuous development of MEMS fabrication processes, Lamb wave resonators have shrunk in size and become perfectly compatible with CMOS technology. In addition, multiple bands are available on one single chip.

Different from surface acoustic waves that propagate along the surface of the bulk material, the Lamb wave travels in the entire piezoelectric film, and the film thickness-to-wavelength ratio is usually less than 1. According to the displacement distribution, these waves are classified into symmetric modes and asymmetric modes. Among all modes, LWRs that use the lowest order symmetric (S0) mode are primarily investigated due to high phase velocity and weak velocity dispersion.

A typical LWR is composed of a suspended piezoelectric film and interdigital transducers (IDTs). Reflectors are indispensable to formation of a standing acoustic wave, and according to the reflector topology, LWRs can be classified into two categories. One category uses metal reflectors, and the other employs etched slots. By etching the piezoelectric film, the interface of the solid and air results in great acoustic impedance mismatch, which contributes to effective reflection of acoustic waves, and thus this topology produces a higher figure of merit and is most frequently implemented.

The AlN thin film is mostly used in LWR, but its moderate coupling coefficient and quality factor restrict its applications. Hence, alternative materials are investigated and exhibit promising features. Lithium niobate (LN) or lithium tantalate have higher coupling coefficients, and Lamb wave resonators based on these materials hold promise in high bandwidth filters. The fabrication of LN thin film can be realized using the Smart Cut™ technology. The process begins with implanting ions into an LN donor wafer. The donor wafer is bonded to an LN or silicon carrier wafer with a thin layer of silicon dioxide, and the ionized LN is split from the bulk material. Using this technique, LN films with a thickness of nanometers can be obtained. An alternative approach to fabrication of LN or thin LT film is use of the chemical-mechanical polarization (CMP) technique to thin the bulk material to micrometers. The LN thin-film-based wafer has been commercialized for years and can be purchased from external vendors. Followed by physical vapor deposition of the metal electrodes, reactive ion etching, and release of the oxide or silicon, the LN thin-film-based LWR is obtained. The  $k_t^2$  values of LN-based LWR exceed 24.1% with an impedance of less than 3 Ohms (Wang et al. 2016). However, due to the anisotropic characteristics of LN, it suffers from various spurious modes, and techniques to suppress them are required.

To promote the quality factor of the resonator, the AlN thin film or zinc oxide (ZnO) thin film can be deposited on bulk materials with high-quality factors, such as silicon or diamond, which is donated from a silicon on insulator (SOI) wafer or a silicon on sapphire wafer. This uniquely structured resonator is also referred to as the thin-film piezoelectric-on-substrate (TPoS) resonator. Because silicon is advantageous for energy storage, TPoS resonators possess a Q factor of over 10,000, which is nearly fourfold that of the conventional AlN LWRs (Tabrizian et al. 2013).



**Fig. 4** (a) SEM image of the LWR filter monolithically integrated with capacitors. (b) Electrical response of the matched filter

### LWR Filter

The resonance frequency of the Lamb wave resonator is dependent on the pitch of the IDTs and is thus restricted by the ability of the mask aligner. Typically, LWR is useful for intermediate frequency bandpass filters.

The ladder-type topology is the most widely used structure in filter design due to low insertion loss and steep roll-off. A high-performance ladder filter must be constructed using resonators with a high figure of merit. Liang et al. investigated the impact of the resonator's geometrical dimensions on the performance of LWR and proposed the design principle for the LWR filter (Liang et al. 2015a). Still, the filter performance is far from satisfactory due to the input impedance mismatch in a 50-Ohm system. Therefore, auxiliary inductors and capacitors are necessary for impedance matching. Due to compatibility with COMS technology, the LWR ladder filter can be integrated with on-chip capacitors in which a thin AlN film serves as both the piezoelectric layer of the resonator and the dielectric layer of the capacitors. As a result, only a pair of inductors are needed, which further contributes to size reductions (Liang et al. 2015b). Figure 4a shows the compact LWR AlN filters, and Fig. 4b shows the electrical response of the matched filter. The matched filter has an insertion loss of 2.78 dB and a notably sharp roll-off.

Another LWR filter structure is based on self-coupled transduction of acoustic waves. Zuo et al. reported a single-chip multi-frequency channel-select filter based on two-port LWRs connected in series. This device features a low insertion loss ( $\sim 2.3$  dB) and small form factor (Zuo et al. 2010). This topology is also extended to LN thin-film-based LWR filters. By taking advantage of the high-coupling coefficient factor of LN, Gong et al. improved the fractional bandwidth of LWR filter to 3.9% (Gong and Piazza 2014). The two-port resonators can also be cascaded in a lattice form (Kim et al. 2013), and by carefully designing the sub-resonators and the matching network, a flexible bandwidth of the filter is achieved.

### LWR Oscillator

Featuring multi-frequency on one chip, Lamb wave resonators are promising candidates for reconfigurable multi-frequency sources. With small form factors, high-quality factors, and perfect compatibility with CMOS, LWRs can outdo conventional quartz or SAW resonators in oscillator circuits in the intermediate frequency range.

The literature (Rinaldi et al. 2011) has demonstrated a reconfigurable CMOS oscillator based on LWRs operating at four different frequencies (268, 483, 690, and 785 MHz). A group of multi-frequency switchable AlN LWRs were connected to a single CMOS oscillator circuit that can be configured to selectively operate in four different states with distinct oscillation frequencies. The advantage of this prototype is that it has the potential to generate multiple stable frequencies without the need for cumbersome and power-consuming phase-locked-loop circuits.

Integrated Device Technology, Inc. presented a pMEMS<sup>TM</sup> oscillator based on a TPoS resonator for high-performance timing applications (Bhugra et al. 2014). The resonator has a high-quality factor ( $Q$ ) of several thousand, enduring stability, and immunity to vibration. This device has potential as a reliable replacement for quartz crystal oscillators in high-frequency timing reference applications.

### Improving Performance

The performance of a typical acoustic resonator is evaluated according to three aspects: quality factor, coupling coefficient, and spurious modes. In this section, we will review techniques used to promote the resonator performance.

#### Q Enhancement

Quality factor refers to the rate of energy loss relative to the stored energy of the resonator. High- $Q$  resonators contribute to construction of oscillators with low noise, filters with low insertion loss, and sensors with low limit of detection (LOD). Hence, enhancement of  $Q$  is one of the most crucial tasks on which researchers focus.

As the definition implies, to maintain high  $Q$  is to reduce energy losses. Typical origins of energy loss include phonon–phonon interactions, electron–phonon interactions, air damping, electrical loading, anchor losses, additional temperature-dependent dissipation, etc. Among all of these causes of poor  $Q$ , anchor losses dominate (Segovia-Fernandez et al. 2015). One way to preserve the energy is to reflect acoustic waves back to the resonator. Harrington et al. developed an in-plane



acoustic reflector at the anchors of the resonators (Harrington and Abdolvand 2011). By etching a pair of trenches with specific distances to the anchors, acoustic waves can be effectively reflected due to the acoustic impedance mismatch. Average Q enhancement of up to 560% is realized for specific designs with reflectors over the same resonators without reflectors. Lin et al. proposed a biconvex-shaped resonator to confine the acoustic energy in the resonant cavity, and the curved edge raises Q by a factor of 2.6 (Lin et al. 2011).

### $K_t^2$ Enhancement

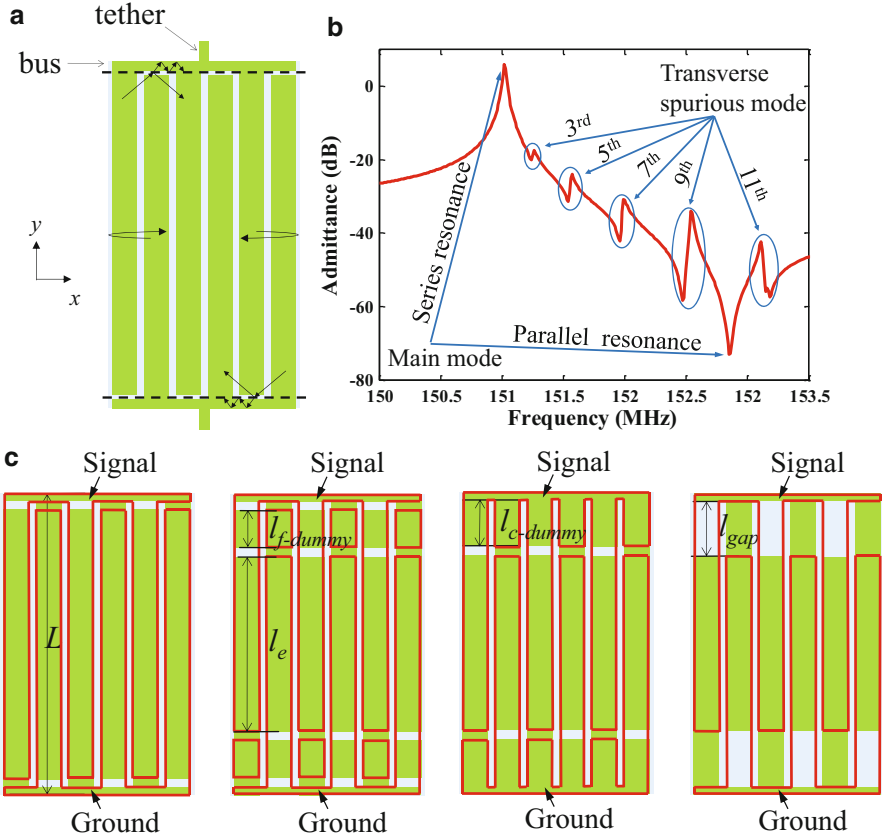
The coupling coefficient factor, i.e.,  $k_t^2$ , represents the efficiency of the mutual transition of the energy between its mechanical and electrical forms and is dominated by two factors: one is the material coupling coefficient ( $K^2$ ) of the piezoelectric layer and the other is the topology of the resonator. As mentioned, LN has a larger  $K^2$  than AlN, and thus LN LWRs have great advantages over AlN LWRs with respect to  $k_t^2$ . In addition, similar to BAW devices, thin Sc-doped AlN films are studied as a high-coupling piezoelectric layer for LWR. Konno et al. developed a 2.6 GHz Lamb wave resonator using a thin  $\text{Sc}_{0.4}\text{Al}_{0.6}\text{N}$  film that exhibits a  $k_t^2$  of 8.3%, which is five times higher than that of the resonators with pure AlN films (Konno et al. 2013).

For an AlN LWR,  $k_t^2$  is primarily mainly dependent on the electrode topology and stack parameters. The electrode topology is classified in four families, i.e., structures with IDTs on both sides, IDT on one side, IDT on one side and a floating electrode underneath, and IDT on one side and a grounded electrode underneath. Of all four structures, those with IDT on both sides yield the highest  $k_t^2$ , which is maximized when the electrode coverage is near 75%. For either topology, the  $k_t^2$  relates to the ratio of the thickness of AlN to the wavelength (which is also two times the IDT pitch). Using the FEA simulation, an AlN LWR is predicted to be able to reach a  $k_t^2$  of up to 10% (Zuo et al. 2012). Enlightened by this work, other researchers have shifted their efforts to achieving a higher  $k_t^2$ . By controlling the ratio, a unique form of Lamb wave, i.e., the Lamé mode, is excited, and a  $k_t^2$  of over 6% has been achieved (Cassella et al. 2016; Zhu et al. 2016).

### Spurious Mode Suppression

Spurious modes are unwanted forms of acoustic resonance caused by various reasons. In LWR filters, these modes generate ripples in the passband or unwanted response in rejection band, and thus mitigation of the spurious modes is of great importance. Spurious modes typically stem from several sources, including flexural modes or antisymmetric Lamb modes, transverse modes or fine-frequency modes, and acoustic waves related to anchoring. In this section, we focus on the spurious mode in the vicinity of the main mode (S0 mode), namely, the transverse mode.

First, it is crucial to investigate the origin of the transverse mode. Acoustic waves travel in the resonator and are reflected at the boundaries, as illustrated in Fig. 5a. The wanted waves travel along the x-axis direction, whereas the transverse waves mainly propagate along the y-axis direction (Zhang et al. 2015b).



**Fig. 5** (a) Illustration of the propagation behavior of the S0 mode in rectangular LWR. (b) Simulation results demonstrating the electric performance of the main mode and the lowest five transverse modes in admittance curve. (c) Illustration of the four kinds of IDT structures, including the conventional structure, the f-dummy IDT structure, the c-dummy IDT structure, and the gap structure. The bottom electrodes are outlined by red lines and the top electrodes are patterned in green

To form a standing wave in the resonator, waves all satisfy the condition:

$$W = \frac{n}{2} \lambda_x, L = \frac{m}{2} \lambda_y, \tag{1}$$

where  $W$  is the pitch of the IDT,  $L$  is the aperture of the IDT,  $n$  and  $m$  are both integers, and  $\lambda_x$  and  $\lambda_y$  are the wavelengths in the x-direction and the y-direction, respectively. The resonance frequency of the  $(n, m)$  mode can be expressed as follows:

$$f_{n,m} = v_0 \sqrt{\left(\frac{n}{2W}\right)^2 + \left(\frac{m}{2L}\right)^2} = \frac{nv_0}{2W} \sqrt{1 + \left(\frac{mW}{nL}\right)^2}, \tag{2}$$



where  $v_0$  is the wave velocity of the S0 mode. The (1, 1) mode is defined as the fundamental S0 mode or the main mode. When  $n = 1, m = 3, 5, 7, \dots$ , such modes are the  $m^{\text{th}}$  transverse spurious modes. The  $k_t^2$  of the transverse modes in LWR can be expressed as follows:

$$k_t^2 \approx \frac{\tilde{\epsilon}_{31}^2}{\tilde{c}_{11}\tilde{\epsilon}_{33}} \frac{(\int_S S_1 \Phi dS)^2}{(\int_S S_1^2 dS)(\int_S \Phi^2 dS)} \propto \frac{\left(\int_0^L \sin\left(\frac{m\pi}{L}y\right)\Phi dy\right)^2}{\left(\int_0^L \sin^2\left(\frac{m\pi}{L}y\right)dy\right)\left(\int_0^L \Phi^2 dy\right)} \propto \frac{1}{m^2}. \quad (3)$$

A 3-D COMSOL FEA simulation is used to demonstrate the transverse modes, as shown in Fig. 5b. The electric field distribution in the cavity can be changed, thus decreasing the  $k_t^2$  of the spurious modes.

Zhang et al. proposed three structures to suppress the third transverse modes, i.e., the floating dummy (defined as “f-dummy”) electrode structure, the connecting dummy electrode (defined as “c-dummy”) structure, and the gap structure, as illustrated in Fig. 5c (Zhang et al. 2015b).

For the f-dummy and c-dummy electrodes, according to Eq. 3, the  $k_t^2$  can be rewritten as follows:

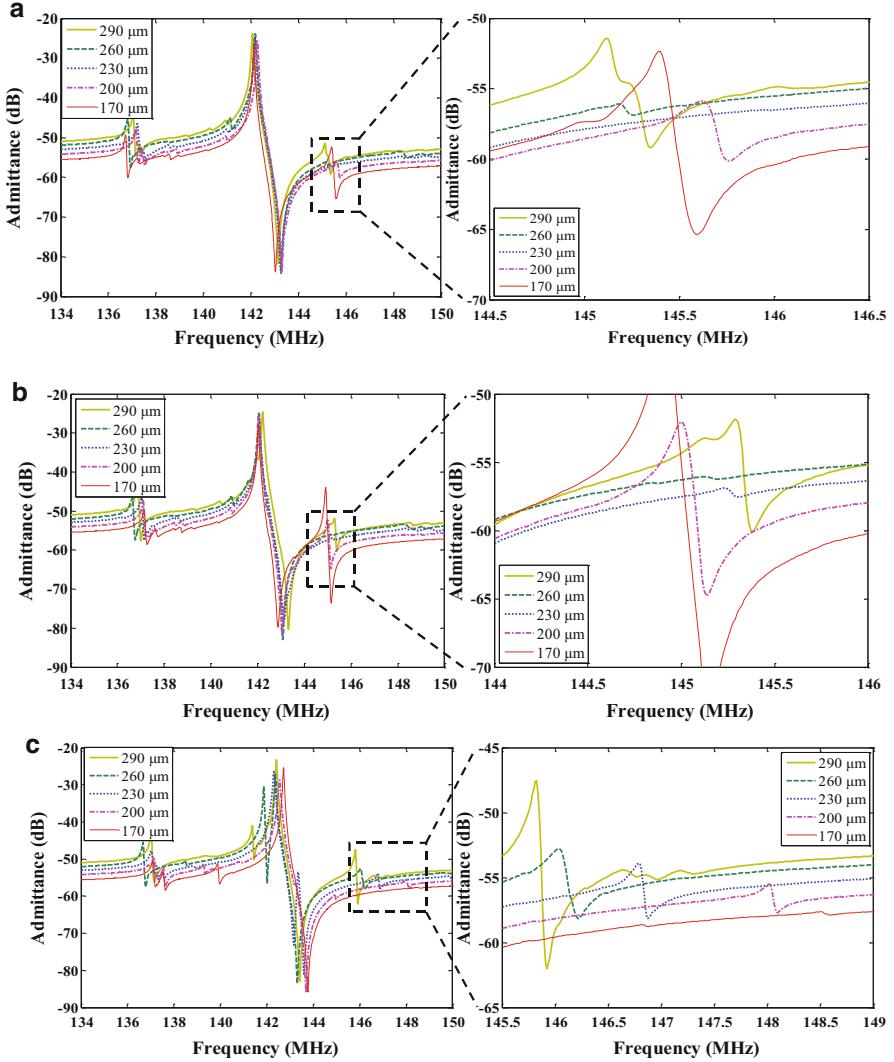
$$k_t^2 = \frac{\tilde{\epsilon}_{31}^2}{\tilde{c}_{11}\tilde{\epsilon}_{33}} \frac{(\int_S S_1 \Phi dS)^2}{(\int_S S_1^2 dS)(\int_S \Phi^2 dS)} \propto \frac{\left(\int_{\frac{L-l_e}{2}}^{\frac{L+l_e}{2}} A(x,y) \sin\left(\frac{m\pi}{L}y\right)\Phi dy\right)^2}{\left(\int_0^L A^2(x,y)\sin^2\left(\frac{m\pi}{L}y\right)dy\right)\left(\int_{\frac{L-l_e}{2}}^{\frac{L+l_e}{2}} \Phi^2 dy\right)} \propto \frac{\sin^2\left(\frac{m\pi l_e}{2L}\right)}{m^2 \frac{l_e}{L}}, \quad (4)$$

where  $l_e$  is the effective electrode length. When  $m = 3$  and  $l_e = (2/3)L$ , the  $k_t^2$  of the third transverse mode becomes 0, indicating an elimination of this mode.

For the gap structure, the desirable  $l_e$  for the gap structure can be written as follows:

$$\frac{v_e^2}{(L-l_e)^2} - \frac{v_g^2}{(2l_e)^2} = \frac{v_e^2 - v_g^2}{W^2}. \quad (5)$$

Resonators with different sizes are designed and fabricated ( $L = 380 \mu\text{m}$ ,  $l_e = 290 \mu\text{m}$ ,  $260 \mu\text{m}$ ,  $230 \mu\text{m}$ ,  $200 \mu\text{m}$ , and  $170 \mu\text{m}$ ). Figure 6 shows the admittance curves of all the devices, and Fig. 6a, b shows the suppression results of the f-dummy and the c-dummy configuration, respectively. In agreement with Eq. 5, the third transverse



**Fig. 6** Measured electrical response for (a) f-dummy electrode, (b) c-dummy electrode, and (c) gap structure with effective length equaling to 290, 260, 230, 200, and 170 μm

mode is suppressed when  $l_e$  is approximately  $2/3$  of the aperture. The gap structure with  $l_e = 170 \mu\text{m}$  eliminates the third transverse mode completely, as shown in Fig. 6c. All of the methods with appropriate geometries can effectively eliminate the third transverse spurious mode in proximity to the main mode, thus demonstrating great potential for synthesis of spurious-free filters.

An alternative approach to suppressing the transverse mode is to modulate the mechanical boundary conditions of the resonator, thus mitigating the coupling

coefficient of the transverse mode. Zhang et al. proposed vertical and lateral protrusion structures for the suppression (Zhang et al. 2015a). Finite element analysis verifies that the use of these structures effectively restrains the transverse modes, and the measured electrical performance of the LWR with protrusions demonstrates a reduction in the spurious response of 11 dB. Giovannini et al. proposed an apodization method which was developed for SAW transducers to suppress the transverse mode in LWR. By patterning the IDT fingers in a sine-function form, the amplitude of the transverse mode is effectively reduced (Giovannini et al. 2014). However, as a side effect, the electromechanical coupling coefficient is slightly compromised. The issue of spurious modes still attracts researchers' attention, and effective countermeasures that do not affect the figure of merit of the resonator are consistently pursued.

---

## Piezoelectric Micro-/Nano-Mechanical Sensors

### Sensing Mechanism

Known for their excellent mass resolution, piezoelectric micro- and nano-mechanical sensors have been studied extensively. In this section, we offer a brief introduction to different resonant modes and sensing mechanisms to lay a theoretical foundation for applications.

#### Thickness Extensional Mode (TE)

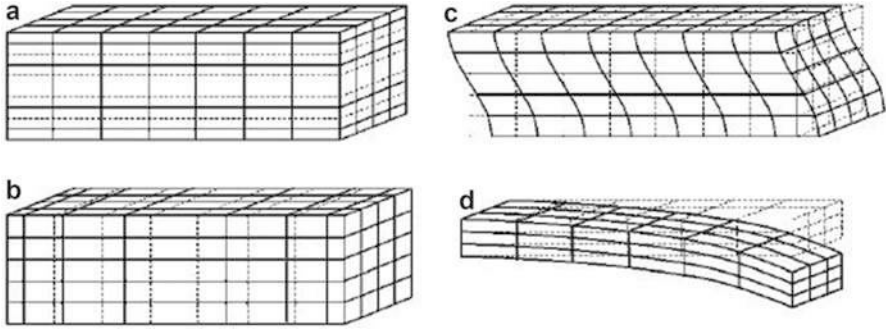
The thickness extensional mode is a longitudinal mode excited in a vertically grown piezoelectric material film by coupling the vertical electric field through the  $d_{33}$  piezoelectric coefficient. Both the wave velocity and resonance frequency of the TE mode are higher than those of any other possible modes for a given piezoelectric film and can offer the highest sensitivity. However, when a TE-mode resonator is immersed in a liquid, it suffers greatly from damping due to acoustic energy absorption by the liquid (Zhang et al. 2011; Zhang and Kim 2005).

#### Thickness Shear Mode (TS)

Excitation of the shear mode in thin-film bulk acoustic resonators commonly requires piezoelectric films with an inclined c-axis. Alternatively, a shear wave can be generated with a lateral electric field applied between two patterned electrodes on one surface of the piezoelectric film. Another approach is to make the bottom electrode larger than the top electrode to create a lateral electric field in a piezoelectric film. The TS-mode is considered superior to TE mode for operation in liquid. However, the wave velocity and resonance frequency of a shear wave are less than those of a longitudinal wave, which results in a lower mass sensitivity.

#### Lateral Extensional Mode (LE)

Lateral extensional mechanical stress can be produced by an electric field applied across the piezoelectric film thickness through the  $d_{33}$  piezoelectric coefficient



**Fig. 7** Grid diagrams for plane uniform particle displacement waves propagating in solids associated with (a) thickness extensional (*TE*) mode, (b) lateral extensional (*LE*) mode, (c) thickness shear (*TS*) mode, and (d) flexural mode, which are four common modes of waves propagating in piezoelectric resonant sensors (Pang et al. 2012)

(Pang et al. 2006). In this case, the resonance frequency is primarily determined by the resonator's lateral dimension, which can be defined lithographically. The sidewalls of the piezoelectric film are used in mass sensing.

### Flexural Mode

Flexural mode vibration can be excited in cantilevers and clamped-clamped beams. The resonance frequency of a cantilever is dependent on its effective spring constant and mass. Thus, the changes in absorption-induced surface stress and mass contribute to the resonance frequency shifts.

To summarize, the deformation patterns associated with four modes mentioned previously are depicted in the form of finite elements, as shown in Fig. 7.

The following linear equation linking the resonance frequency shift and mass loaded onto the resonant sensor surface was first developed by Sauerbrey (Pang et al. 2012):

$$\Delta f = -\frac{2f_0^2}{A\sqrt{\rho_q\mu_q}}\Delta m \quad (6)$$

This equation applies only to systems that meet the following three conditions: the deposited mass must be rigid, the deposited mass must be distributed evenly, and the ratio of frequency change  $\Delta f/f$  must be less than 2%. This description can be understood from the fact that the layer of loaded mass enlarges the path length of the acoustic wave, thus lengthening its half wavelength and decreasing the resonance frequency.

The natural resonance frequency of an acoustic resonator can be estimated by the equation:

$$f_0 = \frac{1}{2d}\sqrt{\frac{c}{\rho}} \quad (7)$$

From Eqs. 6 and 7, we can derive an approximate equation for nearly all acoustic mass sensor types:

$$\frac{\Delta f}{f_0} \approx -\frac{\rho_m d_m}{\rho_0 d_0} \quad (8)$$

where  $f_0$  is the resonance frequency;  $\Delta f$  is the frequency shift;  $\rho_m$  and  $d_m$  are the density and thickness of the added layer, respectively; and  $\rho_0$  and  $d_0$  are the density and thickness of the resonator, respectively. This equation indicates that the resonance frequency of an acoustic resonator is linearly related to the mass of the material absorbed onto the resonator. The minus sign in the equation indicates that the resonance frequency of a resonant sensor decreases as the loaded mass increases.

## Biological Macromolecule Sensing

Biomolecule detection plays an essential role in clinical diagnosis, drug discovery, and counter-bioterrorism. Conventional methods of biomolecule detection require chemical labels or fluorescent tags, which are not only time consuming and expensive to manufacture but might also compromise the activity of biomolecules. Certain biosensors are used to study biomolecular interactions in real time and in situ, without the use of chemical tags. Currently, several sensing methods are under exploration for label-free study of biomolecular interactions and include surface vibration spectroscopy, light interference methods, mass-sensitive QCMs, and micro-cantilevers.

FBAR with a functionalized surface can be used to detect biomolecules (proteins and DNA) in real time. Because the sensing mechanism of FBAR is the resonance frequency shift (decrease in the resonance frequency) caused by the mass of any material added to the surface, this technique does not require fluorescent tags or chemical labels. For FBARs operating in the gigahertz region, the mass sensitivity is much higher than that of a QCM with an operating frequency in the megahertz range (Pang et al. 2012). Thus, label-free, highly sensitive and selective, and real-time detection of biomolecules can be achieved by FBAR.

An FBAR sensor is basically an Al/ZnO/Al/SiN composite resonator. For various sensing purposes, a highly thin gold layer ( $\sim 800 \text{ \AA}$ ) is deposited on the surface of the resonator for chemical modification. For protein detection (Zhang et al. 2010), real-time sensing of protein–ligand interaction uses the biotin–streptavidin system as a model because such a system is well characterized. The gold layer of FBAR allows immobilization of probe molecules using thiolation chemistry. The FBAR device was treated with  $5 \times 10^{-4} \text{ M}$  biotin-HPDP in ethanol solution for 24 h at room temperature, and streptavidin solution was added. Upon binding to streptavidin in buffer, the modified FBAR produced a significant resonance frequency shift, allowing real-time monitoring of the biotin–streptavidin interaction. To rule out any possibility of nonspecific or false resonance shifts, an 11-mercaptop-1-undecanol-treated

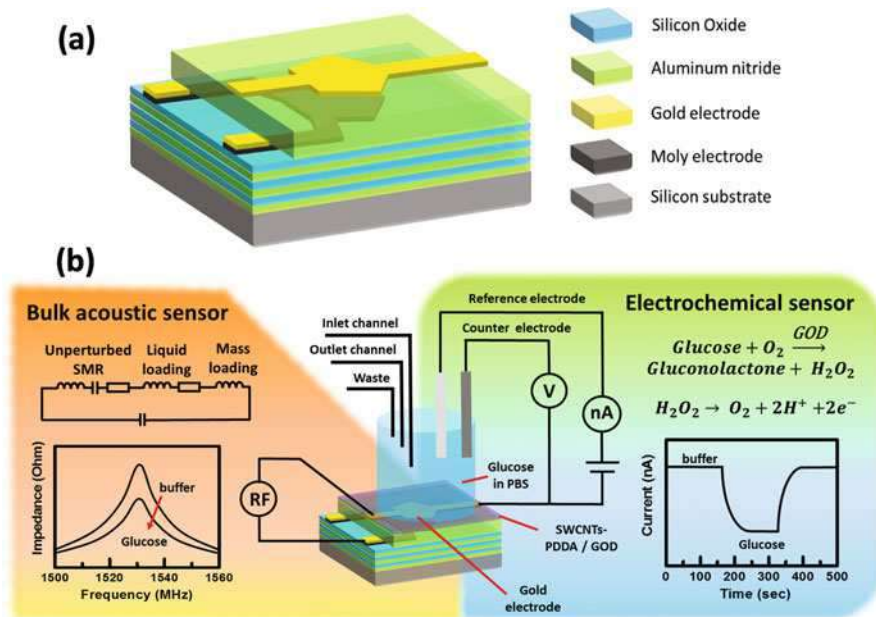
FBAR was also prepared and tested against streptavidin solution, and the device did not show any noticeable resonance frequency shift. Therefore, the large resonance frequency shift produced on the biotin-modified Au-coated FBAR resulted from the specific biotin–streptavidin interaction.

FBAR is also a promising candidate in sequence-specific DNA sensing (Zhang et al. 2007). A 15-mer probe nucleotide was functionalized at the 5' end with an HS-(CH) group for immobilization on the Au-coated FBAR device. An Au-coated FBAR device was treated with HS-ssDNA for 2 h followed by immersion for 1 h in 1.0 mM mercapto-hexanol (MCH) in aqueous solution to form a mixed monolayer. The resonance frequency decreased significantly ( $\sim 70$  kHz) when the DNA-probe-immobilized FBAR was exposed to the target complementary sequence ( $1 \mu\text{M}$ ). To rule out any possibility of nonspecific or false resonance shifts, an MCH-only coated FBAR was treated with the same complementary DNA strain. No resonance shift was observed, indicating that the resonance shift is due to a specific probe–target DNA interaction. With the single-mismatch sequence, the resonance frequency shift was 35% ( $\sim 25$  kHz) of that ( $\sim 70$  kHz) obtained with the complementary sequence, demonstrating the ability to distinguish complementary DNA from a single-nucleotide mismatch DNA sequence.

## Multimode Sensing System

Biological and chemical sensors based on fluorescence labeling, surface plasmon resonance (SPR), electrochemistry, the field-effect transistor (FET), and the acoustic wave resonator have been successfully realized. Although each of these techniques has advantages, detection based on single method can only deliver limited information. For example, fluorescence labeling enables visual monitoring of biomolecular interaction but cannot quantify the interaction from additional aspects, such as the change of mass, which can be detected using QCM and thin-film piezoelectric resonators, or the change of electric charge, which can be acquired by FET and electrochemical sensors. However, the existing multimode sensors cannot be used in the form of a small-sized array, which makes the sensing process time consuming and costly and requires a large amount of sample.

The development of microfabrication technology has offered new opportunities for miniaturization of devices and integration of multifunctional sensors on a single chip. Zhao et al. introduced a novel CMOS-compatible microelectromechanical sensing system that integrates four pairs of electrochemical working electrodes and SMR on one chip (Zhao et al. 2016). This system enables detection of analytes with three sensing modes: electrochemistry, gravimetry, and viscometry. Figure 8 shows the schematic and working principle of the chip. By taking advantage of the compatibility of the material and structure, the two sensors share the same electrode and the film underneath. The gold electrode layer on the surface serves as the top electrode of SMR and also as the working electrode of electrochemical sensor, which facilitates surface modification and characterization of the sensing electrode. The piezoelectric AlN film and the acoustic reflector of SMR also work as dielectric films



**Fig. 8** (a) Schematic of the chip from an oblique view in which the electrochemical sensor and SMR share the same film stack with good compatibility; (b) schematic of the electrochemical-piezoelectric integrated sensing system

underneath the electrochemical working electrode, which help to block DC cross talk between the electrodes.

A glucose concentration monitoring experiment was performed to verify the device performance. The electrochemical sensor can be used to detect the redox reaction of glucose, and SMR shows the ability to monitor viscosity changes induced by concentration variation and the modification process of functional materials. For specific detection of glucose, the electrode surface was functionalized with single-walled carbon nanotube–poly(dimethyldiallylammonium chloride)/glucose oxidase (SWCNTs-PDDA/GOD) composite films. The SMR frequency shift was measured each time after one single layer was formed to precisely determine the mass increment. The real-time electrochemical detection relies on the redox current produced by the enzymatic product. With the successive addition of glucose concentration (0.05–256 mM), a progressively increasing current was observed. The sensitivity and the detection limit are estimated as 0.78 nA/(mM mm<sup>2</sup>) and 6 M (S/N = 3), respectively. However, the impedance magnitude of SMR at the resonance frequency decreases with the increase in glucose concentration due to the heavy load caused by higher viscosity. Thus, glucose concentration can be successfully monitored with a sensitivity of 0.422 mΩ/mM in logarithmic form and a detection limit of 0.05 mM.



Liu et al. used the same integrated dual sensor device for static and non-static assays for detection of polyelectrolyte (PETs) surface assembly and antibody–antigen interactions (Liu et al. 2017). Due to the different sensing mechanism, complementary and inconsistent information can be extracted, which emphasize the critical dependence of the sensing results on the sensing mechanisms and offer valuable guidance for better understanding of bioreactions.

## Piezoelectric Lab on a Chip

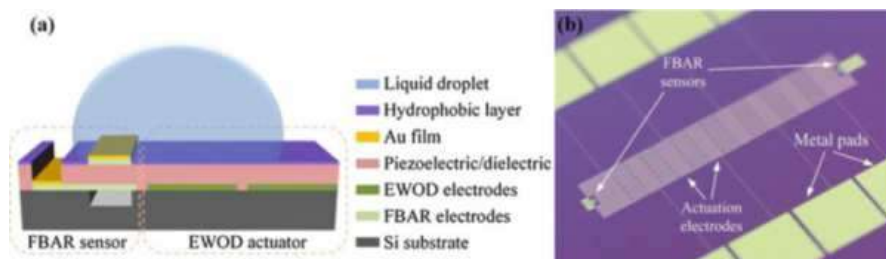
The lab-on-a-chip system is able to perform fast high-resolution separations and detections on an integrated chip with low cost and low consumption of samples and reagents. This approach has been used in a broad range of applications, such as molecular analysis, biodefense, and molecular biology. Digital microfluidics, which is based on the electro-wetting-on-dielectric principle, involves several basic operations, such as droplet creation, transport, cutting, merging, and mixing and has been widely applied among biological and chemical assays. By virtue of the ability to integrate powerful sensors, microfluidic technologies are making strides in enhancing competitiveness for lab-on-a-chip applications. Piezoelectric microsensors have inherently smaller sizes compared with conventional sensors and have the potential to become fully integrated with digital microfluidics because both of their fabrication processes are IC compatible. The batch fabrication property is also a key advantage in point-of-care-testing and high-throughput detection.

Zhang et al. demonstrated a lab-on-a-chip system with digital droplet actuator and FBAR sensors integrated on a single silicon chip (Zhang et al. 2014a). The digital microfluidic actuator uses patterned actuation electrodes, a dielectric film on top of the electrodes, and a hydrophobic layer on the dielectric film. Thin AlN film is a critical component of the FBAR piezoelectric layer and the digital microfluidic dielectric layer, which facilitates the integration of the two devices on a single silicon chip (Zhang et al. 2015c). Photos of the fabricated microsystem are displayed in Fig. 9.

The digital feature of digital microfluidics aids in substantially improving the signal-to-noise ratio of the FBAR sensor in liquid solution. Although the wet-sensing mode is real-time monitored, it has a much worse LOD than the dry-sensing mode because the  $Q_s$  of the FBAR resonator in a liquid environment is degraded by nearly two orders of magnitude (from 3745 to 50). Experimental results show that random variation of the resonance frequency in the wet-sensing mode is approximately 5 kHz, whereas the value is as small as 0.3 kHz in the dry-sensing mode. The mass sensitivity is known to be approximately 900 cm<sup>2</sup>/g, and the LODs are calculated as 2.2 and 0.13 ng/cm<sup>2</sup> in the wet- and the dry-sensing modes, respectively.

Droplets containing Hg<sup>2+</sup> of various concentrations are transported and sensed to demonstrate the capability of the integrated system. The sensitivity and LOD of the system for Hg<sup>2+</sup> detection are found to be 250 kHz/μM and 10<sup>-9</sup> M, respectively. In addition to ion detection, this approach has great potential for use in more sophisticated studies, such as protein–ligand interactions, DNA hybridization, and immunoassays. In addition, the integrated single chip has the potential to detect





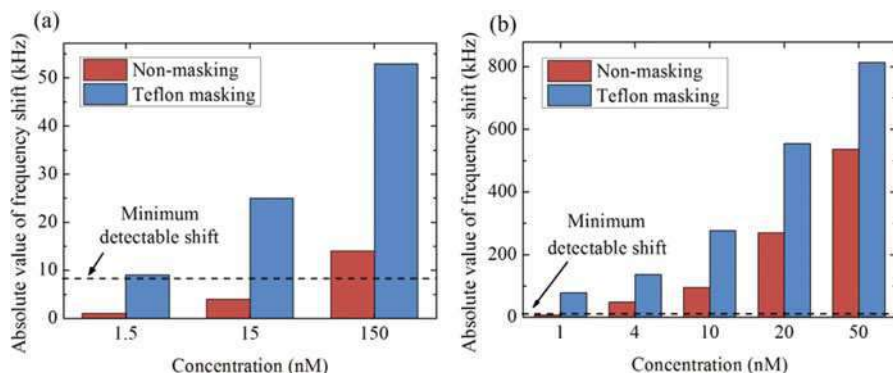
**Fig. 9** Digital microfluidic chip integrated with a monolithic FBAR sensor. (a) Cross-sectional schematic; (b) photo of the fabricated device

various biotargets in parallel with different bioreceptors functionalized on the corresponding sensors by digital microfluidic actuator (Huang et al. 2015). The tiny droplet manipulation ability facilitates different functionalizations on a high spatial density biosensor array, which is nontrivial high-throughput work.

FBAR was also used to detect particulate matter content in a microfabricated air-microfluidic sensor developed by Paprotny (Paprotny et al. 2013). The PM sensor is only  $25 \times 21 \times 2$  mm in size and is two orders of magnitude smaller than commercially available direct PM<sub>2.5</sub> mass sensors. The small size allows the sensor to be used in continuous recording of personal PM exposure levels. The sensor contains an air-microfluidic conduit that separates the particles by size (virtual impactor) and subsequently transports and deposits the selected particles using thermophoretic precipitation onto the surface of a microfabricated mass-sensitive FBAR. The mass-loading of the FBAR causes a change in its resonance frequency, and the rate of the frequency change corresponds to the particle concentration in the sampled air volume. The paper presented experimental results that demonstrate the performance for measurement of PM mass emitted from diesel exhaust and tobacco smoke and showed a sensitivity approaching  $2 \text{ g/m}^3$  over a period of particle accumulation (10-min maximum).

## Sensor Signal Enhancement Strategy in Liquid

Microscale biosensors have an intrinsically higher sensitivity and lower LOD partially due to their miniaturized size. Researchers have studied various methods for material and structure optimization to improve the LODs for analyte detection at low concentrations. Local adsorption techniques have been demonstrated as an effective method in which target adsorption on the non-sensing area is minimized and the target molecules only bind on the exposed sensing area. The target molecule loss on the unwanted area is effectively reduced, thus preserving the bulk concentration of target near the sensing area. A higher fraction of the target molecule is expected to adsorb on the sensing area. Therefore, biosensors constructed with the local adsorption method should have a larger signal response and hence an improved LOD for analyte detection.



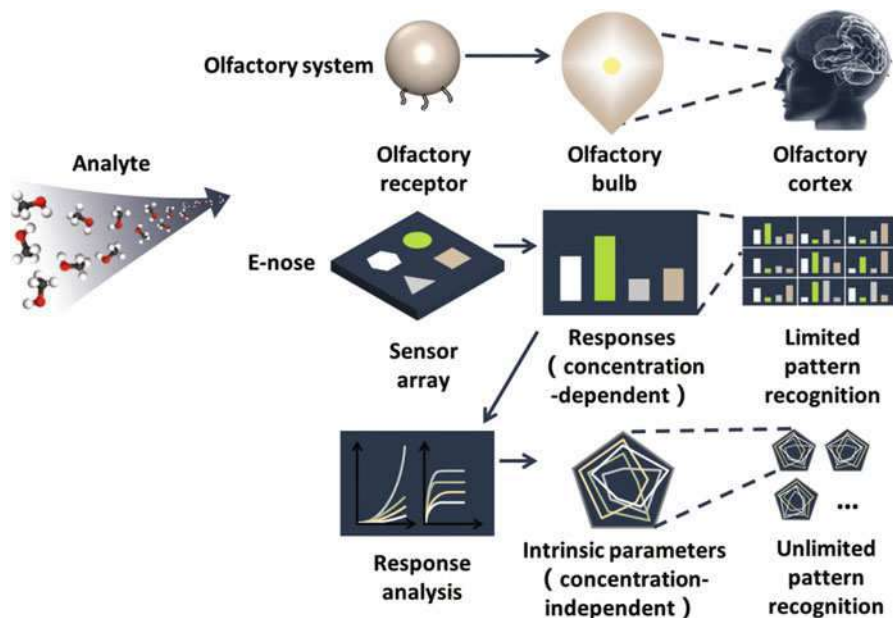
**Fig. 10** Comparison of FBAR sensor signals for both (a) BSA detections at 150, 15, and 1.5 nM, and (b) anti-IgG detections at 50, 20, 10, 4, and 1 nM. The incubation time is 20 min

Zhang et al. studied the signal enhancing mechanism and kinetic issues of the local masking adsorption method using FBARs. A thin film of Teflon is first formed, covering both the sensing and non-sensing areas, and the sensor is subsequently patterned using a microfabrication process to ensure that the sensing area is exposed. The simulation and experimental results show that FBAR biosensors integrated with the local adsorption mask have a larger signal response (Zhang et al. 2014b, 2015d).

For both BSA and anti-IgG detection experiments, the author applied masked FBAR sensors to detect protein concentrations down to the nanomolar scale, as shown in Fig. 10. The sensor responses of the Teflon-masked sensor are always higher than those of the non-masked sensors in a low-protein concentration range with both BSA physical adsorption and anti-IgG-specific binding. The larger signals on the Teflon-masking sensor versus the control group at various concentrations clearly confirm the amplifying effect, thus proving the hypothesis. In detection of both BSA and anti-IgG, comparisons between the LODs of masked and non-masked FBAR sensors were conducted. The absolute value of the minimum detectable frequency shift (device noise) is near 8 kHz. The signal of 15 nM BSA measured from the non-masked sensor is below the level of device noise, whereas the Teflon-masked sensor shows a significant frequency shift under the same conditions. The LOD of a Teflon-masked FBAR sensor for BSA detection is approximately 1.5 nM. Anti-IgG detection shows similar results at 1 nM. Therefore, the LOD of an FBAR sensor is experimentally improved by at least one order of magnitude via the Teflon-masking method.

## VOC Gas Sensing

Volatile organic compounds (VOCs), which are analytes emitted from indoor environments and industry production, are believed to have harmful effects on human



**Fig. 11** Schematic illustration of the mechanism of an olfactory system and e-nose system (both concentration dependent and concentration independent)

health. Moreover, certain VOCs in exhaled breath are found to be effective biomarkers for the detection of diseases (e.g., lung cancer, breast cancer, and diabetes). As a result, highly sensitive and selective detection of VOCs is of vital importance for indoor air-quality control and pollution/environmental monitoring as well as diagnostics purposes. Hence, a portable system for on-site analysis of VOCs components is preferred for these applications. However, most current commercial VOC sensors are developed for concentration information, leading to the problem of analyte discrimination. The electronic nose (e-nose) system consists of an array of chemical sensors and is a powerful tool for VOC analysis, as shown in Fig. 11. Each sensor is chemically modified with gas-sensitive materials, creating a response “fingerprint.” Generally, the discrimination capability of the e-nose is improved with a highly integrated sensor array.

The acoustic wave resonator is one type of sensor with potential for e-nose applications. Based on gravimetric measurement, the response is directly linear to the VOC amount, and real-time sorption can be recorded as well. Thus far, many studies have been performed on VOC detection with acoustic wave resonators, particularly the SAW device and QCM, with good results. However, the operation frequencies of those resonators are limited and cannot be used in highly sensitive VOCs detection. With frequencies extended to GHz, the FBAR sensor shows higher sensitivity (detection limit at the sub-picogram level) for chemical sensing and has yielded progress in e-nose applications. Moreover, FBARs can be easily integrated

into the portable e-nose system, which benefits from micrometer-scale size and CMOS-compatible fabrication.

For the e-nose system, surface modification of the sensors plays an important role in molecular recognition. Due to ambient temperature operation and diverse sensor coatings, polymer coating is one type of popular surface modification methods. A novel method of polymer coating is the layer-by-layer (LbL) self-assembly approach, which is introduced by assembling positively and negatively charged polyelectrolytes in a layer-by-layer fashion via electrostatic interactions. It has been demonstrated that LbL assembly of poly(acrylic acid) (PAA) and poly(4-vinylpyridine) (PVP) on the FBAR surface could be used in tuning the resonance frequency of devices. Moreover, the LbL polymer coating offered a direct method for modification of the FBAR surface and facilitation of VOC sensing applications (Liu et al. 2014).

The conventional fingerprints of the e-nose systems are based on the sensor response to a known concentration of vapor. Different VOCs can produce similar fingerprint patterns, and the discrimination capability is limited. Hence, multiple concentration-independent parameter fingerprints are of greatest importance. However, the polymer modification process is complex, and the thick layer suffers from the irreversible gas adsorption, which might induce “e-nose poisoning.” Self-assembled monolayers (SAMs) are highly ordered molecular assemblies that show resistance to “e-nose poisoning” due to the thin layer. It has been reported that different SAM-functionalized FBARs were integrated into e-nose sensor arrays to build a concentration-independent fingerprint library (Chang et al. 2015, 2016). The adsorption isotherms of VOCs on SAMs were obtained by gravimetric detection and fitted to a BET equation, and the desorption kinetics were fitted to a JMA equation. Relying on the parameters obtained from the affinities and kinetic fitting (monolayer adsorption capacity, adsorption energy constant, and desorption rate), concentration-independent fingerprints were established. Such a fingerprint library facilitates the discrimination capacity of VOC analytes with the FBAR array, offering a more powerful tool for VOC analysis. Moreover, the VOC adsorption effects of the surface chemical groups were analyzed to study the gas–surface interaction, which indicates that such an e-nose can be used in gas-sensitive material selection.

In the last few years, supramolecular functionalization has been applied in VOC detection. Based on specific “host–guest” interactions, the specificity and sensitivity of sensors can be highly increased with supramolecular coatings. An e-nose-type integrated FBAR array in which each sensor was coated with different supramolecular monolayers has been proposed for VOC detection (Lu et al. 2015). The supramolecular monolayers used in this work were fabricated using the Langmuir–Blodgett techniques as a specific sensing interface for VOC recognition (Lu et al. 2015). The obtained affinity constants for each VOC indicate that vapor molecule adsorption occurs inside and outside of the supramolecular cavities. Kinetic information for adsorption and desorption was obtained from the exponential fitting results as well. The parameters can be used as multiparameter fingerprint patterns for highly selective detection and discrimination of VOCs.

Generally, the discrimination capability of an e-nose system is limited with only one type of sensor. Hence, multimode sensing with integration of different sensors is required. Based on multimode sensing, the information from one type of sensor can be complementary to the information from the other type of sensor. A multimode e-nose that combines FBAR with the silicon nanowire field-effect transistor (Si-NWFET) has been proposed for analysis of VOCs. Relying on gravimetric detection, the response of FBAR is related to the concentration of VOCs because the absorbed mass increased with VOC concentration. Si-NW FET can measure the electrical potential change due to the absorption of dielectric VOCs molecules. Based on these two complementary sensors, the multimode e-nose can be used in detection of VOC dual mixtures, such as ethanol and hexane. It was shown that Si-NW FET only responds to the polar ethanol and remains unchanged with nonpolar hexane. In combination with FBAR, concentration information for both VOCs can be directly detected. These two complementary sensors contribute to a novel e-nose for multimode sensing of VOCs and show potential for analysis of polar/nonpolar VOC mixtures (Ye et al. 2016).

High-Q piezo-transduced single-crystal silicon bulk acoustic resonators (PSBARs) were integrated into an e-nose sensor array as well (Yuan et al. 2016). Because these unique acoustic devices possess both a high-quality factor and relatively low motional resistance, PSBARs meet the requirements for high-performance clocking circuits and have been successfully used as vital frequency selection and phase-shifting segments in preeminent oscillators with low phase noise and high stability. Three types of SAMs were functionalized on three PSBARs to act as a highly sensitive e-nose sensor array. The different frequency responses of the modified PSBARs eventually produce a unique code bar for the individual VOC, therefore realizing discrimination of different VOCs. Moreover, it was demonstrated that the PSBAR sensor has the ability to detect low concentration of VOCs (50 ppm for ethanol), which was the lowest detection limit that has been reported for monolayer functionalized BAW VOCs sensors.

Recently, research studies on the Internet of Things (IoTs), which can consistently monitor the vital signs of human or environmental factors, are rapidly increasing. Duan et al. have developed a wireless and passive gas sensor by integrating the Lamb wave resonator with an inductance antenna (Duan et al. 2016). The operation of the Lamb wave resonator is experimentally demonstrated as a temperature sensor and a VOC gas sensor. Wireless and passive detection is achieved through near-field electromagnetic coupling between the spiral inductance coils (sensor antenna) and the readout coil connected to a vector network analyzer (VNA), which displays the resonance frequency of the piezoelectric resonators. Wireless and real-time gas detection of two different gases (ethanol and cyclohexane vapor) by two chips decorated with different monolayer materials (OTES and PFDTs) has been demonstrated. The wireless interrogated piezoelectric microsensors without an additional power source are ideal for wearable human activity monitoring and personal healthcare applications.

## Impact of Nanotechnology

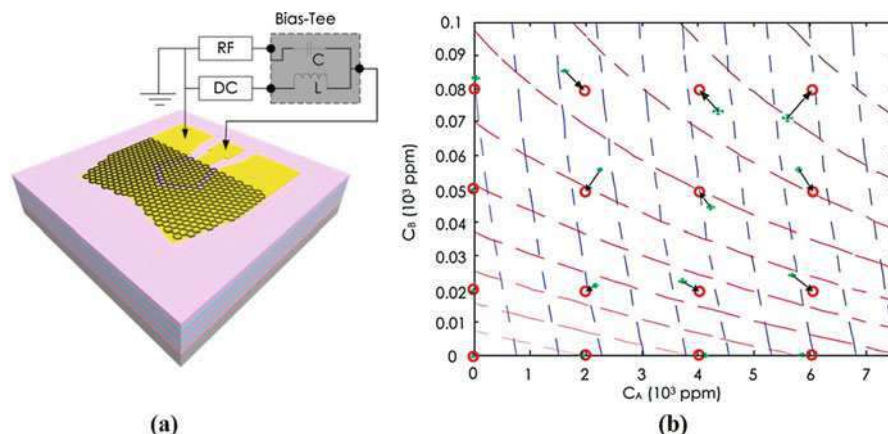
The development of modern sensors is targeted toward smaller sizes to promote high integration with other electronic devices and to create new application scenarios such as real-time air-quality monitoring. However, researchers have made an effort to construct ultrahigh sensitive sensors to fulfill environmental sensor network requirements. The various novel sensors based on nanomaterials have the opportunity to push the current frontier (Feng et al. 2016; Schedin et al. 2007). Nanomaterials have demonstrated their great applied potential to detect and measure dangerous gaseous chemicals under laboratory conditions. The emerging nanomaterials possess the advantages of an ultrathin structure and outstanding physisorption properties, and thus they can operate as a sensitive component in solid-state sensors. The integration of piezoelectric sensors with nanomaterials creates a promising experimental platform for trace gas detection.

A dual-mode gas sensor based on the AlN SMR was developed for simultaneous and independent acquisition of electrical and mechanical signals from the same gas adsorption event (Chen et al. 2016a). The device integrated a graphene field-effect transistor with a piezoelectric microsensor by leveraging multiple structural and functional synergies, as shown in Fig. 12. The dual signals resulting from independent physical processes, i.e., mass attachment and charge transfer, can reflect the intrinsic properties of gas molecules and potentially enable simultaneous target recognition and quantification. These simultaneous measurements of mass adsorption and charge transfer are expected to produce a more precise understanding of the interactions between graphene and various gas molecules. In addition to its practical functions, the device served as an effective tool to quantitatively investigate the physical processes and sensing mechanisms for a large library of sensing materials and target analyses. Because gas sensors for human health and air-quality monitoring have received extensive attention, researchers from both academia and industry are actively seeking innovative solutions that deliver satisfactory selectivity and form factors at the same time. The gas sensor of such a type satisfies the above concern.

Acoustic charge transport in monolayer graphene on the AlN SMR was also investigated (Chen et al. 2016b). The acoustic waves induced macroscopic current flow that linearly scaled with input power. The AE current exhibited unique frequency dependence due to the special configuration and piezoelectric properties of the transducer, which led to transitions between traveling and standing acoustic waves across a characteristic frequency.

A graphene/LiNbO<sub>3</sub> hybrid device was used to investigate the acoustic induced charge transport in chemically doped graphene (Zheng et al. 2016). The chemical doping of graphene via physisorption of gas molecules affected the piezoelectric surface acoustic wave charge carrier transport in a manner different from that of electric field drift. This transport induces a doping-dependent macroscopic current. Chemical doping could be used to manipulate the majority carriers and induces unique acoustoelectric features. Moreover, the device based on such a current was demonstrated to outperform the common chemiresistor for chemicals. This study inspired potential resonator applications for chemical detection involving emerging nanomaterials.





**Fig. 12** (a) SMR sensor integrated with a graphene field-effect transistor. The bias tee separates *RF* and *DC* signals from the same port (TIE) and feeds them to the network analyzer and semiconductor parameter analyzer, respectively. (b) Experimental data for the gas mixture. (The *red circles* indicate actual concentrations, and the *green dots* indicate the values derived from the  $\Delta I/\Delta f_s$  measurements.) The *dashed lines* show the  $\Delta I$  and  $\Delta f_s$  contour lines projected on the CA-CB plane. Coordinates of each intersection point correspond to definitive (CA, CB) values

## References

- Akiyama M, Kamohara T, Kano K, Teshigahara A, Takeuchi Y, Kawahara N (2009) Enhancement of piezoelectric response in scandium aluminum nitride alloy thin films prepared by dual reactive cosputtering. *Adv Mater* 21(5):593–596
- Bhugra H, Lee S, Pan W, Pai M, Lei D (2014) Commercialization of world's first piezomems resonators for high performance timing applications. In: 2014 I.E. 27th international conference on micro electro mechanical systems (MEMS). IEEE, pp 204–205
- Cassella C, Chen G, Qian Z, Hummel G, Rinaldi M (2016) RF passive components based on aluminum nitride cross-sectional lamé-mode MEMS Resonators. *IEEE Trans Electron Devices* 64(1):237–243
- Chang Y, Lu Y, Tang N, Pang W, Zhang H, Zhang D, Duan X (2015) Fingerprint library of volatile organic compounds by monolayer functionalized film bulk acoustic resonator arrays. In: 19th international conference on miniaturized systems for chemistry and life sciences, MicroTAS 2015, 25 Oct, 2015 – 29 Oct, 2015, Gyeongju. Chemical and Biological Microsystems Society, pp 1737–1739
- Chang Y, Tang N, Qu H, Liu J, Zhang D, Zhang H, Pang W, Duan X (2016) Detection of volatile organic compounds by self-assembled monolayer coated sensor array with concentration-independent fingerprints. *Sci Rep* 6:23970
- Chen Y, Zhang H, Feng Z, Zhang H, Zhang R, Yu Y, Tao J, Zhao H, Guo W, Pang W (2016a) Chemiresistive and gravimetric dual-mode gas sensor toward target recognition and differentiation. *ACS Appl Mater Interfaces* 8(33):21742–21749
- Chen Y, Zhang H, Zhang H, Feng Z, Zhao H, Sun C, Zheng S, Pang W, Zhang D (2016b) Acoustically induced current in graphene by aluminum nitride transducers. *Appl Phys Lett* 108(3):033107
- Duan Y, Chang Y, Liang J, Jiang Y, Pang W, Duan XX, Zhang ML (2016) Wireless and passive piezoelectric microelectromechanical resonant sensors for realtime physical and biochemical

- sensing. In: 20th international conference on miniaturized systems for chemistry and life sciences, MicroTAS 2016, 9 Oct 2016–13 Oct 2016. Chemical and Biological Microsystems Society, Dublin, pp 1386–1387
- Feng Z, Chen B, Qian S, Xu L, Feng L, Yu Y, Zhang R, Chen J, Li Q, Li Q (2016) Chemical sensing by band modulation of a black phosphorus/molybdenum diselenide van der Waals heterostructure. *2D Mater* 3(3):035021
- Giovannini M, Yazici S, Kuo N-K, Piazza G (2014) Apodization technique for spurious mode suppression in AlN contour-mode resonators. *Sensors Actuators A Phys* 206:42–50
- Gong S, Piazza G (2014) Monolithic multi-frequency wideband RF filters using two-port laterally vibrating lithium niobate MEMS resonators. *J Microelectromech Syst* 23(5): 1188–1197
- Harrington B, Abdolvand R (2011) In-plane acoustic reflectors for reducing effective anchor loss in lateral? Extensional MEMS resonators. *J Micromech Microeng* 21(8):085021
- Huang J, Zhang M, Cui W, Duan Y, Zheng S, Duan X, Pang W, Zhang D, Zhang HA (2015) High-throughput micro TAS integrating digital microfluidic driver and on-chip biosensor array for massively parallel detection. In: 19th international conference on miniaturized systems for chemistry and life sciences, MicroTAS 2015, 25 Oct 2015–29 Oct 2015. Chemical and Biological Microsystems Society, Gyeongju, pp 1220–1222
- Kim B, Olsson RH, Wojciechowski KE (2013) AlN microresonator-based filters with multiple bandwidths at low intermediate frequencies. *J Microelectromech Syst* 22(4):949–961
- Konno A, Sumisaka M, Teshigahara A, Kano K, Hashimo K-Y, Hirano H, Esashi M, Kadota M, Tanaka S (2013) ScAlN lamb wave resonator in GHz range released by XeF<sub>2</sub> etching. In: 2013 I.E. international ultrasonics symposium (IUS). IEEE, pp 1378–1381
- Kreuzer S, Volatier A, Fattinger G, Dumont F (2015) Full band 41 filter with high Wi-Fi rejection-design and manufacturing challenges. In: 2015 I.E. international ultrasonics symposium (IUS). IEEE, pp 1–4
- Liang J, Zhang H, Xie H, Pang W, Zhang D, Zhang H (2014) Aluminum nitride lamb wave resonators with high figure of merit for narrowband filter applications. In: 2014 I.E. international frequency control symposium (FCS). IEEE, pp 1–4
- Liang J, Zhang H, Zhang D, Duan X, Zhang H, Pang W (2015a) Design and fabrication of aluminum nitride lamb wave resonators towards high figure of merit for intermediate frequency filter applications. *J Micromech Microeng* 25(3):035016
- Liang J, Zhang H, Zhang D, Zhang H, Pang W (2015b) 500-terminated ALN MEMS filters based on lamb wave resonators. In: 2015 transducers-2015 18th international conference on solid-state sensors, actuators and microsystems (TRANSDUCERS). IEEE, pp 1973–1976
- Lin C-M, Lai Y-J, Hsu J-C, Chen Y-Y, Senesky DG, Pisano AP (2011) High-Q aluminum nitride lamb wave resonators with biconvex edges. *Appl Phys Lett* 99(14):143501
- Liu W, Wang J, Yu Y, Chang Y, Tang N, Qu H, Wang Y, Pang W, Zhang H, Zhang D (2014) Tuning the resonant frequency of resonators using molecular surface self-assembly approach. *ACS Appl Mater Interfaces* 7(1):950–958
- Liu W, Zhang H, Zhao H, Tang Z, Wang Y, Sun C, Pang W, Duan X (2017) Comparative analysis of static and non-static assays for biochemical sensing using on-chip integrated field effect transistors and solidly mounted resonators. *Sensors Actuators B Chem* 243:775–783
- Lu Y, Chang Y, Tang N, Qu H, Liu J, Pang W, Zhang H, Zhang D, Duan X (2015) Detection of volatile organic compounds using microfabricated resonator array functionalized with supramolecular monolayers. *ACS Appl Mater Interfaces* 7(32):17893–17903
- Matloub R, Artieda A, Sandu C, Milyutin E, Muralt P (2011) Electromechanical properties of AlO<sub>2</sub>ScO<sub>3</sub>. 1N thin films evaluated at 2.5 GHz film bulk acoustic resonators. *Appl Phys Lett* 99(9):092903
- Moreira M, Bjurström J, Katardjev I, Yantchev V (2011) Aluminum scandium nitride thin-film bulk acoustic resonators for wide band applications. *Vacuum* 86(1):23–26
- Pang W, Ruby RC, Parker R, Fisher PW, Unkrich MA, Larson JD (2008) A temperature-stable film bulk acoustic wave oscillator. *IEEE Electron Device Lett* 29(4):315–318



- Pang W, Yan L, Zhang H, Yu H, Kim ES, Tang WC (2006) Femtogram mass sensing platform based on lateral extensional mode piezoelectric resonator. *Appl Phys Lett* 88(24):243503–243503
- Pang W, Zhao H, Kim ES, Zhang H, Yu H, Hu X (2012) Piezoelectric microelectromechanical resonant sensors for chemical and biological detection. *Lab Chip* 12(1):29–44
- Paprotny I, Doering F, Solomon PA, White RM, Gundel LA (2013) Microfabricated air-microfluidic sensor for personal monitoring of airborne particulate matter: design, fabrication, and experimental results. *Sensors Actuators A Phys* 201:506–516
- Qi M, Zhang D, Pang W, Zhang H (2012) High performance TD-SCDMA band-pass filter based on film bulk acoustic resonator technology. In: 2012 Asia Pacific microwave conference proceedings. IEEE, pp 547–549
- Rai S, Su Y, Pang W, Ruby R, Otis B (2010) A digitally compensated 1.5 GHz CMOS/FBAR frequency reference. *IEEE Trans Ultrason Ferroelectr Freq Control* 57(3):552–561
- Rinaldi M, Zuo C, Van der Spiegel J, Piazza G (2011) Reconfigurable CMOS oscillator based on multifrequency AlN contour-mode MEMS resonators. *IEEE Trans Electron Devices* 58(5):1281–1286
- Schedin F, Geim A, Morozov S, Hill E, Blake P, Katsnelson M, Novoselov K (2007) Detection of individual gas molecules adsorbed on graphene. *Nat Mater* 6(9):652–655
- Segovia-Fernandez J, Cremonesi M, Cassella C, Frangi A, Piazza G (2015) Anchor losses in AlN contour mode resonators. *J Microelectromech Syst* 24(2):265–275
- Shealy JB, Hodge MD, Patel P, Vetry R, Feldman AY, Gibb SR, Boomgarden MD, Lewis MP, Shealy JB, Shealy JR (2016a) Single crystal AlGaN bulk acoustic wave resonators on silicon substrates with high electromechanical coupling. In: Radio frequency integrated circuits symposium (RFIC). IEEE, pp 103–106
- Shealy JB, Shealy JB, Patel P, Hodge MD, Vetry R, Shealy JR (2016b) Single crystal aluminum nitride film bulk acoustic resonators. In: 2016 I.E. radio and wireless symposium (RWS). IEEE, pp 16–19
- Sun Y, Yang Q, Zhang H, Zhang D, Pang W A (2015) Compact dual-passband filter based on bulk acoustic wave technology. In: 2015 I.E. MTT-S international microwave symposium. IEEE, pp 1–4
- Tabrizian R, Casinovi G, Ayazi F (2013) Temperature-stable silicon oxide (SiO<sub>x</sub>) micromechanical resonators. *IEEE Trans Electron Devices* 60(8):2656–2663
- Tasnadi F, Alling B, Höglund C, Wingqvist G, Birch J, Hultman L, Abrikosov IA (2010) Origin of the anomalous piezoelectric response in wurtzite Sc<sub>x</sub>Al<sub>1-x</sub>N alloys. *Phys Rev Lett* 104(13):137601
- Umeda K, Kawai H, Honda A, Akiyama M, Kato T, Fukura T (2013) Piezoelectric properties of ScAlN thin films for piezo-MEMS devices. In: IEEE 26th international conference on micro electro mechanical systems (MEMS). IEEE, pp 733–736
- Wang R, Bhavne SA, Zhgoon S, Bhattacharjee K (2016) Multi-frequency LiNbO<sub>3</sub> lamb wave resonators with < 3?? impedance. In: 2016 I.E. 29th international conference on micro electro mechanical systems (MEMS). IEEE, pp 679–682
- Warder P, Link A (2015) Golden age for filter design: innovative and proven approaches for acoustic filter, duplexer, and multiplexer design. *IEEE Microw Mag* 16(7):60–72
- Yang Q, Pang W, Zhang D, Zhang H (2016) A modified lattice configuration design for compact wideband bulk acoustic wave filter applications. *Micromachines* 7(8):133
- Ye C, Hemi Q, Xuexin D, Luye M, Reed MA VOC (2016) Detection using multimode E-nose composed of bulk acoustic wave resonator and silicon nanowire field effect transistor array. In: Proceedings of 2016 I.E. Sensors, 30 Oct–3 Nov 2016. IEEE, Piscataway, p 3. <https://doi.org/10.1109/ICSENS.2016.7808606>
- Yokoyama T, Iwazaki Y, Onda Y, Nishihara T, Sasajima Y, Ueda M (2015) Highly piezoelectric codoped AlN thin films for wideband FBAR applications. *IEEE Trans Ultrason Ferroelectr Freq Control* 62(6):1007–1015
- Yuan Z, Qingrui Y, Ye C, Rui Z, Jin T, Hemi Q, Xuexin D (2016) Detection of volatile organic compounds by high-Q piezotransduced single-crystal silicon bulk acoustic resonator arrays. In:

- Proceedings of 2016 I.E. Sensors, 30 Oct–3 Nov 2016. IEEE, Piscataway, p 3 <https://doi.org/10.1109/ICSENS.2016.7808661>
- Zhang H, Kim ES (2005) Micromachined acoustic resonant mass sensor. *J Microelectromech Syst* 14(4):699–706
- Zhang H, Marma MS, Bahl SK, Kim ES, McKenna CE (2007) Sequence specific label-free DNA sensing using film-bulk-acoustic-resonators. *IEEE Sensors J* 7(12):1587–1588
- Zhang H, Pang W, Marma MS, Lee C-Y, Kamal-Bahl S, Kim ES, McKenna CE (2010) Label-free detection of protein-ligand interactions in real time using micromachined bulk acoustic resonators. *Appl Phys Lett* 96(12):123702
- Zuo C, Sinha N, Piazza G (2010) Very high frequency channel-select MEMS filters based on self-coupled piezoelectric AlN contour-mode resonators. *Sensors Actuators A Phys* 160(1):132–140
- Zhang H, Pang W (2011) A novel single-ended to balanced bulk acoustic wave filter for wireless communications. *IEEE Microwave Wireless Compon Lett* 21(7):347–349
- Zhang H, Pang W, Kim ES (2011) Miniature high-frequency longitudinal wave mass sensors in liquid. *IEEE Trans Ultrason Ferroelectr Freq Control* 58(1):255–258
- Zuo C, Yun CH, Stephanou PJ, Park S-J, Lo C-ST, Mikulka R, Lan J-HJ, Velez MF, Shenoy RV, Kim J (2012) Cross-sectional dilation mode resonator with very high electromechanical coupling up to 10% using AlN. In: 2012 I.E. international frequency control symposium proceedings. IEEE, pp 1–4
- Zhang Z, Lu Y, Pang W, Zhang D, Zhang H A (2013) High performance C-band FBAR filter. In: 2013 Asia-Pacific microwave conference proceedings (APMC). IEEE, pp 923–926
- Zhang M, Cui W, Liang J, Zhang D, Pang W, Zhang H A (2014a) single-chip biosensing platform integrating FBAR sensor with digital microfluidic device. In: 2014 I.E. international ultrasonics symposium. IEEE, pp 1521–1523
- Zhang M, Cui W, Zhang D, Pang W, Zhang H (2014b) Response signal enhancement of film bulk acoustic resonator mass sensor with bounded hydrophobic Teflon film. In: 2014 I.E. international frequency control symposium (FCS). IEEE, pp 1–4
- Zhang M, Cui W, Chen X, Wang C, Pang W, Duan X, Zhang D, Zhang H (2015a) Monolithic integrated system with an electrowetting-on-dielectric actuator and a film-bulk-acoustic-resonator sensor. *J Micromech Microeng* 25(2):025002
- Zhang M, Huang J, Cui W, Pang W, Zhang H, Zhang D, Duan X (2015b) Kinetic studies of microfabricated biosensors using local adsorption strategy. *Biosens Bioelectron* 74:8–15
- Zhang H, Liang J, Zhang H, Zhang D, Pang W (2015c) Spurious-free lamb wave resonators with protrusion structures. *Appl Phys Lett* 107(24):243502
- Zhang H, Liang J, Zhou X, Zhang H, Zhang D, Pang W (2015d) Transverse mode spurious resonance suppression in lamb wave MEMS resonators: theory, modeling, and experiment. *IEEE Trans Electron Devices* 62(9):3034–3041
- Zhao H, Guo X, Wang Y, Duan X, Qu H, Zhang H, Zhang D, Pang W (2016) Microchip based electrochemical-piezoelectric integrated multi-mode sensing system for continuous glucose monitoring. *Sensors Actuators B Chem* 223:83–88
- Zheng S, Zhang H, Feng Z, Yu Y, Zhang R, Sun C, Liu J, Duan X, Pang W, Zhang D (2016) Acoustic charge transport induced by the surface acoustic wave in chemical doped graphene. *Appl Phys Lett* 109(18):183110
- Zhu Y, Wang N, Sun C, Merugu S, Singh N, Gu Y (2016) A high coupling coefficient 2.3-GHz AlN resonator for high band LTE filtering application. *IEEE Electron Device Lett* 37(10):1344–1346



# Microsensors and Systems for Water Quality Determination

Shanhong Xia, Jianhua Tong, Chao Bian, Jizhou Sun, and Yang Li

## Contents

Introduction .....	848
Sensing Elements .....	850
Phosphate Microelectrodes .....	850
Nitrate Microelectrodes .....	855
Heavy Metal Ion Microelectrodes .....	862
Pre-treatment Units .....	870
Micro Digestion Reactors .....	870
Pre-concentration .....	883
Integrated System .....	889
Integrated Microchip System .....	889
Buoy-Based Multiparameter Monitoring Systems .....	891
Conclusions .....	895
References .....	895

## Abstract

Water pollution is serious in many countries, and eutrophication and heavy metals are the main pollutants. Cost-effective, miniaturized and highly sensitive sensors for on-site detection and online monitoring of water quality have attracted more attentions. In this chapter, several microsensors and systems are presented for highly sensitive detection of total phosphorus (TP), total nitrogen (TN), and heavy metal ions ( $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Hg}^{2+}$ ), which are main indicators for eutrophication and heavy metals pollution. As the effective ways to improve the sensitivity and the limit of detection, several sensitivity-enrichment methods such as nano-materials modification on microelectrodes, and pretreatment

S. Xia (✉) · J. Tong · C. Bian · J. Sun · Y. Li  
Institute of Electronics, Chinese Academy of Sciences, Beijing, China  
e-mail: [shxia@mail.ie.ac.cn](mailto:shxia@mail.ie.ac.cn)

processes such as thermally assisted ultraviolet digestion and ionic-liquid based preconcentration are also introduced. The electrochemical microsensors were fabricated by MEMS-based bulk fabrication process, and integrated with microfluidic units to realize a continuous monitoring of trace pollution targets. A buoy-based automated analytical system was developed by integrating central controller module, sequential injection module, and electrochemical analytical module together, and it was applied to online monitoring of water quality in fresh waters.

---

**Keywords**

Determination · Electrochemical · Microelectrode · Microsensor · MEMS · System · Water Quality

---

## Introduction

Over the past few decades, environmental pollution has become more serious in many countries, leading to ill health, death, and disabilities of millions of people annually along with the rapid industrialization and urbanization. Environmental pollution consists of five basic types of pollution, namely, water, air, soil, noise, and light (UNEP (United Nations Environment Programme) Collaborating Centre on Energy and Environment 2006). Among them, water pollution is one of the most serious ecological threats we face today since water is essential to human, plant, and other animal life, which is why most people always live near a water source. There are many types of water pollution such as eutrophication, heavy metals, oxygen depleting, and other toxic substances occurring in surface water, ground water, and seawater as well. Therefore, the quality of water has become a vitally important issue of concern not only for the health of human being, but also for the development of the global economy and society.

Water quality determination, as an important and effective approach to optimize water resource allocation and control the water pollution, has drawn great public attention around the world. A number of indicators are used to assess water quality in general. Some simple indicators, such as the temperature, pH, dissolved oxygen (DO), oxidation-reduction potential (ORP), conductivity, and turbidity, are pretreatment free in the detection process. Therefore, the sensors, also called electrodes, can get the value of the indicators directly just after putting the electrodes into the water, and most of these products are integrated probes on which many electrodes are integrated together according to the requirement of applications. However, for the indicators which require pretreatment process, such as total phosphorus (TP), total nitrogen (TN), and heavy metal ions ( $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Hg}^{2+}$ ,  $\text{Cd}^{2+}$ , and  $\text{Cr}^{6+}$ ), it is complicated or impossible to read the value of the concentration just after putting the sensors into the water.

Total phosphorus (TP, defined as the measure of orthophosphate, condensed phosphates (pyro-, meta-, and polyphosphates), and organically bound phosphorus) and total nitrogen (TN, defined as the sum of nitrate-nitrogen ( $\text{NO}_3\text{-N}$ ), nitrite-nitrogen ( $\text{NO}_2\text{-N}$ ), ammonia-nitrogen ( $\text{NH}_3\text{-N}$ ), and organically bonded nitrogen)

are two key parameters of water quality. Excessive TP and TN may present the phenomenon of eutrophication, bursts algae bloom, and reduces dissolved oxygen in lakes and reservoirs, which causes both environmental and physiological problems (Galloway et al. 2008; Wilson et al. 2011). As reported by Ministry of Environmental protection of the People's Republic of China (Ministry of Environmental protection of the People's Republic of China 2012), TP is one of the three main pollution indicators in Chinese surface water environment. According to European Union Environment Council standards, over 0.1 mg/l of phosphorus in water may cause over growth of algae. Thus, the determinations and monitoring of TP and TN are very important for water quality control, especially for inland water like lakes and reservoirs. The determinations of TP and TN usually consist of two processes. Firstly, a digestion process is required to convert all forms of phosphorus and nitrogen to orthophosphate and nitrate, and then, the concentration of TP and TN can be obtained by detecting the concentration of phosphate and nitrate by spectrometry or electrochemical methods.

Among the heavy metal ions, mercury ( $\text{Hg}^{2+}$ ) is considered to be a highly toxic substance that can cause kidney injury, central nervous system disorders, intellectual deterioration, and death (Taylor et al. 2006). Because of its bioaccumulation and biomagnification abilities, mercury at a trace level in aquatic system is harmful to environment and human (Leopold et al. 2009). Strict legislation limits of its concentration are imposed in many countries and generally below 0.1 ppb for surface water, and the China's surface water standard of class I ~ II is 50 ng/L. Inductively Coupled Plasma Mass Spectrometry (ICP-MS) is an analytical technique used for elemental determinations with a limit of detection of 1 ng Hg/L (EGA 2007), as well as the Cold Vapor Atomic Absorption Spectroscopy (CVAAS), as one of the primary techniques for mercury analysis, has a limit of detection of 0.01  $\mu\text{g}$  Hg/L. Although these two methods have a low detection limit, their equipment are all very big in size, which are more suitable for the analysis in laboratory. Anodic Stripping voltammetry (ASV) is an effective way for quantitative determination of heavy metal ions at low concentration level. It takes advantages of simplicity, high sensitivity, minimal space and power requirements, and low instrumentation cost. The reported limit of detection by using bare gold electrodes was around 0.40  $\mu\text{g}$ /L (Bonfil and Brand 2000; Giacomino et al. 2008), which still does not meet the requirement for water environment monitoring.

In this chapter, several electrochemical determination methods are presented for the detection of complex parameters such as total phosphorus (TP), total nitrogen (TN), and heavy metal ions ( $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Hg}^{2+}$ ). The related sensitivity-enrichment methods, pretreatment methods such as preconcentration and digestion are also demonstrated. The fabrication process of these microsensors and pretreatment units are presented, and the fabricated devices were used to detect the real water samples. An integrated microfluidic system, in which a digestion unit and a phosphate detection unit were integrated in a single chip, was developed for fully functional detection of TP. A buoy-based automated analytical system was developed for online monitoring of water quality in fresh waters, and it was tested in Wuxi Tai Lake for several months.

## Sensing Elements

Sensing Element is the device that receives signals and stimulus, and converts them to electrical signals (Wikipedia 2017). For the detection of total phosphorus (TP), total nitrogen (TN), and heavy metal ions, the sensing elements are phosphate microelectrode, nitrate microelectrode, and heavy metal ion microelectrode, respectively.

### Phosphate Microelectrodes

Many phosphate detection methods have been developed over the past few decades, including chromatographic method (Ao and Zhang 2009; Yokoyama et al. 2009), electrochemical methods (Sten 1998), and spectrophotometric method (Bulatov et al. 2009; Shyla and Nagendrappa 2011), etc. The conventional approach for phosphate determination is spectrophotometric method, which has been developed and improved over decades and have been widely used among chemists. It is considered to be a highly accurate determination method that is also very sensitive and therefore extremely precise (Mavrodineanu et al. 1973). This method is very convenient for use in laboratory experiments because it is an inexpensive and relatively simple process; however, big analysis equipment such as spectrophotometer are required in the analysis process, then this method is not suitable for online or in situ determination of phosphate. Electrochemical method uses electrodes, thus is suitable for the requirement of miniaturization and distributed sensor system in the development of in situ monitoring of TP in surface water. Xiao et al. introduced cobalt metal as phosphate sensing material in 1995 (Xiao et al. 1995). Polished cobalt wire was used as the phosphate sensitive electrode, and the potentiometric response towards phosphate in the concentration ranging from  $10^{-5}$  to  $10^{-2}$  mol/L. Meruva and Meyerhoff (1996), in 1996, explained the response mechanism of cobalt electrodes to phosphate by a theory of mixed potential resulting from the slow oxidation of cobalt. This mixed potential response mechanism is more acceptable and it is referred by most of the researchers. Recently, cobalt-based microelectrodes were developed for their numerous advantages by integrating on-chip working electrodes with reference electrodes and micro fluidic channels for phosphate detection. Zhiwei Zou et al. developed a disposable on-chip phosphate sensor with planar cobalt microelectrodes on polymer substrate (Zou et al. 2007). Jin-Hwan Lee et al. (2009) presented a cobalt-coated needle-type microelectrode array sensor (Lee et al. 2009). These two researches realized cobalt microelectrode by MEMS method, and cobalt was electroplated on Au substrate. Another electrochemical method for phosphate detection is based on the reduction current of molybdophosphate on inert electrodes, such as modified glassy carbon electrode (Berchmans et al. 2011), carbon paste electrode (Lu et al. 1999), and gold microelectrode (Justyna et al. 2011, 2013). This method has good selectivity, and the electrodes need no modification, which makes the fabrication and preservation of electrodes quite easy. However, the low current response of this method limits its application. Considering the current is

limited by chemical reaction rate instead of diffusion rate (Carpenter et al. 1997), it is hard to effectively raise the current response and sensitivity simply by increasing electrode's surface area (Bai et al. 2014a).

### Cobalt-Based Microelectrode

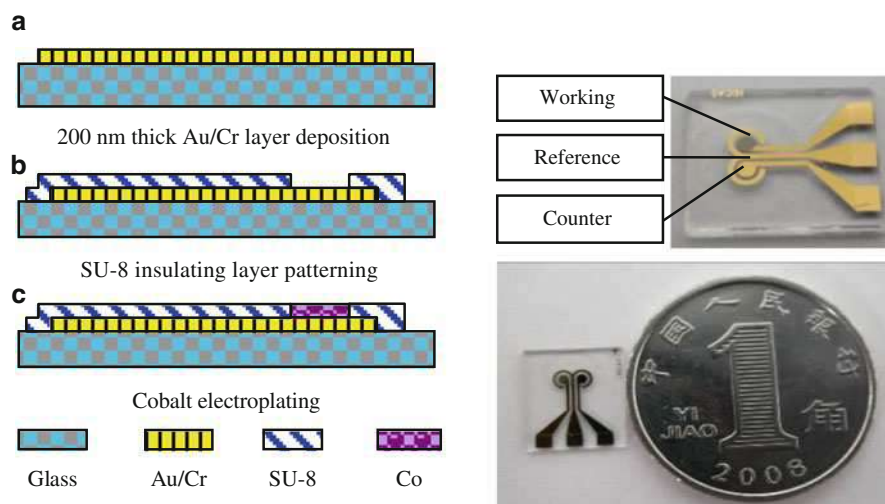
The previously reported cobalt electrodes could not meet the detection need for phosphate concentration lower than 0.02 mg P/L, which is the standard of class I water according to environmental quality standards for surface water (GB3838-2002) (Ministry of Environmental Protection of the People's Republic of China 2002) in China.

Xia group (Bai et al. 2012) developed a cobalt-based microelectrode for phosphate detection. This microelectrode was fabricated by MEMS technique, and a cobalt film was electrodeposited on the surface of the microelectrode as the sensing material. The fabrication process is shown in Fig. 1a.

A layer of 30 nm/200 nm Cr/Au layer is deposited and patterned on a glass substrate by lift-off process. The Cr layer is used as a seed layer and the Au layer is for electrical connection. After covered by SU-8 negative photo resist as insulating layer, electrode area was discovered for cobalt electroplating. Photographs of the fabricated microelectrode are shown in Fig. 1b.

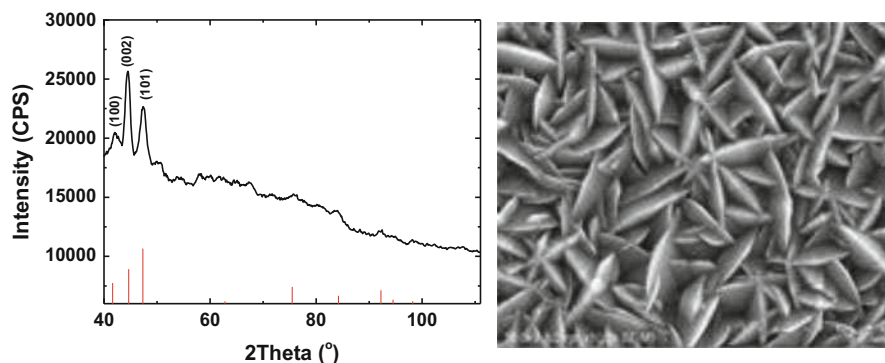
The cobalt electrodes were examined by XRD, and SEM, as shown in Fig. 2. This photo shows that the cobalt layer obtained in previous described electroplating conditions mainly grew in (002) crystal direction, and it has a uniform spam-like nanostructure ( $70 \text{ nm} \times 2 \mu\text{m}$ ).

Calibration and measurements were carried out with a two-electrode-system including an SCE (saturated KCl) electrode as reference electrode, and open circuit



**Fig. 1** (a) Brief fabrication process of cobalt microelectrode and (b) the photograph of the fabricated device





**Fig. 2** XRD examination and SEM image of electrodeposited cobalt film

potential was measured as the response. The electrodes were first flushed by DI water for 2 min, and then dipped in buffer solution for 400 s to complete a simple conditioning process before calibration. For each measurement, after signals became stable in about 200 s, a washing step was performed using DI water, and then the next standard solution was applied.

Electrochemical character of cobalt electrodes was examined, and over 26 chips were tested. The calibration curve with standard deviation bar is shown in Fig. 3a. Reasonably low chip-to-chip deviation has been obtained by measuring  $\text{KH}_2\text{PO}_4$  at  $10^{-4}$  M on 14 different sensors with variances of 2.34% relative standard deviation (R.S.D) as shown in Fig. 3b.

The linear fitting equation of the calibration curve is

$$y = 36.727x + 701.81 \quad (1)$$

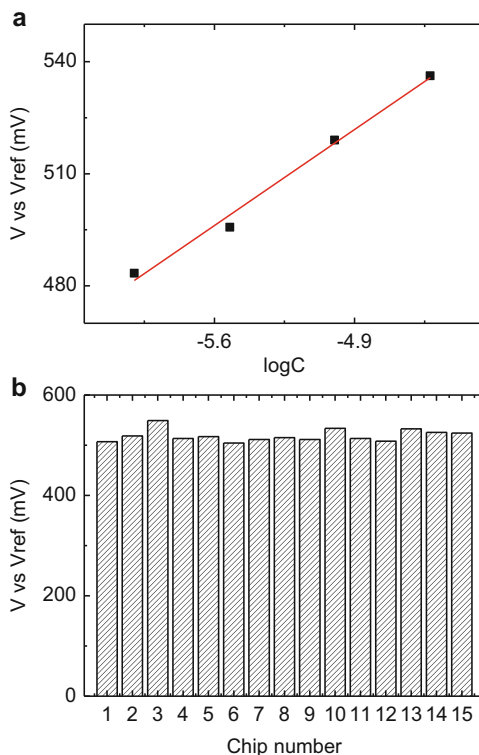
and linear fit error is  $R^2 = 0.9908$ . The response voltage of digested water sample was 490.2 mV. Calculating from Eq. 1, the total phosphorus concentration should be  $1.7 \times 10^{-6}$  mol/L (0.0527 mg P/L). The results using different testing methods are listed in Table 1.

### Microelectrode Based on Reduction of Molybdophosphate Complex

Molybdate spectrophotometry has already been widely used and set as the standard detection method in China (Ministry of Environmental Protection of the People's Republic of China 2002). However, it suffers from interference of turbidity and operational complexity. The electrochemical method based on the reduction of molybdophosphate complex for phosphate detection has been demonstrated in an admirable way for its good selectivity, simple fabrication process, low detection limit, and easiness of combination with flow injection systems (Jońca et al. 2011; Udnan et al. 2005). Several studies using conventional gold electrodes have already been published; Xia group (Bai et al. 2014a) developed a gold microelectrode by microfabrication technique, which could easily realize the batch fabrication of micro devices. The fabrication process and the image of the microelectrode are shown in



**Fig. 3** (a) Calibration curve for water sample detection (b) chip-to-chip deviation of 15 different electrodes in measuring  $10^{-4}$  mol/L  $\text{KH}_2\text{PO}_4$  at pH 7.5

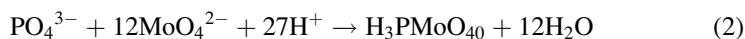


**Table 1** Detection results using different methods

Detection method	Cobalt electrode	Spectrometric equipment	Qualified institute certification
Total phosphorus (mg/L)	0.0527	0.044	0.05

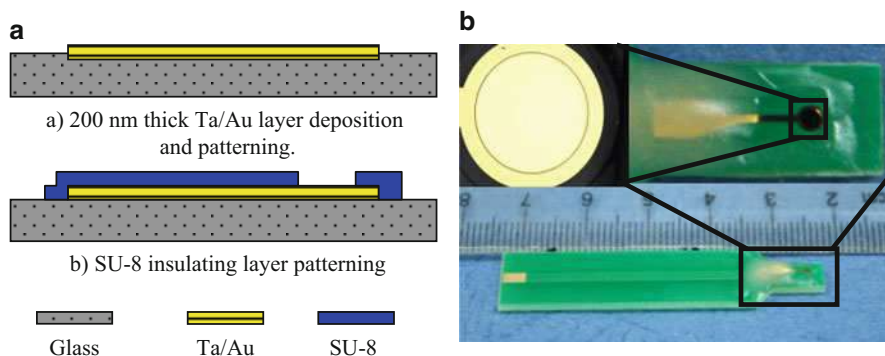
Fig. 4. The gold micro plate electrode was used as working electrode, and the concentration of phosphate was obtained by adding phosphate solution into excessive amount of molybdate to form molybdophosphate complex.

Phosphate is nonelectroactive species, thus the electrochemical determination of this ion involves the treatment of the sample with an acidic molybdate solution to form a Keggin anion,  $\text{P}(\text{Mo}_{12}\text{O}_{40})^{3-}$ , as shown in Eq. 2.



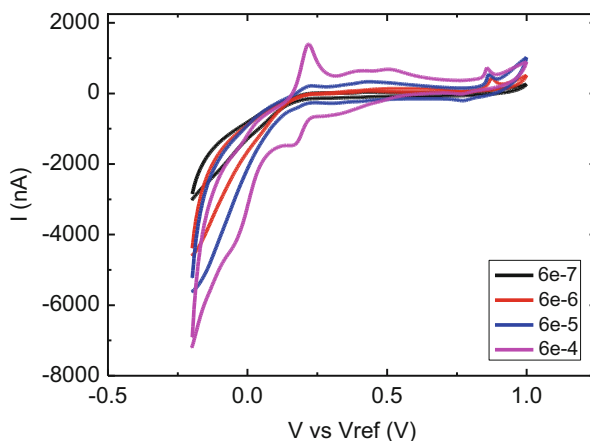
The complex shows two reduction waves and two oxidation waves on gold electrode, as shown in Fig. 5.

The peak value of reduction wave showed relevance with concentration of phosphate, but the linearity and LOD could not meet the need of TP detection. To enhance sensitivity, amperometry was selected for the detection of phosphate.



**Fig. 4** (a) Fabrication process and (b) photos of gold microelectrode

**Fig. 5** Electrochemical reduction of molybdophosphate complex

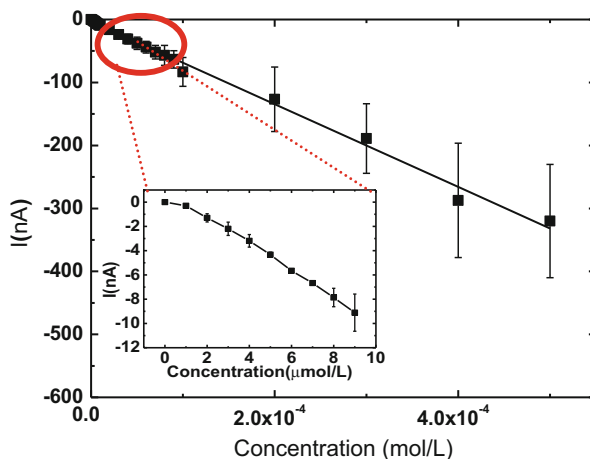


To avoid the effect of background current, the voltage was chosen as 0.27 V, which was at the valley of the reduction wave.

As reported by Neil G. Carpenter et al. (1997), the amperometric response was not proportional to the electrode radius, which confirms that the rate of molybdophosphate reduction is mainly controlled by the preceding chemical reaction. However, the steady state currents are always proportional to the concentration of phosphate added into the acidic molybdate sodium solution in an appropriate range. So the steady state currents at gold microelectrode could be the basis of an analytical method of phosphate in water.

The diluted TP standard samples were digested using the digesting unit of 5B-2P total phosphorus detection equipment manufactured by Lian-hua Tech. Co., Ltd. Eight milliliter standard samples were first mixed with 1 ml digesting solution in sealable reaction tubes and shook well. The mixtures were then put into digest equipment and heated at 120 °C for 30 min. The digested samples were cooled to room temperature in cool water bath, and molybdate sodium was added into the

**Fig. 6** Calibration curve of the microelectrode in the range of 1~500  $\mu\text{mol/L}$



**Table 2** Detection results of TP standard solutions

Nominal value (mg/L P)	Measurement result (mg/L P)	Recovery rate
0.02	0.01826	91.3%
0.2	0.186	93%
0.3	0.284	94.6%

samples to get the concentration of 20  $\text{mmol/L}$ . The mixed samples were then measured with gold microelectrodes.

The current response increases as phosphate was added into the molybdate sodium solution. The response time was less than 20 s under stirring, forming obvious current steps for different concentrations.

The calibration curve in Fig. 6 shows that the microelectrode has the sensitivity of  $-0.657 \text{ nA per } \mu\text{mol/L}$  ( $R^2 = 0.994$ ), a wide linear range of 1–500  $\mu\text{mol/L}$ , and a low LOD of 0.66  $\mu\text{mol/L}$ .

Referring to environmental quality standards for surface water in China (Ministry of Environmental Protection of the People's Republic of China 2002), most water samples got from land surface contain phosphorus less than 0.4 mg/L. Three different concentrations of TP standard samples were tested, and results are listed in Table 2. The measurement results show good consistency with their nominal values, which indicates that the digesting process has little influence to the detection, and this method could be used for TP detection in water.

## Nitrate Microelectrodes

Various analytical techniques, including ultraviolet (UV) spectrometry (Sempere et al. 1993), high-performance liquid chromatography (Akyuz and Ata 2009),

ion-exchange chromatography (Helaleh and Korenaga 2000), Raman spectroscopy (Kim et al. 2012), colorimetry (Raimbault et al. 1990), capillary electrophoresis (Jimidar et al. 1995), and chemiluminescence (Kodamatani et al. 2009), have been developed to quantitate nitrate ions in liquid samples. These techniques have higher detection precision and lower detection limit, but most of their analytical processes also reveal several drawbacks such as low sample throughput, usage of toxic reagents, time-consuming procedures, and expensive instruments with large volume (Noroozifar et al. 2007). Therefore, the development of simple, sensitive, and reliable methods for nitrate detection in liquid samples is highly demanded by departments of healthcare and environmental protection.

Based on advantages of high sensitivity, rapid response, low power consumption, cheap instrumentation, favorable portability, and suitability for automation, electrochemical method is considered as a perfect alternative for nitrate concentration measurement. Great efforts have been invested into the research and development of electrochemical electrodes for nitrate detection (Isoda et al. 2007; Yu et al. 2015). But due to the slow kinetics of the charge transfer, bare unmodified electrodes are usually characterized by low sensitivity and irreproducibility in nitrate determination. To improve the selectivity and sensitivity in the process of quantifying nitrate, researchers intend to modify the electrode with sensitive materials prior to electrochemical analysis. Since then a wide variety of materials have been investigated, and according to their properties most of them can be classified into three categories: metallic electrocatalyst, organic conducting polymer, and biocatalyst. Reactive metallic electrocatalysts, such as copper (Silva et al. 2013), silver (Dohyun et al. 2007), palladium (Wang et al. 2006a), platinum (Dima et al. 2005), can catalyze the electrochemical reduction of nitrate on working-electrode surface to generate current response. Doped-polypyrrole, as one of the most important and commonly used organic conducting polymers, can be employed as sensitive materials in both voltammetric and potentiometric methods for nitrate detection (Aravamudhan and Bhansali 2008; Zhang and Tian 2006). Biocatalysts, such as nitrate reductase enzyme immobilized on the electrode surface, are getting more attentions in nitrate determination (Silva et al. 2004). Among these sensitive materials, doped-polypyrrole is expensive and electrodes modified with doped-polypyrrole are usually characterized by low stability and sensitivity towards trace nitrate. Nitrate reductase enzymes are always showing good sensitivity and selectivity towards nitrate, but it is too sophisticated to be field-deployable. Correspondingly, metallic catalysts for nitrate reduction can not only offer marked improvements in selectivity and sensitivity in the determination process but also be fabricated easily and reproducibly. So this kind of electrocatalyst has been widely employed in electrochemical sensors for nitrate detection. As a nonnoble metal catalyst, copper has been demonstrated that it is a reactive and selective electrocatalyst for direct nitrate reduction in several literatures (Dohyun et al. 2007; Wang et al. 2006a). Meanwhile, microelectrodes with geometrical diameters ranging from micrometers to submicrometers have many advantages in electrochemical analysis, such as reduced ohmic drop and charging currents, small time constant, fast diffusion rate of substances, high rate of electron transfer, and increased current densities allowing lower detection limits and greater

sensitivities to be achieved in comparison to macroelectrodes (Stortini et al. 2015; Yang et al. 2011).

So this section herein mainly presents the characterization of a microsensor modified by copper nanoparticles for nitrate sensing in liquid samples. First, a microelectrode was fabricated by microfabrication technology. Copper nanoparticles, as the electrocatalyst material, were electrodeposited onto the working surface of the microelectrode to construct an electrochemical sensor for nitrate detection. Second, the developed sensor was employed to implement recognition and quantification for nitrate ions by linear sweep voltammetry under acidic electrolytes. Finally, concentrations of two real water samples were detected using the developed sensor. The detection results were compared with the data obtained by standard ultraviolet (UV) spectrophotometric method for nitrate measurement.

## **Fabrication and Modification of Microelectrode Chip**

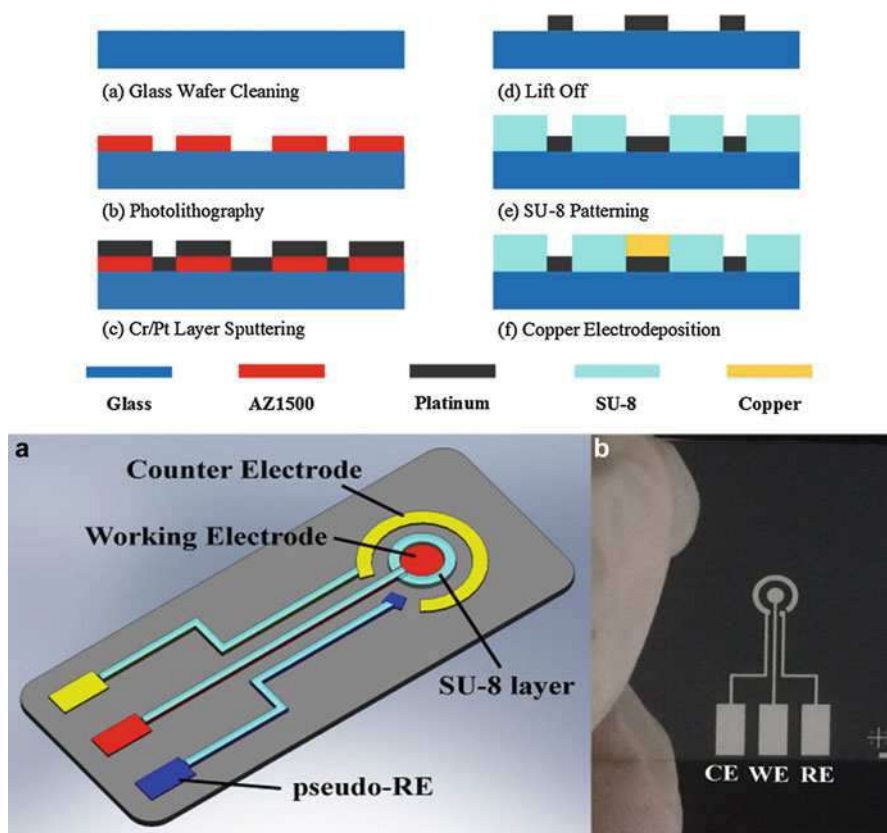
### **Microelectrode Fabrication**

The microelectrode chip was fabricated by standard microfabrication process. Firstly, a layer of platinum with the thickness of 200 nm is deposited and patterned on a glass substrate by photolithography and lift-off process to form the working electrode (WE) and counter electrode (CE). A layer of tantalum with the thickness of 30 nm was used as the adhesion layer material between the platinum layer and the glass substrate. Then a SU-8 layer was patterned around the platinum electrodes to form a micro pool to define the sensitive area of working electrode. Finally, the glass wafer was diced into individual chips, which were subsequently wire-bonded and encapsulated on a customized print circuit board, then ready to be used. Figure 7 shows the major fabrication process and photograph of the microelectrode chip. The sensitive area of the working electrode is 1 mm<sup>2</sup>.

### **Microelectrode Modification**

Firstly, the microelectrode chip was sonicated in acetone, ethanol, and deionized water for 180 s sequentially. Then, the microelectrode was electrochemically pretreated by cyclic voltammetry scan in 0.05 molL<sup>-1</sup> H<sub>2</sub>SO<sub>4</sub> solution until a reproducible voltammogram was obtained. Secondly, the electrodeposition of copper material was performed by chronoamperometry scan in 0.15 molL<sup>-1</sup> CuSO<sub>4</sub> (pH = 2.0). A stable potential of -500 mV was applied on the working electrode versus the Ag/AgCl reference electrode for 200 s. After the electrodeposition process, the modified microelectrode was washed by pouring deionized water on the electrode surface for 30 s, and stored in deionized water.

The surface morphology of the modification layer was examined by scanning electron microscopy (SEM). The microstructural analysis showed that the deposited layer was constructed by copper nanoparticles with an average diameter of 400 nm and the spatial distribution of these particles was very uniform, as shown in Fig. 8. There are several benefits for this kind of micromorphology. First, uniform copper nanoparticles can provide large effective surface area in the following electrochemical analysis, which can provide increased current response for nitrate ions. Second,

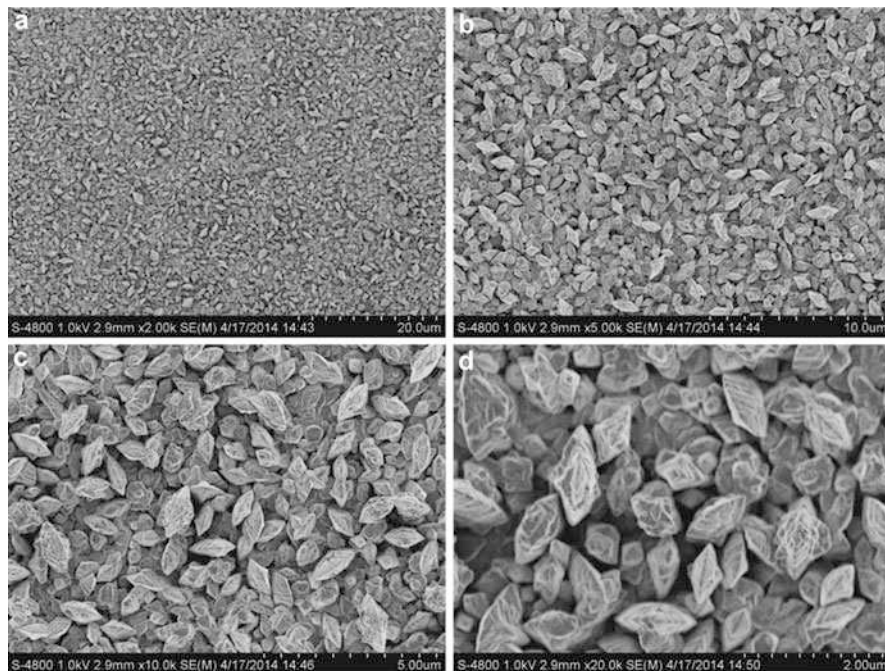


**Fig. 7** The fabricated microelectrode: (a) the fabrication process; (b) photograph of the microelectrode

the microenvironment of the fabricated nanoparticles can provide more channels for proton transfer in the detecting reactions which can facilitate the exchange of electroactive species between the copper surface and the electrolyte. So according to classical electrochemical theory, compared with compact copper layer, the fabricated copper layer herein is a more effective electrocatalyst in facilitating the related reactions.

The reaction mechanism and kinetics of copper nucleation on platinum substrate are used to explain the phenomena that unique morphologies could be achieved under this condition. Since the electrodeposition process was performed in acidic electrolyte, the reduction process of copper ion on working electrode is probably accompanied by a hydrogen evolution reaction according to the following equations:



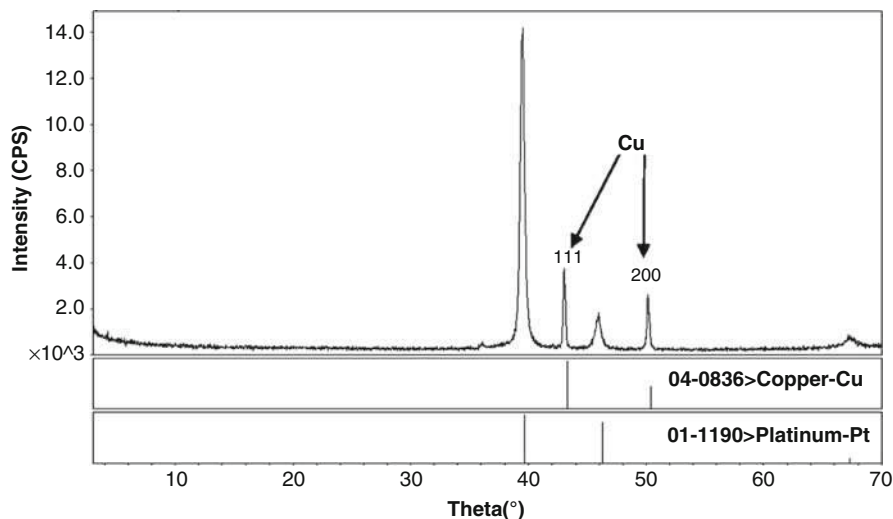


**Fig. 8** SEM micrographs of copper layer deposited by controlled-potential scan with  $-0.5$  V in  $0.15 \text{ molL}^{-1}$   $\text{CuSO}_4$  for 200 s, scan rate  $50 \text{ mV s}^{-1}$ : (a)  $\times 2,000$ ; (b)  $\times 5,000$ ; (c)  $\times 10,000$ ; (d)  $\times 20,000$

When the deposition voltage is lower than  $-300$  mV, the speed of copper nucleation on platinum surface is very slow and the deposition process was not disturbed by other reactions such as hydrogen evolution reaction. In this condition, the copper atoms were stacked compactly onto the electrode surface. The deposited copper layer tends to be flat and compact. While when the deposition voltage is higher than  $-500$  mV, the copper nucleation is speeded up, and the electrodeposition process is accompanied by hydrogen evolution reaction at the same time. The electrodeposition was performed by chronoamperometry scan, so the hydrogen evolution reaction is relatively stable and slight. The evolution of  $\text{H}_2$  with constant speed makes it convenient to form tiny hydrogen flow as a dynamic template and thus influences the morphologies of deposited copper layers (Li et al. 2011a). The deposited copper will not be stacked compactly and tends to develop along three dimensional directions, which results in porous copper particles on working electrode.

Materials deposited on the surface of working electrode were characterized by X-ray diffraction (XRD). The result is shown in Fig. 9. Four peaks of crystal directions could be clearly recognized in the spectrum, and it could demonstrate that the material deposited on platinum working-electrode was copper mainly grown in Cu (111) and Cu (200) crystallographic orientation.

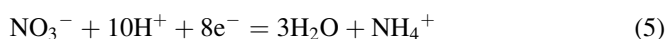




**Fig. 9** XRD spectrum for electrodeposited copper layer

### Electrochemical Measurement for Nitrate

Previous literatures have demonstrated that under acidic conditions nitrate can be reduced to ammonium ions on a copper surface according to the following equation (Ward-Jones et al. 2005):



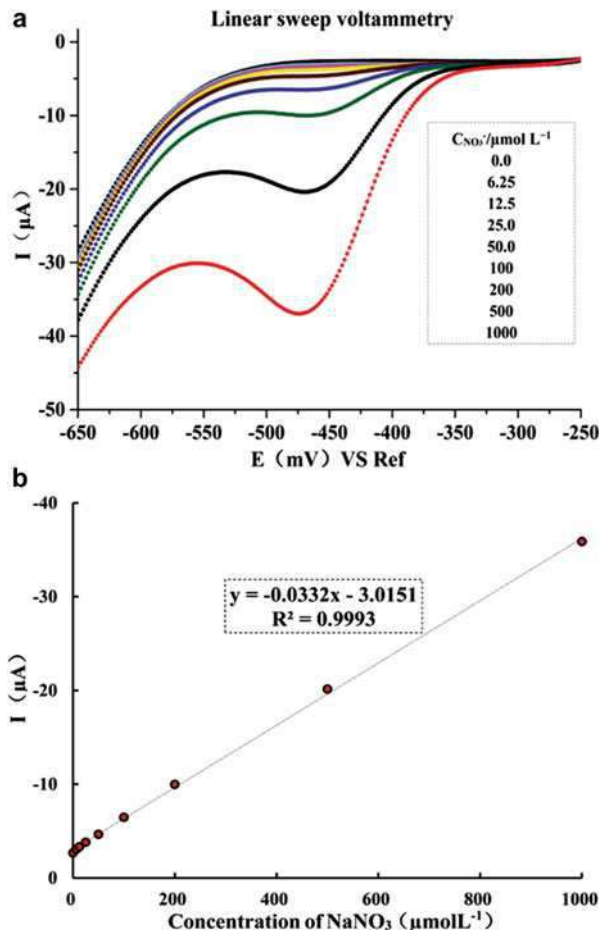
The electrocatalytic activity of the deposited copper layers was verified by quantifying the nitrate standard solutions with various concentrations based on Eq. 5. Linear sweep voltammetry was chosen for the electroanalysis of nitrate. Different concentrations of nitrate solution were prepared in  $0.1 \text{ molL}^{-1} \text{ Na}_2\text{SO}_4$  supporting buffer adjusted at pH 2.0. Linear sweep voltammograms of the fabricated sensors in electrolytes with varied nitrate concentrations ranging from 0 to  $1,000 \text{ }\mu\text{molL}^{-1}$  were recorded in Fig. 10a. There is a cathodic peak appeared around  $-470 \text{ mV}$  in the voltammograms, which is the reduction peak of nitrate ions on copper surface. It is also found that the reduction peak current was linearly increasing with the nitrate concentration in these testing solutions. The calibration plots of current values as a function of nitrate concentration in this range shows quite good linearity and yields straight lines, as shown in Fig. 10b:

$$y = -0.0332x - 3.0151 (R^2 = 0.9993). \quad (6)$$

So the amperometric responses of the calibrated sensor towards water samples can be used to calculate the corresponding nitrate concentrations based on the fitted calibration plots. The determination procedure mainly contains three steps. Firstly, three standard nitrate samples with the concentrations of 0, 200, and  $500 \text{ }\mu\text{molL}^{-1}$



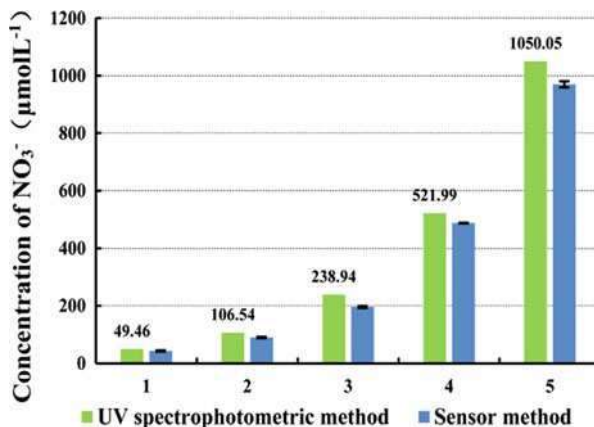
**Fig. 10** The electrocatalytic activity of the deposited copper layers: **(a)** Linear sweep voltammograms recorded in electrolytes with different  $\text{NaNO}_3$  concentration; **(b)** Corresponding calibration plots of current value and the  $\text{NO}_3^-$  concentration recorded in  $-470$  mV



were detected with the sensor to measure and record the corresponding current values. The fitted relation between nitrate concentration and amperometric response was calculated. Secondly, the sample was also detected by linear sweep voltammetry scan to measure the corresponding current response. Finally, the corresponding nitrate concentrations were derived based on the fitted relation and the amperometric response values.

Based on this procedure, the concentrations of five standard nitrate samples given by the proposed sensor and standard nitrate measurement method based on standard ultraviolet (UV) spectrophotometry were detected and compared, as shown in Fig. 11. It is found that the results given by the sensors are usually smaller than those obtained by the standard ultraviolet (UV) spectrophotometric method. But the maximal difference between these two methods was of 13.2%, which means the sensor has a good analytical performance for these standard nitrate samples. We think that the difference of the results given by these two methods is mainly

**Fig. 11** Comparison of results given by two detection methods for five standard nitrate samples with concentration of 50, 100, 200, 500, and 1,000  $\mu\text{molL}^{-1}$  (repeated number,  $N = 3$ )



related with the cumulative passivation effect on copper surface. The freshly deposited copper layer would exhibit a large and highly active surface area, which has a remarkable electrocatalytic effect for the reduction of nitrate. But along with the calibration process and continuous measurements for nitrate samples with increased concentrations, more and more reaction products would adsorb onto copper surface so that the effective surface area of the copper layer was decreased gradually, which was called cumulative passivation effects in other literatures (Gamboa et al. 2009). This phenomena would lead to a small deviation of the results given by the sensor from the data obtained by the standard method.

Two running water samples were also detected using the proposed sensor method and the UV spectrophotometric method. As shown in Table 3, the results were in good agreement with each other and the maximal difference was of 14.36%, which shows the proposed micro electrochemical sensor can be used to detect the real water samples.

## Heavy Metal Ion Microelectrodes

The heavy metal pollution in the natural waters is increasingly becoming a serious environmental problem throughout the world because of the increased use of heavy metals in industrial processes and products. Heavy metal poisoning may cause many disorders in plant and animal kingdoms, and their accumulation leads to various systemic diseases of nervous and other systems (Praline et al. 2007).

Owing to the potential danger to the environment and human, there is a growing need for monitoring of heavy metals in water environment. Many methods have been investigated for the detection of heavy metals, such as atomic absorption spectroscopy (Vijan and Sadana 1980; Xie et al. 2008), X-ray fluorescence (Lau and Ho 1993; Yamamoto et al. 1985), mass spectrometry (Fu et al. 1983; Lu et al. 1993), and electrochemical voltammetry (Lin et al. 2009; Liu et al. 1999). Electrochemical voltammetry has been proved to be a particularly suitable method

**Table 3** Experimental results using different methods for real water samples

Samples	Sensor's number	UV method (mg N L <sup>-1</sup> )	Sensor method (mg N L <sup>-1</sup> )	RSD (%)
Sample 1	1#	0.41	0.37	9.51
	2#	0.43	0.46	-7.44
	3#	0.39	0.33	14.36
Sample 2	4#	0.86	0.82	5.00
	5#	0.81	0.76	6.54
	6#	0.77	0.80	-3.51
Sample 3	7#	1.49	1.62	-8.99
	8#	1.57	1.68	-7.20
	9#	1.61	1.45	9.93

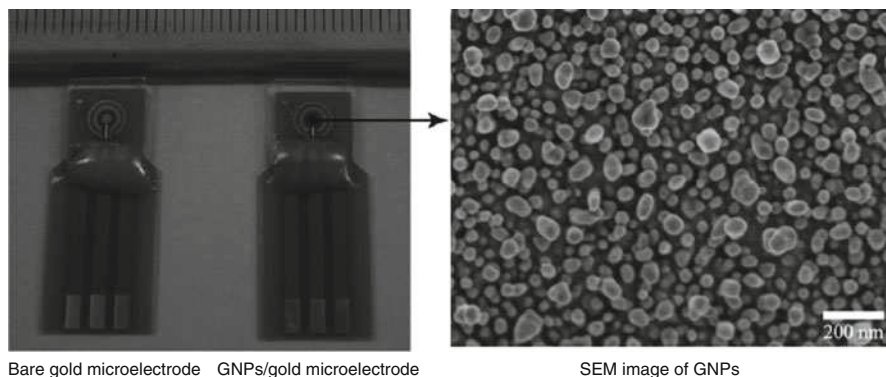
for on-site determination, due to its properties such as small in size, simple, portable, reliable, and low-cost. Electrochemical sensor offers a suitable platform for the development of a simple system for determination and monitoring of trace heavy metals in water (Jung et al. 2011; Moujahid et al. 2011).

### Environment-Friendly Electrode for Heavy Metals Detection

Simultaneous detection of heavy metals in water with simple and efficient methods is of great importance. Simultaneous detection of more than one heavy metal on mercury electrodes has been widely investigated owing to its wide cathodic potential range (Jakmunea and Junsomboon 2008; Rodrigues et al. 2011). The traditional electrodes for heavy metals detection are hanging-drop mercury electrode and mercury film electrode. The application of these electrodes was limited for the extreme toxicity of mercury and its tedious handling. New electrode materials with satisfied performance and environment-friendly merit were required. Recently, the alternative materials which are less toxic and more environmentally friendly than mercury were investigated.

As an environment-friendly electrode, the gold electrode has been used for the detection of heavy metals in a positive potential window (Giacomino et al. 2008; Gibbon-Walsh et al. 2010; Jena and Raj 2008; Salaüun and van den Berg 2006; Salaüun et al. 2007; Wang et al. 2012a), such as copper, arsenic, and mercury. Due to the overflow of hydrogen on gold electrodes, it is difficult for a gold electrode to detect heavy metals at highly negative potentials, such as Zn. According to the facts mentioned above, simultaneous detection of heavy metals on a gold electrode in a wide potential window is hard to realize. A gold microelectrode modified with GNPs and Sn film for simultaneous detection of heavy metals in both positive and negative potential range was developed by Wang et al. (2012b) in Xia group. The microelectrode chips, as shown in Fig. 12 including one gold microelectrode and two platinum microelectrodes, were fabricated by Micro Electro-Mechanical System technique. Gold microelectrode acted as the working electrode and platinum microelectrode acted as counter electrode. The area of the gold microelectrode is 1 mm<sup>2</sup>.

Gold nanoparticles (GNPs) have been extensively investigated in electrochemistry for sensor fabrication. GNPs modified electrode showed high surface area and



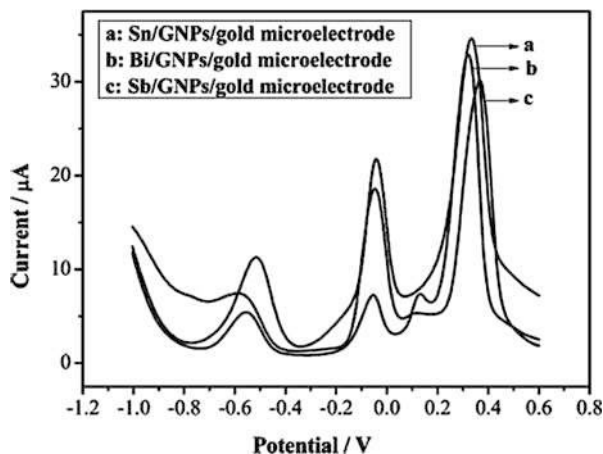
**Fig. 12** The picture of microelectrode chips (bare gold microelectrode and GNPs/gold microelectrode) and the SEM image of GNPs

fast electron transfer, both of which can largely improve the performance of sensors (Wang et al. 2006b; Xu et al. 2009; Zhang et al. 2005). GNP-modified electrodes were also applied in heavy metal detection. Bernalte et al. (2012) reported GNPs modified screen-printed carbon electrodes for anodic stripping voltammetric determination of  $\text{Hg}^{2+}$ . Hezard et al. (2012) electrodeposited GNP on glassy carbon using the method of cyclic voltammetry to detect  $\text{Hg}^{2+}$ . Compton et al. (Dai and Compton 2005; Dai et al. 2004) reported GNPs modified glassy carbon electrode for anodic stripping voltammetry of  $\text{As}^{3+}$  where the interference of  $\text{Cu}^{2+}$  to  $\text{As}^{3+}$  was reduced after the modification of GNPs. GNPs were modified on gold microelectrodes in this chapter (Zhang et al. 2005). The influence of hydrogen overflow on gold electrode was effectively reduced, owing to the shifting of the stripping potential to more positive values after the modification of GNPs. These GNPs modified gold microelectrode was applied in the simultaneous detection of heavy metals in a wide potential range (positive potential for  $\text{Cu}^{2+}$  and negative potential for  $\text{Pb}^{2+}$  and  $\text{Zn}^{2+}$ ).

Bismuth film, antimony film, and tin film have been widely reported in detection of heavy metals. In situ plating of bismuth on the screen-print carbon nanotubes electrode (Injang et al. 2010) and Nafion/2,2-bipyridyl/bismuth composite film coated glassy carbon electrode (Torma et al. 2008) were prepared for simultaneous determination of  $\text{Pb}^{2+}$ ,  $\text{Cd}^{2+}$ , and  $\text{Zn}^{2+}$ . Antimony powder-modified carbon paste electrodes (Svobodova-Tesarova et al. 2011) and macroporous antimony film electrode (Urbanová et al. 2010) were fabricated for detection of  $\text{Pb}^{2+}$  and  $\text{Cd}^{2+}$ . Antimony film microelectrode based on a single carbon fiber substrate microelectrode (Slavec et al. 2010) was fabricated for detection of  $\text{Pb}^{2+}$ ,  $\text{Cd}^{2+}$ , and  $\text{Cu}^{2+}$ . In situ plated tin film electrodes (Czop et al. 2011) were investigated for the determination of  $\text{Cd}^{2+}$  and  $\text{Zn}^{2+}$ , and the potential application in detection of  $\text{Cu}^{2+}$  and  $\text{Pb}^{2+}$  were also mentioned.

In situ plated bismuth film, antimony film, and tin film on GNPs modified gold microelectrodes were investigated and compared. Stripping voltammograms on three different microelectrodes in the solution containing  $500 \mu\text{g L}^{-1}$  of  $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ ,

**Fig. 13** SWSVs on the Sn/GNPs/gold microelectrode (curve a), Bi/GNPs/gold microelectrode (curve b), and Sb/GNPs/gold microelectrode (curve c) in solutions containing  $500 \text{ mgL}^{-1}$  of  $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ , and  $\text{Zn}^{2+}$



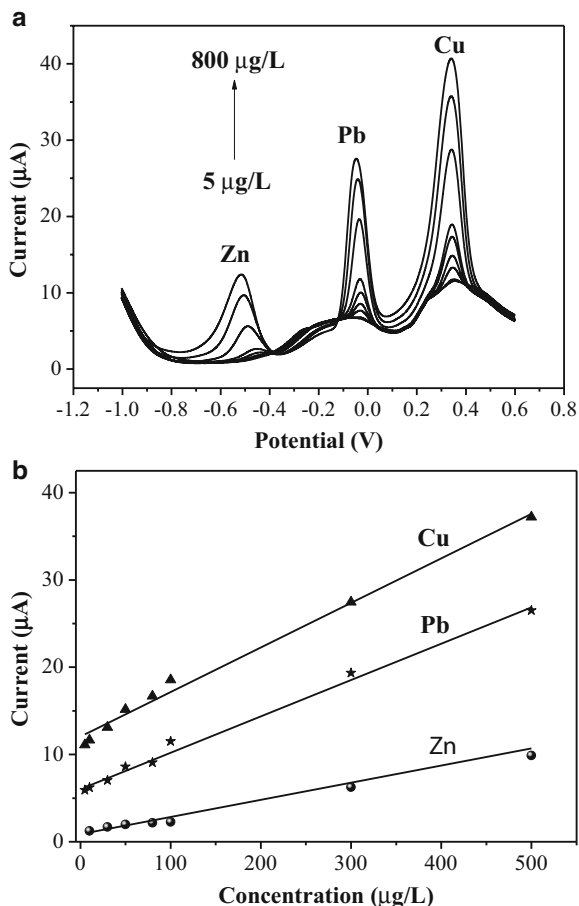
and  $\text{Zn}^{2+}$  are shown in Fig. 13. The oxidation peak potentials of  $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ , and  $\text{Zn}^{2+}$  were located at 0.30 V,  $-0.05$  V, and  $-0.50$  V, respectively. Compared with a Sb/GNPs/gold microelectrode and a Bi/GNPs/gold microelectrode, the Sn/GNPs/gold microelectrode (Fig. 13 curve a) showed the best stripping performance to  $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ , and  $\text{Zn}^{2+}$ . The Bi/GNPs/gold microelectrode (Fig. 13 curve b) had almost the same response as Sn/GNPs/gold microelectrode for  $\text{Cu}^{2+}$ , but the worst response to  $\text{Pb}^{2+}$  and  $\text{Zn}^{2+}$ . Owing to the strong Bi-Au interaction, fast and uniform Bi film deposited on the GNPs/gold microelectrode, which results in the decreasing of current efficiency for  $\text{Pb}^{2+}$  and  $\text{Zn}^{2+}$  deposition (Baldrianova et al. 2006). Compared with Sn/GNPs/gold microelectrode, Sb/GNPs/gold microelectrode (Fig. 13 curve c) had lower current response to the target metal, especially for  $\text{Zn}^{2+}$ . Because the detection solution for Sb was more acidic, the formation of hydrogen bubbles at the accumulation step blocked the reduction of the target metals onto the electrode surface. The results indicate that the tin film can accelerate the reduction of target metals onto the obtained gold electrode surface and enhance the sensitivity of the microelectrode.

As shown in Fig. 14, after accumulation at  $-1.4$  V for 300 s in acetate buffer solution ( $0.1 \text{ molL}^{-1}$ , pH 4.5), the Sn/GNPs/gold microelectrode revealed a good linear behavior in the examined concentration ranges from 5 to  $500 \text{ mgL}^{-1}$  for  $\text{Cu}^{2+}$  and  $\text{Pb}^{2+}$ , and from 10 to  $500 \text{ mgL}^{-1}$  for  $\text{Zn}^{2+}$ , with a limit of detection of  $2 \text{ mgL}^{-1}$  for  $\text{Cu}^{2+}$ ,  $3 \text{ mgL}^{-1}$  for  $\text{Pb}^{2+}$ , and  $5 \text{ mgL}^{-1}$  for  $\text{Zn}^{2+}$  ( $\text{S/N} = 3$ ). The interference of some organic species with sulfhydryl groups was reduced and the selectivity of the microelectrode was improved due to the addition of Sn in the detection solution.

### Ultramicroelectrode for Heavy Metals Detection

In order to meet the requirement of different water samples, improvement of electrochemical sensor in low detection limits and high sensitivity is required. Recently, special attention has been paid to ultramicroelectrodes for improving the characteristics of electrochemical sensors. The ultramicroelectrode can form hemispherical diffusion layers and reduced ohmic resistance effects, resulting in an

**Fig. 14** (a) SWSVs of Sn/GNPs/gold microelectrode with increasing concentrations of  $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ , and  $\text{Zn}^{2+}$ , (b) linear relationship between the peak currents and the concentrations of  $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ , and  $\text{Zn}^{2+}$



increase in current densities, an increase in mass transport rate, and an enhance in signal-to-noise ratios (Gholizadeh et al. 2012; Kokkinos et al. 2011). The benefits of using ultramicroelectrodes in sensor application are the greater sensitivity that arises from the enhanced mass-transport and lower limit of detection that arises from the high signal-to-noise ratios at such electrodes. Thus, ultramicroelectrodes are recognized as an important tool for electrochemical detection, particularly for heavy metals detection (Berduque et al. 2007; Gardner et al. 2009), with high sensitivities, low detection limit, fast response times, and reproducible geometries (Bagal-Kestwal et al. 2008; Cugnet et al. 2009; Li et al. 2011b).

Although ultramicroelectrodes yield extremely large current densities, they generate small currents, which are difficult to detect with electrochemical instruments (Sandison et al. 2002). Therefore, ultramicroelectrode arrays (UMEAs) have been used to fabricate sensors with both high current densities and satisfactory current signals. Nevertheless, the spacing between the electrodes in the array becomes

important (Davies and Compton 2005). In order to reach a regime where hemispherical diffusion layers of the neighboring electrodes do not overlap, the ratio of the interelectrode spacing ( $d$ ) and the radius of an individual electrode ( $R_b$ ) need to be large enough. Many literatures proved that  $d/R_b > 10$  was required with the aim of avoiding this phenomenon (Gardner et al. 2009; Li et al. 2011b). Although deep recesses may reduce the benefits of UMEAs, it has been shown that the overlap between the diffusion fields of the individual electrodes will occur to a lesser degree (Wang et al. 2013). In our study, three types of recessed UMEAs with  $d/R_b$  of 5, 10, and 20 were fabricated and compared to optimize the geometries of UMEAs (Wang et al. 2013).

The microsensor chip was fabricated by MEMS technique. The sensor chip has a Pt counter electrode, a Pt pseudo-reference electrode, and an array of Au working ultramicroelectrode. The Au ultramicroelectrode arrays were fabricated by standard photolithographic and lift-off techniques using SU-8 as the insulating layer. The thickness of the insulating layer is about 3  $\mu\text{m}$ . Three types of recessed UMEA were designed with  $d/R_b$  ratio of 5, 10, and 20, and designated as UMEA-5, UMEA-10, and UMEA-20, respectively. Each UMEA consists of 112 disk-shaped ultramicroelectrodes with individual electrode's radius of 10  $\mu\text{m}$ . The on-chip reference electrode was fabricated by coating Ag/AgCl ink on the substrate electrode of Pt and annealing at 90  $^\circ\text{C}$  for 1 h. The photograph and the cross-sectional view of the microsensor chip are shown in Fig. 15a, b. The surface morphologies of UMEA-5, UMEA-10, and UMEA-20 are characterized by SEM images in Fig. 15c.

The electrochemical behavior of the UMEAs varied significantly with  $d/R_b$ . A peak-shaped voltammogram was produced by the UMEA-5, while a sigmoidal curve was produced by UMEA-20, indicating hemispherical diffusion of ferricyanide towards the working ultramicroelectrode surface. The UMEA-10 produced cyclic voltammograms with mixed diffusion characteristics. This result is in good agreement with the literature data, where a  $d/R_b > 10$  is recommended in order to prevent from overlapping and consequently from a macroelectrode behavior (Gardner et al. 2009). The steady-state limiting current at a disk microelectrode is given by Cugnet et al. (2009):

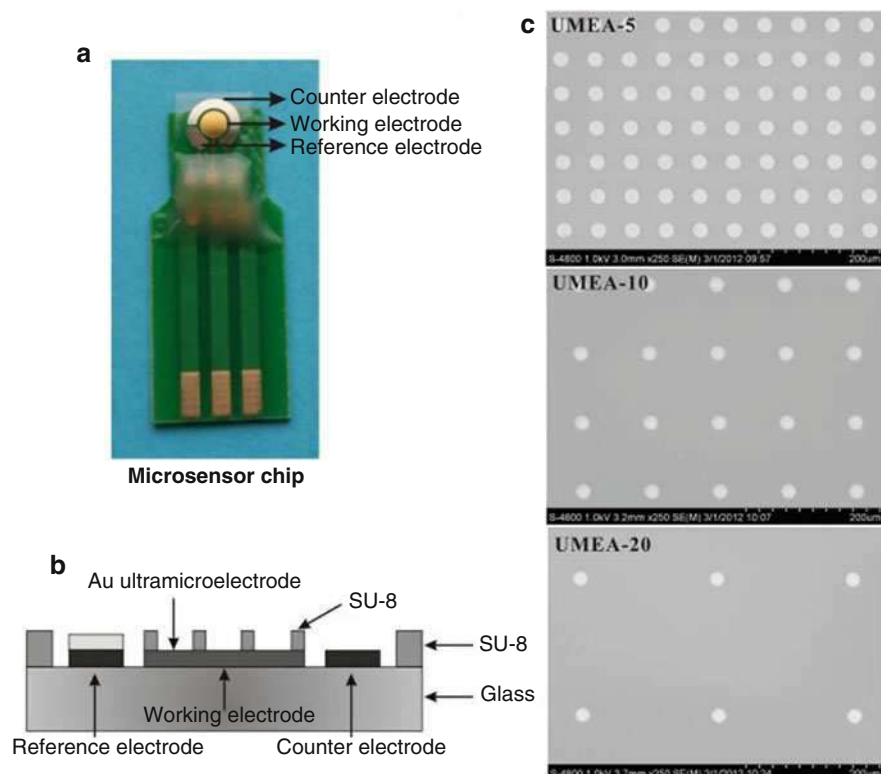
$$I_{\text{lim}} = 4nFDC_0R_b x. \quad (7)$$

Considering the influence of insulator layer to the steady-state limiting current, the expression for recessed microelectrode is defined as (Godino et al. 2009):

$$I_{\text{lim}} = \frac{4nFDC_0R_b}{\frac{4L}{\pi R_b} + 1} x, \quad (8)$$

where  $n$  is electron transferred number of the reaction ( $n = 1$ ),  $F$  the Faraday constant ( $F = 96,485 \text{ C mol}^{-1}$ ),  $D$  the diffusion coefficient of  $\text{Fe}(\text{CN})_6^{3-}$  ( $D = 7.6 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ ),  $C_0$  is the bulk concentration of  $\text{Fe}(\text{CN})_6^{3-}$  ( $C_0 = 2 \text{ mM}$ ),  $R_b$  the radius of the





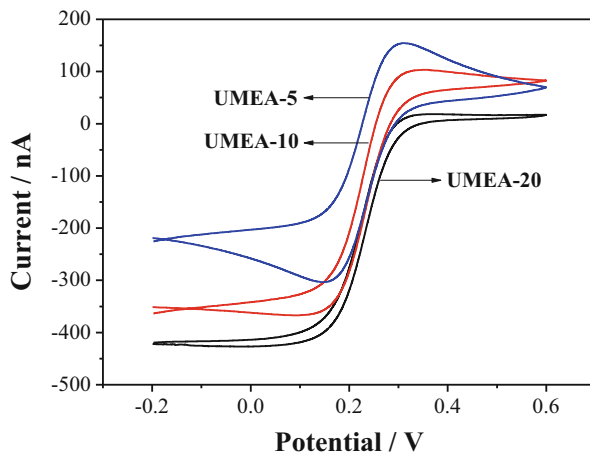
**Fig. 15** The photograph of the microsensor chip (a), the cross-sectional view of the microsensor chip (b), and the SEM images of UMEA-5, UMEA-10, and UMEA-20 (c)

microelectrode ( $R_b = 10 \mu\text{m}$ ),  $L$  the depth of the recess ( $L = 3 \mu\text{m}$ ), and  $x$  is the number of the microelectrode in the array ( $x = 112$ ). According to Eq. 2, the calculated steady current for UMEA is 475 nA. Compared with the experimental values of 303, 367, and 426 nA for UMEA-5, UMEA-10, and UMEA-20, the UMEA-20 showed the best agreement with the theoretical values, as shown in Fig. 16.

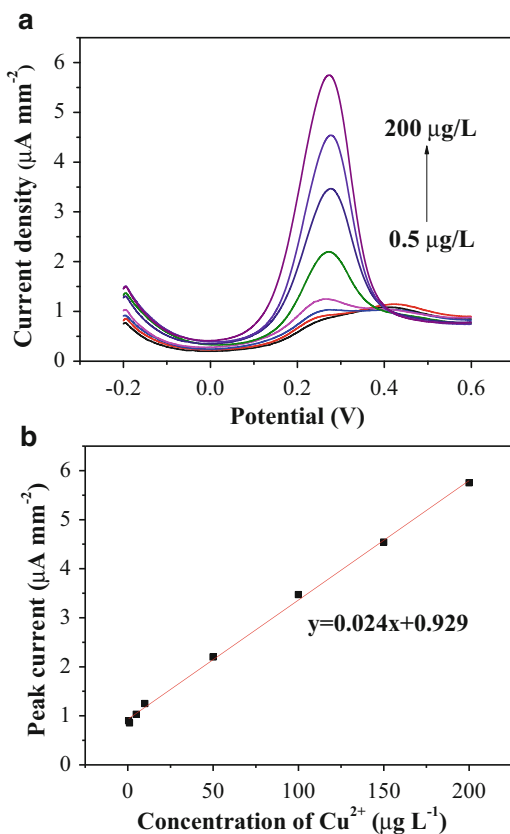
The electrochemical characteristics of gold nanoparticles-modified UMEA-20 were assessed by detecting the standard copper solutions with different concentration. The stripping peak currents increased linearly with copper concentration from 0.5 to 200  $\mu\text{g L}^{-1}$ , as shown in Fig. 17a. The linear equation is  $I_{\text{Cu}}(\mu\text{A mm}^{-2}) = 0.024C_{\text{Cu}}(\mu\text{g L}^{-1}) + 0.929$ , with correlation coefficient of 0.998 and sensitivity of  $0.024 \mu\text{A mm}^{-2} \mu\text{g}^{-1} \text{L}$ , as shown in Fig. 17b. The LOD was estimated to be  $0.2 \mu\text{g L}^{-1}$  ( $S/N = 3$ ). Comparing with the literature (Orozco et al. 2008), the LOD of this work is lower, owing to the favorable background current of the gold nanoparticles-modified UMEA-20 and the using of pulse technique of SWSV which can further minimize the influence of background current.



**Fig. 16** CV curves of UMEA-5, UMEA-10, and UMEA-20 in 2 mM ferricyanide at a scan rate of  $50 \text{ mV s}^{-1}$



**Fig. 17** SWSV curves of gold nanoparticles-modified UMEA-20 at different concentrations of copper ions (a) and the calibration plot (b)



## Pre-treatment Units

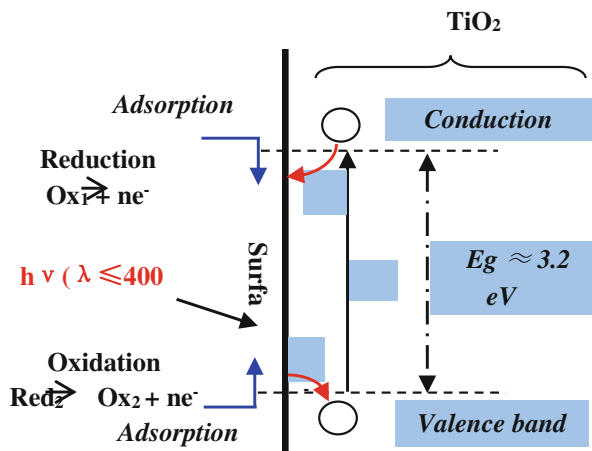
### Micro Digestion Reactors

Before the detection of TP, a digestion pretreatment procedure of the water sample must be proceeded to convert all types of phosphorus to orthophosphate, the readily analyzable form. Subsequently, either electrochemical or optical inspection method is employed to detect the concentration of orthophosphate and complete the detection of TP.

The common used procedures for water treatment (Maher and Woo 1998) mainly rely on an elevated temperature digestion with mineral acids and/or oxidizing agents, for example perchloric acid (Hinkamp and Schwedt 1990), nitric-sulfuric acid (American Public Health Association 2000), and peroxydisulfate (Woo and Maher 1995). Among them, the autoclave persulfate oxidation technique is the most commonly used method for the digestion of TP. This technique was recommended as one of the standard methods for the examination of water and wastewater by the Water Environment Federation (Method4500-P) (Clescerl et al. 1998) and the Environmental Protection Agency (Method365) (O'Dell 1993). This traditional autoclave persulfate oxidation method is a high temperature oxidation (HTO) process with the temperature of 120 °C, the pressure of 1.1 kg/cm<sup>2</sup>, and potassium peroxydisulfate as oxidant. Another commonly used digestion method is microwave digestion (Maher et al. 2002), which is also a HTO process. Although microwave digestion method has many advantages such as high speed and high efficiency, strong oxidant agent such as potassium peroxydisulfate, nitric acid, and perchloric acid are still needed.

Photocatalysis has recently appeared as a new emerging technique with the advantages of high efficiency and environmental friendly (Bozzi et al. 2004; Daneshvar et al. 2005; Daniel and Grutz 2007; Mace and Duce 2002; Oms et al. 2003; Roig et al. 1999). Among many kinds of photocatalyst, titanium dioxide (TiO<sub>2</sub>) has often been proposed for the elimination of organic pollutants, due to the merits of relatively inexpensive, safe, chemically stable, and high photocatalytic activity (Elmolla and Chaudhuri 2010; Siboni et al. 2012; Uğurlu and Karaoğlu 2009). The photocatalytic process is initiated by the photogeneration of hole/electron pairs (e<sup>-</sup>/h<sup>+</sup>) by absorption of UV light with energy equal to or higher than the corresponding band gap (Herrmann 2010), as shown in Fig. 18. An electron will be promoted from the valence band to the conduction band, when TiO<sub>2</sub> absorbs a photon containing the energy equal to or larger than its band gap (3.2 eV). The activation of electrons (e<sup>-</sup>) results in the generation of "holes" (h<sup>+</sup>, electron vacancy) in the valence band. In this reaction, h<sup>+</sup> and e<sup>-</sup> are powerful oxidizing and reducing agents, respectively. The electron-hole pairs may recombine in a short time or take part in chemical reactions depending on reaction conditions and molecular structures of the semiconductors. The strong oxidation power of h<sup>+</sup> enables it to react with OH<sup>-</sup> in water to generate the highly active hydroxyl radical (·OH) which is also a powerful oxidant. Most phosphorous compounds can be oxidized completely by either the hydroxyl radicals or the holes themselves to innocuous final products (e.g.,

**Fig. 18** Principle scheme of TiO<sub>2</sub>-based photocatalysis



PO<sub>4</sub><sup>3-</sup>, CO<sub>2</sub>, and H<sub>2</sub>O). In addition, the reducing power of the electrons can induce the reduction of molecular oxygen (O<sub>2</sub>) to superoxide (O<sub>2</sub><sup>-</sup>). It has been confirmed that the superoxide is almost as effective as the holes and hydroxyl radicals in the chain reactions for the breaking down of organic compounds.

In many photo-decomposition reaction systems, TiO<sub>2</sub> powders are often used as a photocatalyst (Garcia et al. 2007; Lachheb et al. 2002). However, powder catalysts have several problems, such as difficulties in the separation of the catalyst from suspension after the reaction and the prevention of aggregation in high concentration suspensions. To avoid the agglomeration, suspension must be diluted, and then the overall reaction rate tends to be slow. These problems can be solved by the immobility of catalyst. TiO<sub>2</sub> thin films have been prepared by Sol-Gel method (Shen et al. 2009), sputter method (Dorjpalam et al. 2005) and chemical vapor deposition (CVD) method (Hitchman and Tian 2002), as well as by epoxide resin gluing on silicon wafer.

### Reaction Kinetics of TP Digestion Process

Several experimental results indicated that the oxidation rates of photocatalytic oxidation of various compounds in the heterogeneous system over illuminated TiO<sub>2</sub> are in accord with the Langmuir–Hinshelwood (L-H) kinetics model (Attia 2008; Guettaï and Amar 2005; Kima and Hongb 2002; Konstantinou and Albanis 2004):

$$r = -dc/dt = k_1k_2c/(1 + k_2c), \quad (9)$$

where  $r$  is the oxidation rate of the reactant [mg/(L·min)],  $c$  is the concentration of the reactant (mg/L),  $k_1$  is the adsorption constant of the reactant (L/mg), the absorption constants for the compound modules adsorbed on the TiO<sub>2</sub> surface, and  $k_2$  is the reaction rate constant [mg/ (L·min)].

When the initial concentration of the compound to be oxidized is small enough (e.g., mg/L), the term  $k_2c$  could be ignored. Equation 1 could be simplified to a pseudo-first-order rate equation format:

$$-dc/dt = k_1k_2c = kc, \quad (10)$$

where  $k$  represents the product of  $k_1$  and  $k_2$ , and  $k$  could be defined as a pseudo-first-order rate constant. A common form of the pseudo-first-order rate kinetic equation can be obtained by integrating the equation above (Hunt et al. 1991; Konstantinou and Albanis 2004):

$$\ln(c_t/c_0) = -kt, \quad (11)$$

where  $k$  is pseudo-first-order rate constant,  $t$  is the reaction time, and  $c_t$  and  $c_0$  are the compound concentrations at the irradiation time  $t$  and initial concentration, respectively. The slope of the plot of  $\ln(c_t/c_0)$  versus time gives the value of rate constant  $k$  ( $\text{min}^{-1}$ ).

The plots of  $k$  versus temperature for two typical TP components during the thermal-assisted photocatalysis digestion process are shown in Fig. 8. With the increase of the reaction temperature, the values of  $k$  increased, but not be linear. The increase of the rate constant became sharper at the higher temperature. It can be concluded that temperature has a significant enhanced effect on the reaction rate of the photocatalysis digestion. The relationship between the rate constant  $k$  and the temperature was determined quantitatively by the Arrhenius Equation which was used to describe the nonlinear relationship between  $k$  and temperature (Hunt et al. 1991).

The Arrhenius equation is based on the collision theory and it is an experiential equation proposed according to the experiential results (Laidler 1984). Arrhenius argued that when reactants transform into products, they must first acquire a minimum amount of energy to overcome the repulsive force between electron clouds due to the infinitely approach of molecules. This minimum energy is called the activation energy  $E_a$ . At an absolute temperature  $T$ , the fraction of molecules that have a kinetic energy greater than  $E_a$  could be calculated from statistical mechanics, Maxwell-Boltzmann Distribution. The Arrhenius equation is a simple, but remarkably accurate formula for the temperature dependence of the reaction rate constant, and it is available in the elementary reaction, nonelementary reaction, and heterogeneous reaction (Salaüun et al. 2007).

The logarithmic formula is written as follows:

$$\ln k = -E_a/(RT) + \ln A, \quad (12)$$

where  $k$  is the rate constant,  $T$  is the absolute temperature in degrees Kelvin,  $E_a$  is the activation energy,  $A$  is the pre-exponential factor, and  $R$  is the gas constant ( $8.314 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$ ). The unit of the pre-exponential factor  $A$  is identical to those of the rate constant. Given the small temperature range that kinetic studies occur in, it is reasonable to approximate the activation energy and pre-exponential factor as

being independent of the temperature. So the plot of  $\ln k$  versus  $1/T$  produces a straight line according to Eq. 12.

### Photocatalytic Digestion Based on Glued TiO<sub>2</sub> Film

TiO<sub>2</sub> film was prepared by spinning coating on silicon wafer. The silicon wafer from the 46th research institute of Electronic Equipment has a diameter of 50 mm. A 300 nm thick silicon dioxide layer was made by thermal oxidation, and then a 300 nm thick silicon nitride layer was made by low pressure vapor deposition (LPCVD). Firstly, epoxide-resin glue was spun on the surface of the wafer with the spin rate of 2,000 rpm for 1 min. Then 0.05 g TiO<sub>2</sub> nanoparticles were dispersed on the surface of the glue. The TiO<sub>2</sub>/silicon was then dried at 100 °C for 2 h.

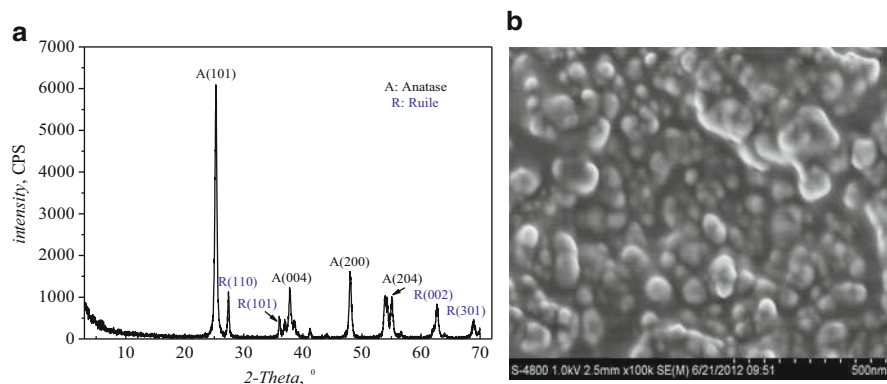
The function of the epoxide-resin glue is only to fix the TiO<sub>2</sub> particles on surface of the silicon wafer. The epoxide-resin glue is inert, resistant to acid and alkali. The characteristics of the epoxide-resin glue will not change at the temperature less than 150 °C. Basically, epoxide-resin glue has no influence on the photocatalyst property.

The photocatalytic activity of TiO<sub>2</sub> film was investigated with respect to the photocatalysis digestion of Na<sub>5</sub>P<sub>3</sub>O<sub>10</sub> and C<sub>3</sub>H<sub>7</sub>Na<sub>2</sub>O<sub>6</sub>P. The reactor is a cylinder made of quartz, with the height of 15 mm and inner diameter of 55 mm. The solution contained in the reactor for digestion is 9 mL in volume, about 5 mm thick. The digestion process of water sample was started by immersing the TiO<sub>2</sub> nanoparticles-attached silicon chip into the cylinder reactor, and then turned on the UV light and heater. The radiated power of UV light on the surface of the chip was around 7,000 μW/cm<sup>2</sup>. The temperature of the reaction was controlled by the heater. The digestion temperature was preset by the heater. The temperature of the solutions elevated with the increase of the time after the heater turned on. At the beginning, the temperature elevated quickly as the increase of the time. The acceleration of the temperature became slower and slower as the increase of the time till the preset temperature achieved. Four minutes, 6 min, and 8 min were needed for the solutions to achieve 40 °C, 50 °C, and 60 °C, respectively.

Before the color reaction, 4.5 ml digested water samples were transferred from reactor to centrifuge tube. The TiO<sub>2</sub> film was used repetitively and showed long-term stability. The silicon wafers with TiO<sub>2</sub> film was swashed with ultra-pure water for 2 min after each digestion process. After the digestion process, the solutions were cooled down for 15 min, and then 1 mL ascorbic acid solution and 1 mL ammonium molybdate solution were added into the digested solution in order. This mixed solution was kept stable for 10 min after mixed thoroughly. The orthophosphate concentration (weight by P) of the digested solution was measured by the ultraviolet spectrophotometer, at the wavelength of 700 nm against a reagent blank.

The digestion method proposed in this work doesn't need oxidant agent. Since the rudimental oxidant in the water sample has a remarkable interference for the electrochemical detection of TP, this method proposed in this work is favored for the electrochemical detection method.

XRD analysis was performed to identify the phase structure and composition of nano titanium dioxide powders purchased from Yunnan University, China, as shown in Fig. 19a. The TiO<sub>2</sub> used in this work is a mixture of anatase and rutile

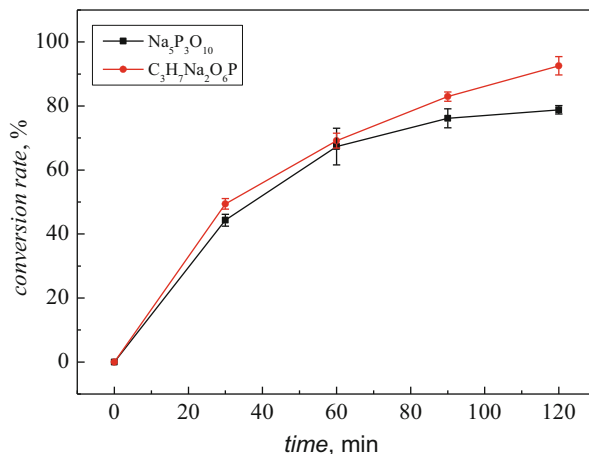


**Fig. 19** XRD (a) and SEM (b) examination of TiO<sub>2</sub> film

(anatase: 84%, rutile: 16%). It has been reported that the mixture of anatase and rutile with a certain proportion has the best catalytic activity (Brillas et al. 1995). The reason could be that the mixture of two different crystal structures have a larger disfigurement density in the crystal lattices, and the disfigurements form the carrier traps, in which the electron or hole could be get trapped. In this way the lifetimes of both hole and electron were extended. Figure 19b shows the SEM image of nano TiO<sub>2</sub> on the surface of silicon wafer. And the average size of the TiO<sub>2</sub> particle was about 50 nm.

In order to examine the effect of photocatalysis oxidation and individual UV irradiation, experiments were conducted under room temperature using initial concentration 4 mg/L (weight by P) Na<sub>5</sub>P<sub>3</sub>O<sub>10</sub> and C<sub>3</sub>H<sub>7</sub>Na<sub>2</sub>O<sub>6</sub>P, respectively. When the UV irradiation (365 nm) was applied alone without TiO<sub>2</sub> catalyst, the experimental results showed that the concentrations of the two phosphorous compounds had no decrease in the solutions, which indicated that the individual UV light could not oxidize the phosphorus compounds. However, a significant increase of the orthophosphate concentration, which means a significant reduction of the concentration of Na<sub>5</sub>P<sub>3</sub>O<sub>10</sub> and C<sub>3</sub>H<sub>7</sub>Na<sub>2</sub>O<sub>6</sub>P, was observed after the treatment of UV irradiation and TiO<sub>2</sub> catalysis, which implies that the phosphorus compounds have been oxidized. The oxidation of the phosphorus compounds could be attributed to the combination of the UV irradiation and TiO<sub>2</sub> through which hydroxyl radical ( $\cdot$ OH) was added to the solutions. The generation of  $\cdot$ OH depends on the absorption of UV irradiation on TiO<sub>2</sub>. Figure 20 shows the conversion rates (the ratio of the concentration of P converted into orthophosphate to the initial concentration of P in phosphorus compound) of the two typical phosphorus substances for different irradiation time at room temperature (25 °C). The conversion rates elevated with the increase of the irradiation time for both substances, but not in a linear relationship. It was observed that the conversion rates increased rapidly during the first 30 min and then with the irradiation time increased, the acceleration of the conversion rate became slower and slower. After 120 min, the conversion rates of Na<sub>5</sub>P<sub>3</sub>O<sub>10</sub> and C<sub>3</sub>H<sub>7</sub>Na<sub>2</sub>O<sub>6</sub>P could reach 78.8% and 92.6%, respectively.

**Fig. 20** The conversion rates of two typical phosphorus substances versus irradiation time at room temperature (25 °C)



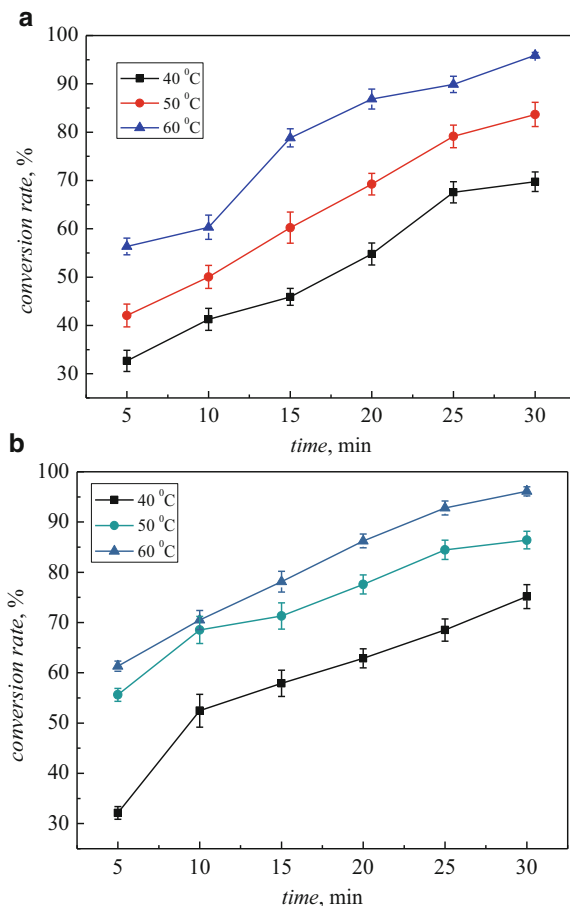
To reduce the digestion time and improve the efficiency, the thermal-assisted UV photocatalysis digestion method is proposed in this work. The experiments were conducted at reaction temperature 40 °C, 50 °C, and 60 °C, respectively, and the effect of temperature on the digestion process was studied. Conversion rates of the two typical phosphorus substances at different irradiation time and different temperature are shown in Fig. 21. It can be seen that the trends of change on the conversion rate versus irradiation time at these reaction temperatures (40 °C, 50 °C, and 60 °C) are similar to that at the room temperature, and the conversion rates are larger at the higher reaction temperatures than that at room temperature, which means the oxidation speed of phosphorus compounds to orthophosphate is much faster at a higher reaction temperature. It can be concluded that heating has an enhanced effect on the photocatalysis oxidation digestion process. At the condition of 30 min and 60 °C, the conversion rates of Na<sub>5</sub>P<sub>3</sub>O<sub>10</sub> and C<sub>3</sub>H<sub>7</sub>Na<sub>2</sub>O<sub>6</sub>P are 95.9% and 96.1%, respectively. When the two kinds of compounds were kept at 40 °C, 50 °C, and 60 °C for 30 min, almost no orthophosphate was detected in the samples, which indicated that the substances chosen in the work were relatively stable during the temperature of 40–60 °C.

The proposed thermal-assisted UV digestion method was also used to digest real water samples collected from lakes and rivers around Beijing. The experimental results were compared with that using standard autoclave persulfate oxidation method. The digestion condition of thermal-assisted digestion process was 60 °C and 30 min. The measurement results are shown in Table 4. It is found that the conversion rate using thermal-assisted UV digestion method is consistent with that using standard autoclave persulfate oxidation method.

### Integrated Photocatalytic Microfluidic Platform

As an important and necessary step of TP detection, a digestion pretreatment procedure must be proceeded to convert all types of phosphorus into phosphate. Subsequently, electrochemical or optical inspection methods could be employed to

**Fig. 21** The conversion rates of  $\text{Na}_5\text{P}_3\text{O}_{10}$  (a) and  $\text{C}_3\text{H}_7\text{Na}_2\text{O}_6\text{P}$  (b) versus irradiation time at different reaction temperature



**Table 4** TP analysis using different digestion methods with real water samples

Water samples	Digestion method		Conversion rate
	Traditional method (mg/L)	Thermal-assisted digestion method (mg/L)	
1#	0.1504	0.1332	88.56%
2#	0.1660	0.1593	95.96%
3#	0.3076	0.2834	92.13%

complete the phosphate detection. The traditional high temperature oxidation (HTO) digestion method using high pressure sterilizer for TP determination is high power consuming, and it is not suitable for portable detection. Especially, with the development of MEMS-based microsensors (e.g., the micro electrochemical electrodes in the previous work (Bai et al. 2014b)), small-sized, low power consumption, and portable detection systems are required.

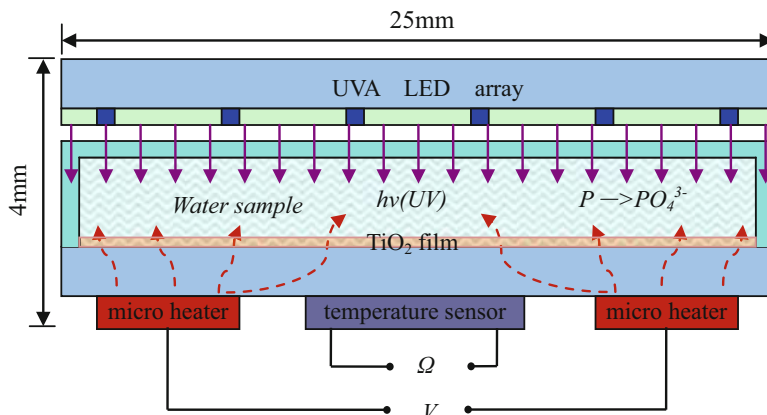


Microfluidic technology is emerging as a platform which is able to manipulate small volumes of liquid in a microchannel network with microscale dimensions (Haerberle and Zengerle 2007). Microfluidic platform equips with different elements on a single chip in order to realize complex functionality. A wide range of chemical reactions have been performed on the microfluidic platform in a rapid manner with higher efficiency due to the large surface to volume ratio and shorter diffusion distance of the reactant solution in the microchannels. Photocatalysis using semiconductor materials is a promising alternative to the conventional treatment processes for a wide variety of organic contaminants in both air and water (Lee et al. 2012; Tsuji et al. 2012; Xi et al. 2011). Recently, several researchers have made their efforts to adopt these microfluidic chips as their choice of microreactor in order to enhance photocatalytic activity. Microreactors can be expected to exhibit some distinct properties on photoreactions, such as higher spatial illumination homogeneity and better light penetration through the entire reactor depth in comparison to large-scale reactors. Among the different photocatalytic materials, titanium dioxide ( $\text{TiO}_2$ ) is preferred for its strong oxidizing power, low cost, bulk nontoxicity, ready availability, activity under a broad range of pH, and its long-term stability (Bozzi et al. 2004; Daniel and Grutz 2007; Siboni et al. 2012). Usually, a  $\text{TiO}_2$  film is immobilized on the inner surface of the microchannel by injecting a  $\text{TiO}_2$  precursor followed by calcinations (Aran et al. 2011). Due to the short diffusion distance in a microchannel, the flow reactant solution can come into contact with the  $\text{TiO}_2$  film more effectively, leading to an enhanced performance in photocatalytic reactions as compared to that seen in macroscale reactors under static conditions. Therefore, these photocatalytic degradation systems are actively used as an economical remediation method for various purification applications.

In recent decades, lighting-emitting diodes (LEDs) have developed rapidly with the advantages of long-lasting, robust, small in size and high in efficiency (Hu et al. 2012; Li et al. 2005; Moiseev et al. 2011). In most studies of microfluidic photocatalytic reactors, the external commercial UV-LED are usually employed as the UV light source, which means the microfluidics reactors and UV light source are discrete (Chen et al. 2005; Lindstrom et al. 2007; Wang and Ku 2006). Besides, the existing photocatalytic microfluidic reactors for organic matter degradation usually work at room temperature (Ghosha et al. 2009; Qin 2013). For the refractory pollutants in wastewater, it usually takes a relatively long time to degrade completely using photocatalytic reaction at room temperature, such as azo dyes, organic hydrogen chloride, and organophosphorus pesticide.

The schematic of the working principle of photocatalytic microchip for TP digestion is shown in Fig. 22. When the water sample is irradiated by UVA light in the presence of  $\text{TiO}_2$  photocatalyst, the phosphorus containing compounds could be converted to phosphate during a period of time; UVA LED array was employed to generate the stable UVA light, and micro heater and micro temperature sensor were used to control the temperature of water sample.

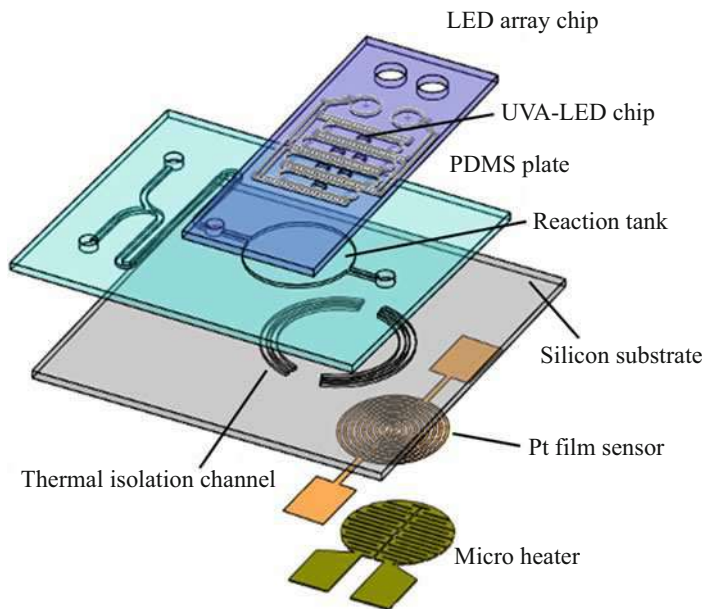
The schematic of the photocatalytic microchip is shown in Fig. 23. This microchip is composed of two parts: the microfluidic-based chip and the UVA-LED array chip. The silicon substrate and PDMS plate was bonded together to form the



**Fig. 22** Schematic of the working principle of photocatalytic microchip for TP digestion

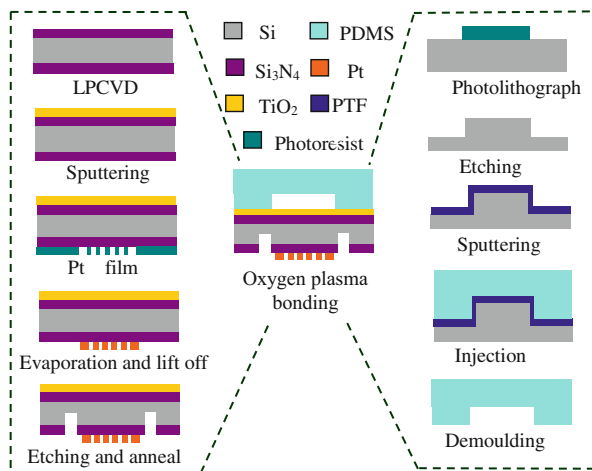
microfluidic chip. A micro Pt film temperature sensor, a copper micro heater, and thermal isolating channels were integrated on the back side of the silicon substrate. The microfluidic channels and reaction tank were made on the PDMS plate by injection molding process. Several micro UVA-LEDs were bonded on the surface of an aluminum PCB board in parallel connection to form the LED array chip. Then the LED array chip was attached on the top side of the microfluidic chip to form a high intensity irradiation above the reaction tank.

The fabrication process of the microfluidic chip is shown in Fig. 24. The silicon substrate was prepared using standard silicon microfabrication techniques. The substrate was a 270  $\mu\text{m}$  thick double-sided polished silicon wafer. Firstly, 300 nm thick layers of silicon nitride ( $\text{Si}_3\text{N}_4$ ) were deposited on both sides of the silicon wafer by low-pressure chemical vapor deposition (LPCVD) as electrical insulation layer. The top surface of the wafer was then coated with a 300 nm thick  $\text{TiO}_2$  layer by sputtering. A 200 nm thick layer of platinum was deposited by sputtering on the back side of the silicon wafer, which was then patterned to form a micro temperature sensor structure. A 10 nm thick adhesion layer of tantalum is necessary in order to improve the adhesion of the platinum film on the silicon nitride layer. To reduce the heat loss, thermal isolation channels were etched around the Pt temperature sensor. After the fabrication, an annealing process was done to make the  $\text{TiO}_2$  film crystallized. The micro copper heater was attached on the surface of the Pt temperature sensor. The PDMS plate with microfluidic channels and reaction tank was made by soft lithography process. Silicon wafer was used to fabricate the mold by deep reactive-ion etching (DRIE), and then a 100 nm thick PTFE layer was sputtered on the top side of the silicon wafer. The mixed PDMS was poured on to the surface of the silicon mold to transfer the patterns. When the PDMS was solidified thoroughly, detach the PDMS plate from the silicon wafer. Figure 25 shows the photograph of the photocatalytic microchip system.

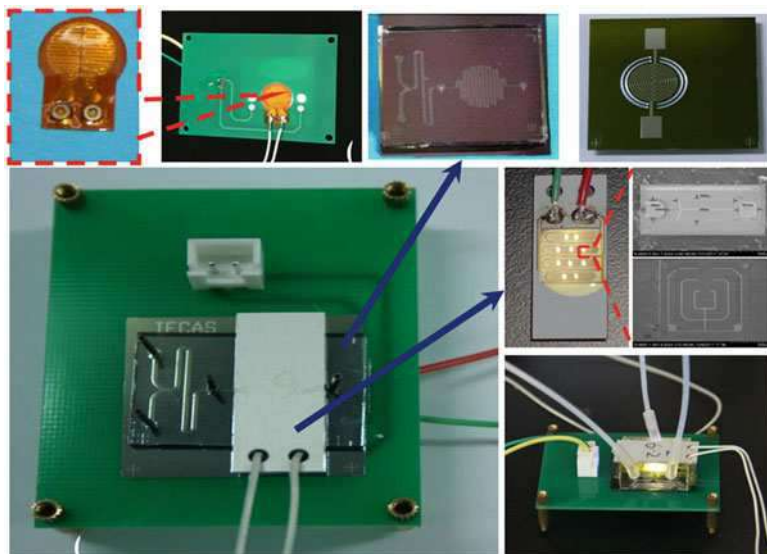


**Fig. 23** Schematic diagram of the MEMS-based photocatalytic microchip

**Fig. 24** The fabrication process of the microfluidic chip



The crystal phase of  $\text{TiO}_2$  film was anatase after annealed at  $450\text{ }^\circ\text{C}$  in air atmosphere for 2 h. The anatase phase has been proved to have the best catalytic activities among the three phase structures of titanium dioxide. Platinum (Pt) was chosen as the material of the temperature sensor to monitor the reaction temperature of the microchip due to its high precision and compatibility with MEMS technology.



**Fig. 25** The photograph of the photocatalytic microchip system

The Pt film temperature sensor sputtered on the back side of the microfluidic chip is not a standard component, so the calibration process was done to set up the relationship between Pt film resistance and the temperature. The relationship was obtained in the temperature range of 22–140 °C as the following:

$$R_T = 290.62(1 + 2.253 \times 10^{-3}T), \quad (13)$$

where  $T$  is the absolute temperature,  $\Omega$ ;  $R_T$  is the Pt film resistance at absolute temperature  $T$ , K.

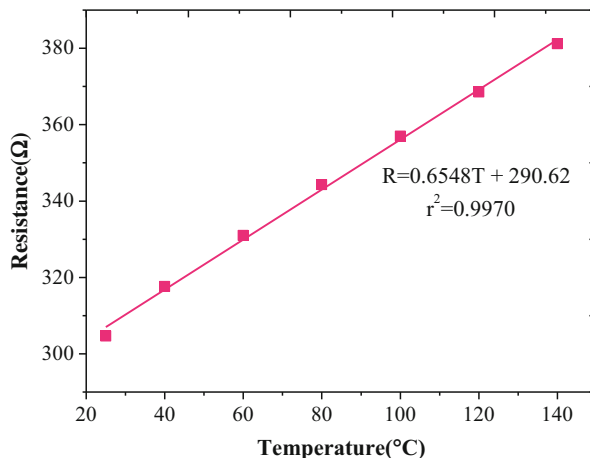
According to Eq. 1, the temperature coefficient could be calculated. And at 22–140 °C, the temperature coefficient of the sputtered Pt film was  $2.253 \times 10^{-3} \text{ }^\circ\text{C}^{-1}$ , and the obtained sensitivity was  $0.6548 \text{ } \Omega/^\circ\text{C}$ . According to Eq. 1, a digestion temperature could be calculated with a tested  $R_T$ .

$$T = (R_T - 290.62)/0.6548 \quad (14)$$

The UV light source employed in this work is the UVA-LED array chip which integrated with ten UVA-LEDs on the surface of the PCB board in parallel connection (Fig. 25).

Figure 27 shows the saturation characteristic and spectral characteristic curves of the LED array chip integrated with ten UVA-LEDs. As shown in Fig. 27a, the radiation power of the LED array chip increased with the increase of the applied current. But when the applied current was higher than 0.25 A, the radiation power decreased. The maximum irradiation power 0.0236 W was obtained when the applied current reached 0.25 A. As the elevation of the temperature, the crystal

**Fig. 26** The calibration curve of the temperature sensor



lattice vibration of semiconductor became intensified, and the number of impurities and defects in semiconductor would increase. In this way, the photo-absorption efficiency of the semiconductor enhanced, which resulted in the increase of the loss coefficient. So when the applied current reached a certain value, the radiation power of the LED array chip decreased.

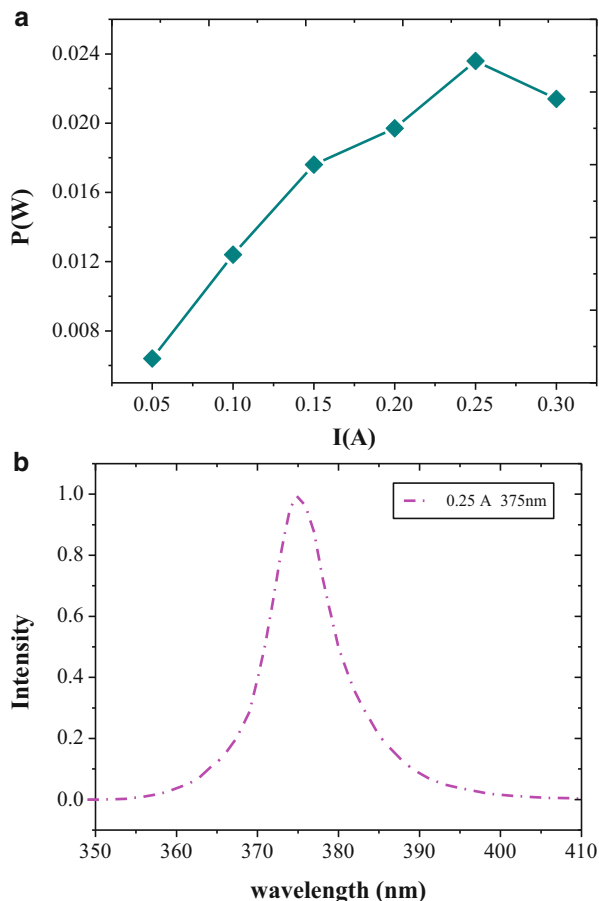
The spectral characteristic curve of the LED array chip integrated with ten UVA-LEDs is shown in Fig. 27b. As shown in this figure, the spectral range of the UVA-LED is relatively narrow, and the peak wavelength is 375 nm.

The digestion process of water sample was started by injecting the water sample into the integrated photocatalytic chip. After the inlets and outlets of the chip were sealed, a certain voltage was applied on the micro copper heater to increase the temperature of the chip. When the chip temperature rose to the expected value, the UV LED array chip was opened and the digestion process started. After the digestion process was done, the chip cooled down for 5 min, and a 5  $\mu\text{L}$  digested water sample was extracted out of the chip. Then a 1  $\mu\text{L}$  ascorbic acid solution and a 1  $\mu\text{L}$  ammonium molybdate solution were added into a 5  $\mu\text{L}$  digested solution in order. This mixed solution was kept stable for 10 min after mixed thoroughly. The orthophosphate concentration of the digested solution was measured by the UV spectrophotometer at the wavelength of 700 nm against a reagent blank.

Sodium tripolyphosphate ( $\text{Na}_5\text{P}_3\text{O}_{10}$ ) and sodium glycerophosphate ( $\text{C}_3\text{H}_7\text{Na}_2\text{O}_6\text{P}$ ) were chosen as the typical components of TP. The current applied on LED array chip was 0.25 A, and the irradiation power was 0.0236 W.

Figure 28a shows the digestion rates at different reaction time without heating process (room temperature is 33 °C). The conversion rates elevated with the increase of the irradiation time for both substances. After 120 min, the digestion rates of  $\text{Na}_5\text{P}_3\text{O}_{10}$  and  $\text{C}_3\text{H}_7\text{Na}_2\text{O}_6\text{P}$  could reach 57.74% and 60.48%, respectively. Digestion rates of two typical phosphorus substances at 30 min and different reaction temperature are shown in Fig. 28b. And as shown in this figure, the conversion rates increased with the elevation of the reaction temperature for both substances. At the

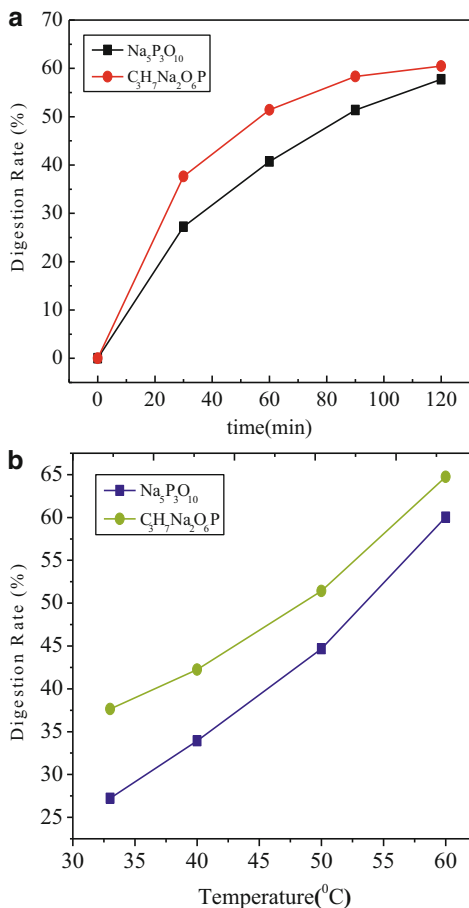
**Fig. 27** Characteristics of the LED array chip: (a) saturation characteristic curve; (b) spectral characteristic curve



condition of 30 min and 60 °C, the digestion rates of  $\text{Na}_5\text{P}_3\text{O}_{10}$  and  $\text{C}_3\text{H}_7\text{Na}_2\text{O}_6\text{P}$  are 60.04% and 64.75%, respectively.

Compared with the results of the UV digestion process at room temperature, the enhancement of heating on UV photocatalytic digestion process could be observed. In the collision theory, it is thought that the collision between reagents is the precondition for the reaction to occur (Goldberger and Watson 2004). Among all the collisions, most of them could not produce reaction, and only very minor collisions are effective. Firstly, molecules in the collisions that could react must have sufficient energy to overcome the repulsive force between electron clouds due to the infinitely approach of molecules. The minimum energy needed is called the activation energy  $E_a$ . In this way, the rearrangement of atoms in molecules could take place, which means the chemical reaction occurs. So in order for a chemical reaction to take place, molecules must collide with energies equal to or greater than the activation energy for the reaction. At an absolute temperature  $T$ , the fraction of molecules that have a kinetic energy greater than  $E_a$  could be calculated from

**Fig. 28** Photocatalytic characteristic of the photocatalytic microchip



statistical mechanics, Maxwell-Boltzmann Distribution. And the function of Maxwell-Boltzmann Distribution has a direct proportion with absolute temperature  $T$ .

Table 5 gives the power consumption of the photocatalytic microfluidics chip at different digestion temperature. The reaction temperature of the photocatalytic chip was controlled by adjusting the voltage applied on the micro heater. As shown in this table, the power consumption was relatively low, less than 2 W.

## Pre-concentration

### Extraction

Extraction is an important preconcentration procedure in analytical chemistry. An effective approach, which helps to achieve higher sensitivity and better tolerance to interfering ions, is using extraction procedure prior to determination (Al-Saidi and Emara 2011). The extraction procedure blocks unwanted components and

**Table 5** The power consumption of the microchip at different reaction temperatures

Digestion temperature of the microchip (°X)	Applied voltage on heating electrode (V)	Applied current on LED array chip (A)	Power consumption of the microchip (W)
33	0	0.25	0.9075
40	1.41	0.25	1.1284
50	2.14	0.25	1.4163
60	2.86	0.25	1.8163

concentrate analytes from bulk sample into a small volume which increase the sensitivity of determination.

However, the procedure of extraction is often the most time consuming and manual intensive part of the whole analysis. The minimization and automation of the extraction procedure are the main tasks in analytical chemistry. Microextraction techniques (Andruch et al. 2013; Kocúrová et al. 2012; Stanisz et al. 2014a; Wei et al. 2003) have been gaining a growing interest in the recent years. Dispersive liquid-liquid microextraction (DLLME) illustrate significant advantages among the microextraction techniques (Al-Saidi and Emará 2011; Andruch et al. 2013). DLLME uses small amount of extraction solvent, which makes the procedure environmentally friendly. It has very short extraction time because it achieves the equilibrium state rapidly. As a result of high aqueous to organic phase ratio, excellent enrichment factors can be achieved.

The central procedure of DLLME is to make extraction phase exhaustively dispersed in the sample solution. Methods for generating dispersed liquid phase include rapid injection, high speed stirring (Suárez et al. 2014), and manual shaking (Wu et al. 2013). Other equipment-assisted techniques such as ultrasound-assisted emulsification (Khan et al. 2015), vortex-aided dispersion (Seebunrueng et al. 2014), and microwave-assisted emulsification (Song et al. 2015) were also demonstrated. These methods inflict intense mechanical stimulation on the sample with either manual operation or equipment. They cannot be easily integrated with online systems.

During the process of extraction, the extracting agents can conflict with the electrochemical determination. For example, common methods for mercury extraction rely on chelates. After the formation of complex with chelates, mercury becomes electrochemically inert. Back extraction will liberate mercury at the cost of introducing strong acid, which is also unfavorable for stripping voltammetry: the troubles in this protocol counteract its benefits.

## Ionic Liquids

Ionic liquids (ILs) are a class of ionic compounds which are liquid below 100 °C (Sun et al. 2010). They are different from traditional organic solvents as they are composed from anions and cations. ILs have been widely used as environment-friendly extractants to replace organic solvents. They have many good features such as high conductivities, wide potential windows, and negligible vapor pressure, therefore, attracted increasing attention in analytical chemistry (Amio and Agaosa 2008; Martinis et al. 2009; Stanisz et al. 2014b) and electrochemical related areas



(Afkhami et al. 2014; Bagheri et al. 2015; Endres et al. 2008; Opallo and Lesniewski 2011). Ionic liquids containing pyridinium cations were reported to extract mercury from water efficiently (Papaiconomou et al. 2008). There were many ILs introduced innovative extraction procedures. Temperature controlled ionic liquid dispersive liquid phase microextraction (TC-IL-DLPME) was demonstrated by Zhou et al. in 2008 (Zhou et al. 2007). Based on the principle of solubility change with temperature, ILs dissolve completely at high temperature and disperse exhaustively at low temperature. Comparing to the traditional liquid-liquid extraction, the procedure is safe, rapid, and provides excellent enrichment factor with low reagent consumption. Stanisiz et al. summarized the applications of ILs in the preconcentration and determination of metals recently (Stanisiz et al. 2014a). In previous studies, determinations were conducted by various spectroscopic and mass spectrum techniques.

ILs can be dispersed in water when its methanol mixture is rapidly sprayed using a syringe (Amio and Agaosa 2008). Several new DLLME techniques with ionic liquids had been developed (Stanisiz et al. 2014b; Sun et al. 2010). Ionic liquid/ionic liquid dispersive liquid-liquid microextraction (IL/IL-DLLME) procedure utilizes a hydrophobic IL and a hydrophilic IL as an extraction solvent and a disperser solvent, respectively (Zhao et al. 2011). Ultrasound-assisted ionic liquid DLLME coupled with liquid chromatography had been developed for the extraction and determination of sulfonamides in milk powder (Gao et al. 2012). Microwave had been used to assist the dispersion of ILs in sample (Xu et al. 2011). Ionic liquids were essentially salts. They can be composed by two other salts in the solution. Homogeneous liquid-liquid microextraction (HLLME) using ionic liquids was performed by adding ion-pairing agents to generate a new dissolvable ionic liquid dispersed in the sample (Berton et al. 2013; Gao et al. 2011; Wang et al. 2014). In situ solvent formation IL-DLLME use similar process (Baghdadi and Shemirani 2009). The ILs in solution will shift from dissolved form to suspended form according to many factors such as solvents, temperature, ionic strength, etc. Salt-controlled homogeneous ionic liquid phase microextraction was developed recently (Adlnasab and Ebrahimzadeh 2013). 1-hexyl-3-methylimidazolium hexafluorophosphate and methanol were used as the extraction and consolute solvents. The homogenous solution was broken by the addition of sodium chloride. Temperature controlled ionic liquid dispersive liquid phase microextraction (TC-IL-DLPME) was based on the solubility correlation of methylimidazolium-type ionic liquids with temperature (Zhao et al. 2011; Zhou et al. 2007). A high amount of IL was fully dissolved into sample at 80 °C. Then the solution was ice bathed to force IL to precipitate as fine droplets. Based on similar principles, cold-induced aggregation microextraction with IL was also developed (Baghdadi and Shemirani 2008).

### **Trace Mercury Determination Based on ILs Preconcentration and Stripping Voltammetry**

In this chapter, the preconcentration-stripping voltammetry protocol by ionic liquid preconcentration prior to stripping N-octylpyridinium cations in the selective extraction of mercury were validated in presence of interferences. The determination of trace

mercury in water was performed by extracting mercury with ionic liquids (ILs). N-octylpyridinium tetrafluoroborate and N-octylpyridinium trifluoromethylsulfonate were found to be efficient and selective extractant for mercury. The whole protocol was described in details (Li et al. 2016). Copper (II), zinc (II), lead (II), and cadmium (II) were chosen for selectivity test against mercury (II). For all other metals, extraction results were rather poor. Mercury was enriched by 17 times while interfering ions were reduced by two orders of magnitude in the organic media under optimum condition.

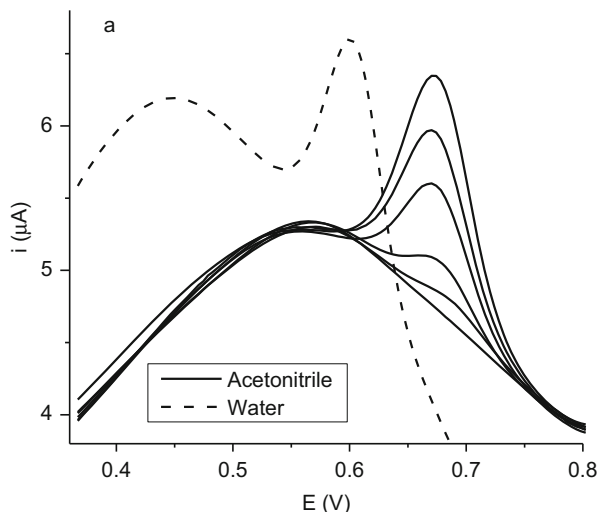
We found that extraction by  $[\text{OPy}]^+[\text{BF}_4]^-$  and  $[\text{OPy}]^+[\text{TfO}]^-$  was consistent in the presence of hydrochloric, sulfuric, nitric, and perchloric acid up to 1 mol/L. Advantage is that the method is tolerant to different sample conditions and compatible with many pre-treatment steps that introduce acids. On the other hand mercury cannot be back extracted into aqueous phase by the common method of increasing proton intensity. Because ILs are miscible with most organic solvents, we needed to select an organic solvent as diluent to liberate mercury ions for determination. Based on the experiments, acetonitrile was chosen as the best working media. The final solution contains enriched mercury and less interference. Formula of supporting electrolyte (40 mM  $\text{HClO}_4$ , 0.6 mM  $\text{LiCl}$ ) was suggested (Fig. 29).

Based on the previous experiments, the detection mechanism can be explained as follows. The interaction between ILs and mercury (II) is dependent on the solvent. Water solvated ILs are discrete cations and anions surrounded by water molecules. We think that these ions do not have strong interaction with mercury ions individually. Simple substance of ILs is formed by neutral molecule clusters. They can extract mercury efficiently. Thus the binding with mercury (II) only lies within the cluster structure of ILs. This explains the different extracting behavior of four N-octylpyridinium ILs. They only differ in the anions. Although the N-octylpyridinium cation might be the key of mercury selectivity, the nonreactive anions determined the cluster structure and altered their characteristics. Because the IL clusters partly existed in the organic solvent, they protected mercury ions from taking part in electrochemical reactions. So all ILs showed some inhibitions on mercury voltammetry.

Trace mercury was detected by differential pulse stripping voltammetry (DPSV) with gold disc electrode in acetonitrile and ILs. Temperature controlled dispersive liquid phase microextraction (TC-DLPME) technique was utilized to improve the performance of preconcentration. It had higher preconcentration factor and less time consumption. It improved the sensitivity and selectivity for electrochemical determination of mercury.

The analytical performance showed that the sensitivity was increased by eight times compared to direct detection in the initial sample using the same electrochemical parameters. The LOD was 0.05  $\mu\text{g/L}$  and the relative standard deviation (R.S.D) of the method is within 10%. Tap, pond, and waste water samples were analyzed with recovery ranging from 81% to 107% and detection limit of 0.05  $\mu\text{g/L}$ . This method could offer an attractive approach for on-site determination of trace mercury contamination in water.

**Fig. 29** Analysis curve for stripping voltammetry determination of mercury in acetonitrile and water. (a) DPSV curve of 0, 10, 25, 50, 75, 100  $\mu\text{g/L}$  Hg (II) in acetonitrile (solid line, from bottom to top) and 100  $\mu\text{g/L}$  Hg (II) in water (dashed line)

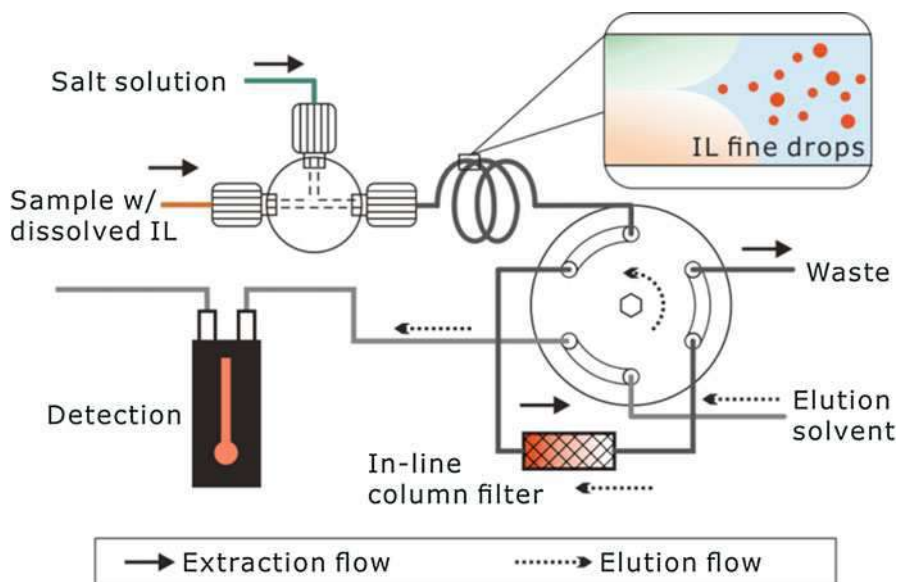


### Salt-Induced ILs Dispersive Liquid-Liquid Microextraction and Online System

Although there were variety of methods to generate dispersed extraction phase, centrifugation which separates the extraction phase by density, needs to be applied in the majority of DLLME procedures up to now. When using an extraction phase lighter than water, special vessels were required. Centrifugation is a time-consuming step and a limitation of DLLME as fast sample preparation methods. It also becomes a major obstacle for automation because centrifuge cannot be integrated into a continuous flow system. In some recent studies, several new procedures had been proposed to replace centrifugation in DLLME. It is possible to use demulsification reagents to break the emulsion. But the process is time consuming and requires the sample to be stationary. Ebrahimpour et al. introduced an online system which uses in-line columns for the collection of solvent drops (Ebrahimpour et al. 2012).

Salt-induced ionic liquid dispersive phase is a powerful tool for extraction. For example, extraction phase volume can be controlled by the amount of injected sodium perchlorate. Using-salt induced ionic liquid dispersive phase, the extraction can be performed by simply mixing and injecting solution. As the volume of extraction phase could be precisely controlled down to a few microliters, common syringe filters and in-line column filters were used to collect the IL fine drops. It eliminates the need for centrifugation, stirring, shaking, etc. It is suitable for sample enrichment from water prior to spectrophotometric determination.

For the dispersive liquid-liquid microextraction, the essential mass transfer process is extremely fast because of the large surface area of fine droplets of extraction phase. Previously, most of the time was consumed in producing dispersive phase and



**Fig. 30** Schematic design of the online SI-DLLME and filter separation system. IL fine drops were generated in-line by the salting out effect. They were then trapped in the column filter. By changing the valve to elution position, trapped IL was rinsed from inverse direction by organic solvents. Detection was done by spectrophotometer with a continuous flow cell

centrifugation. The salt-induced dispersive liquid-liquid extraction (SI-DLLME) and filter separation can be integrated into an automated online system. The system is easily composed by tubing and valves and could perform both extraction and elution by flow injection. Based on the online system, the procedures of phase generation and separation can be improved. Both use little operations and minimal amount of time and reagents. Figure 30 shows a kind of online SI-DLLME and filter separation system.

The online SI-DLLME system takes advantages of rapid and automated extraction with low reagent consumption. In the system, fully dispersed ionic liquid droplets were generated by injecting perchlorate into sample due to the salting out effect. Extraction phase volume was controlled by the salt amount. It was as small as a few microliters. The IL/aqueous phase ratio had reached one in a thousand. Common filters demonstrated their capabilities for collecting such low volume of extraction phase from emulsion. The analytes were recovered by elution. The online system was built by common tubing and valves. Both input and output were continuous-flow which can be easily connected with other instruments. It was successfully applied for the extraction (Li et al. 2015). It can also be used on many other analytes. The proposed microextraction protocol is an attractive approach for both fast manual operation and online system.

## Integrated System

### Integrated Microchip System

Nowadays, the instruments for TP detection are usually big in size, and time- and reagent-consuming. Also, it is not easy to carry and operate them for fieldwork in large water area. With the development of MEMS and microfluidics techniques, micro biochemical sensors and microsystems are attracting much attention for the online monitoring and real-time detection of biological and chemical substances, due to the advantages such as fast-response, and low power- and reagent- consumption. Xia group has developed some microsensors for TP determination (Bai et al. 2014c; Tong et al. 2015). However, it was still a big challenge to integrate the digestion unit and detection unit in a single chip. This chapter presents a fully integrated on-chip microreactor which was developed for the detection of TP in freshwater. An electrochemical three-electrode-system coupled with a thermally assisted UV photocatalytic unit was embedded in a microfluidic chip, and the water samples containing phosphorus could be digested into phosphate and detected by electrochemical unit subsequently.

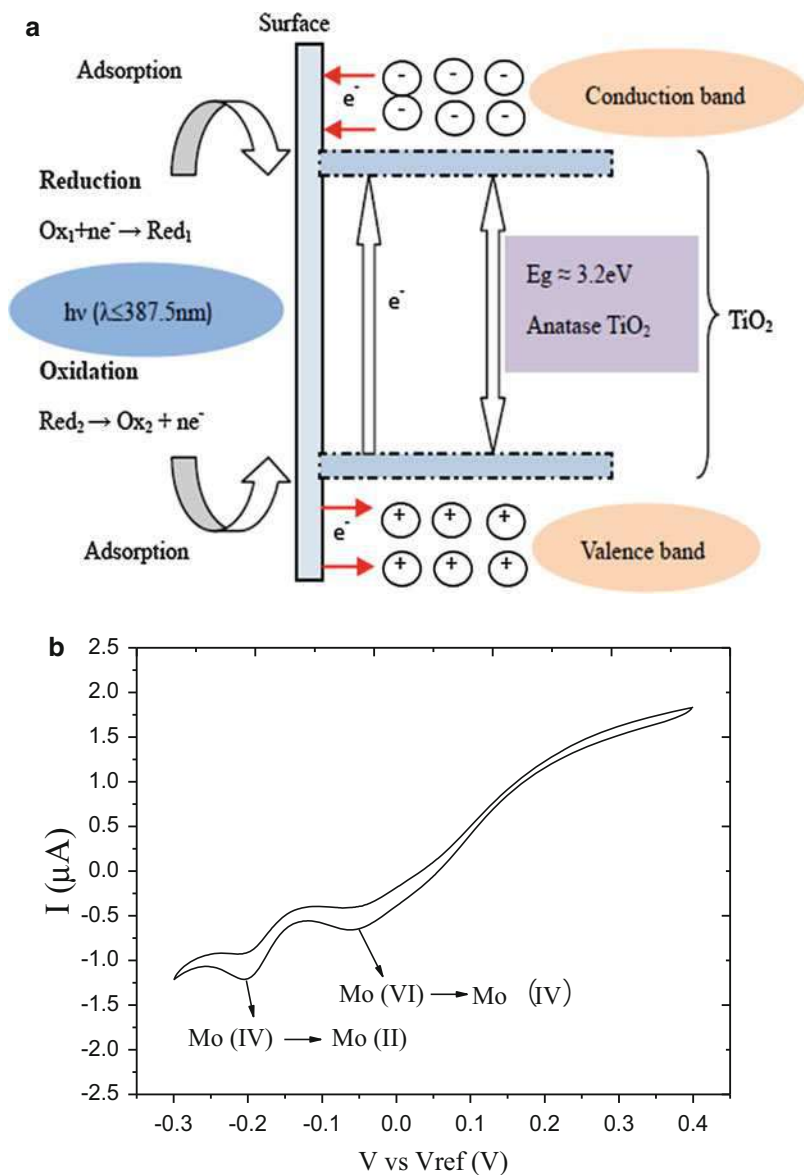
### Working Principle

TP determination includes two processes: digestion pretreatment process and phosphate detection process. The digestion principle of TP is based on thermally assisted UV digestion utilizing  $\text{TiO}_2$  photocatalyst, which has been presented in “[Micro Digestion Reactors](#)” of this chapter. The detection principle of phosphate is based on the reduction of molybdophosphate complex, which has been presented in “[Microelectrode Based on Reduction of Molybdophosphate Complex](#)” of this chapter (Fig. 31).

The schematic of the integrated microreactor for TP determination is shown in Fig. 32. The microreactor includes a digestion pretreatment unit and a phosphate detection unit. An UV-LED array chip is covered on the microfluidic channels fabricated by polydimethylsiloxane (PDMS), in which a  $\text{TiO}_2$  film is prepared on the substrate surface. A micro heater and a Pt micro temperature sensor are assembled on the backside of the digestion area to control the temperature of the digestion process. The electrochemical detection unit is a three-electrode-system including a gold interdigitated ultramicroelectrode array as working electrode (WE) and counter electrode (CE), and Ag/AgCl as reference electrode (RE).

### Device Fabrication

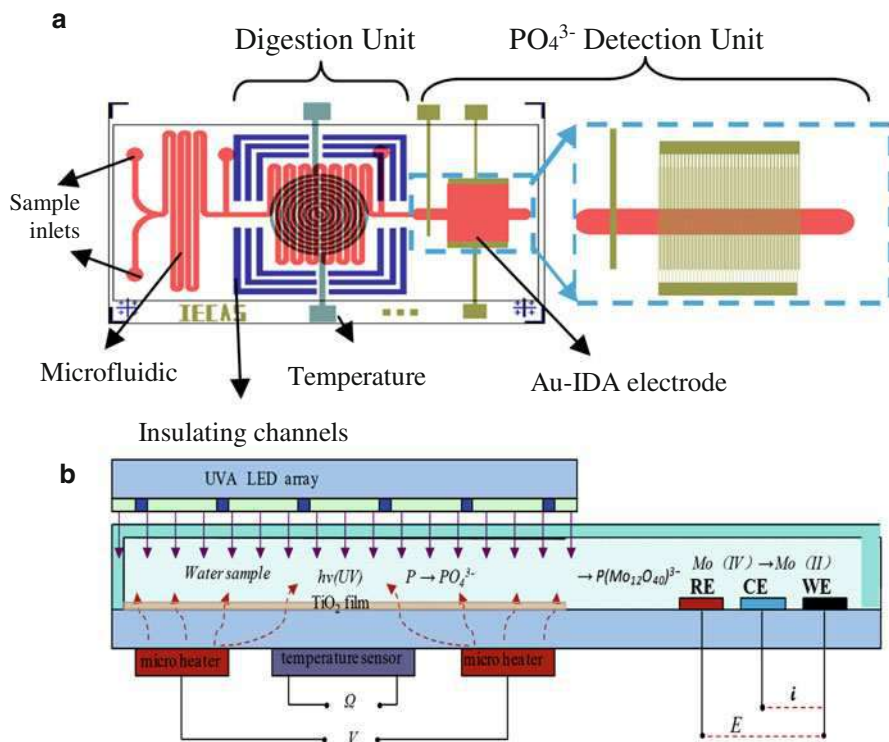
The integrated chip was fabricated with MEMS technique, as shown in Fig. 33. This integrated chip contains two parts: silicon device and PDMS microchannels.  $\text{TiO}_2$  film, Au-IDA electrode, Pt temperature sensor, and insulation channels were fabricated on silicon substrate. And microchannels and reaction reservoir were formed on PDMS membrane by soft-photolithography (Fig. 34).



**Fig. 31** The principle of photocatalytic digestion with  $TiO_2$  (a) and electrochemical determination of TP (b)

### Analytical Performance of Integrated Chip System

The integrated chip performed a fully functional process for TP determination (TP digestion and phosphate detection). The mixed solutions of 4 mg/L  $C_3H_7Na_2O_6P$  and molybdate sodium were injected into the microchannels, digested in digestion



**Fig. 32** The mask layout (a) and schematic (b) of the integrated on-chip electrochemical microreactor

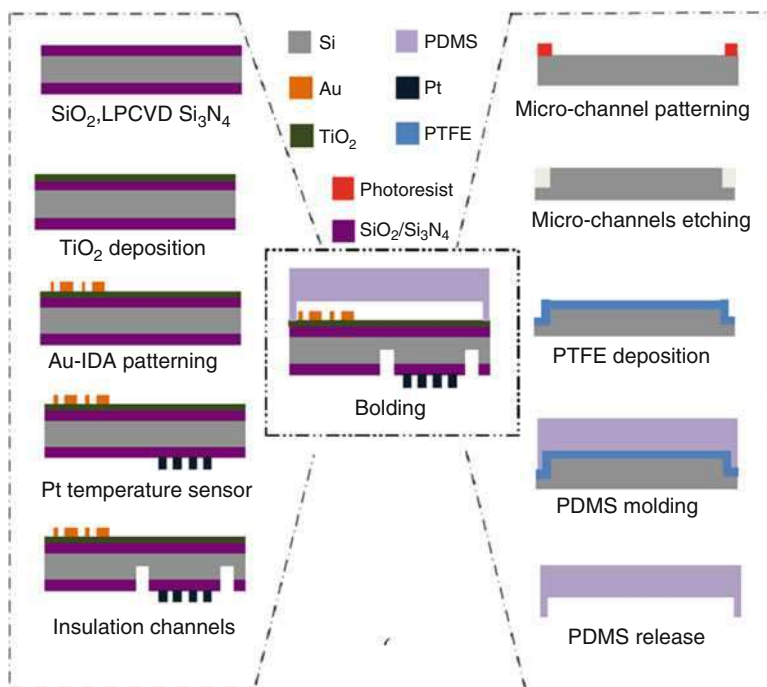
area, and pushed to the electrochemical 3-electrode system for phosphate detection. The digestion process was carried out at two digestion temperatures (25 °C and 40 °C) for 1 h. The results are shown in Table 6. The detected concentration of TP is lower than the nominal value, and the possible reason could be the interference of undigested solution in the movement from digestion area to the detection area.

## Buoy-Based Multiparameter Monitoring Systems

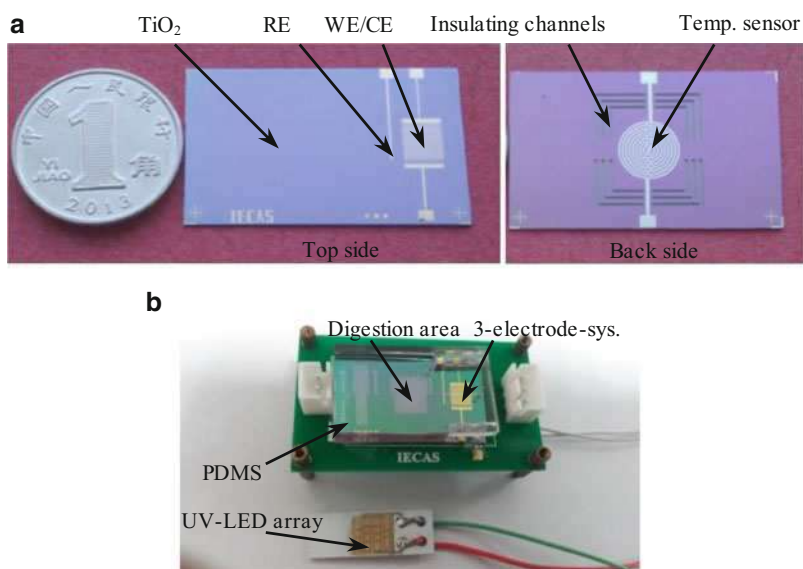
The traditional way for online monitoring of water quality is based on station which is a building with big analysis facilities, and this way has the advantages of high accuracy, and high detection stability. However, the costs of equipment and maintenance are very high, and this station-based monitoring approach is not suitable for large-area application. Buoy-based monitoring system, as an online monitoring method, is an effective solution due to its advantages such as remote monitoring, low power consumption, low cost, and small in-size.

Xia group developed a buoy-based online monitoring system which can monitor the concentration of TP, TN, COD,  $\text{NH}_3\text{-N}$ , and several heavy metal ions ( $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ ,  $\text{Zn}^{2+}$ )





**Fig. 33** The fabrication process of the integrated on-chip microreactor



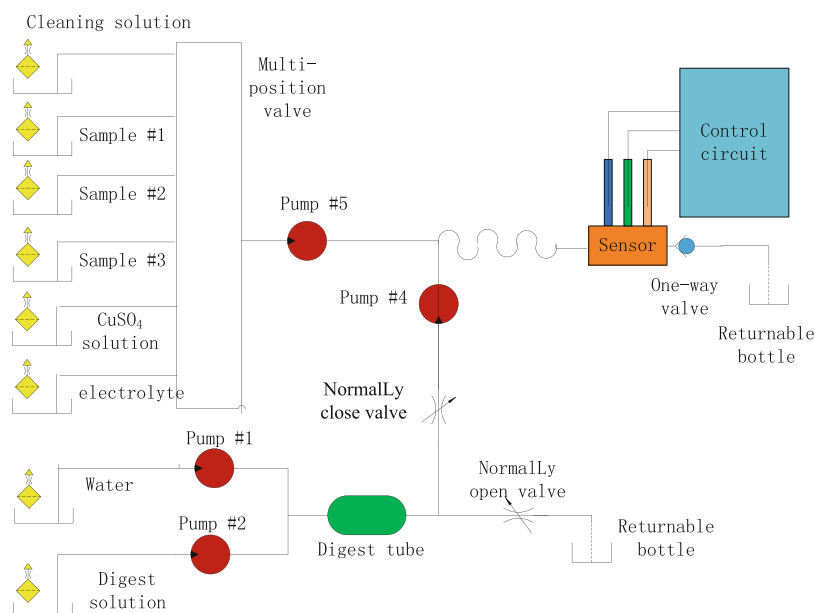
**Fig. 34** The photograph of the *top* side (a), *back* side (b) of the fabricated microchip, and the packaged integrated microreactor for TP determination



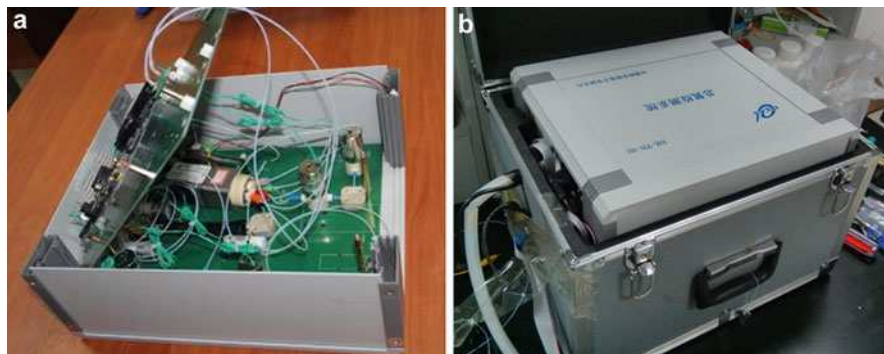
**Table 6** The results of TP determination (mg/L P)

$C_0$	Digestion condition: 40 °C, 1 h		Digestion condition: 25 °C, 1 h	
	$C_x$	RSD	$C_x$	RSD
4.0	1.82	54.5%	0.51	87.25%
	1.54	61.5%	0.21	94.75%

$C_x$ : Detected concentration;  $C_0$ : Real concentration

**Fig. 35** Schematic of the total nitrogen (TN) detection system

in surface freshwaters for the first time. The pretreatment processes for the detections of TP and TN are based on the thermally assisted UV digestion, and electrochemical methods based on reduction of molybdophosphate complex, and copper nanoparticles catalysis were used for the detection of phosphate and nitrate, respectively. COD detection is based on the oxidation-reduction reaction using a hydroxyl-activated glassy carbon electrode.  $\text{NH}_3\text{-N}$  detection is based on a gas-liquid separation and colorimetric method. Heavy metal ions detection is based on the light addressable potentiometric sensors (LAPS) combined with electrochemical sensors. A main program was made to control the logic and working sequence of the buoy system, and the data acquired from the sensors was sent to the remote terminal server by GPRS wireless technique. People could check the real-time data and analyze the pollution trends of these parameters, as well as predict the blooming of blue algae of lakes. The system is powered by solar energy, and it could work for several weeks before the next routine maintenance due to the small amount of reagent consumption.



**Fig. 36** Pictures of the TP/TN detection system. (a) inside, (b) connected with cables and tubes

**Fig. 37** Picture of the buoy floating on a lake



The schematics of typical electrochemical detection system, for example, TN detection system, is shown in Fig. 35. In this system, solenoid operated micropump, solenoid operated isolation valves, and multiposition valves are used to control the flow of the water samples and reagents. Thermally assisted UV catalytic digestion method is used to convert various forms of nitrogen into nitrate. A Pt/Pt/Ag-AgCl electrochemical three-electrode-system is employed as the sensing unit of nitrate. The packaged TP/TN system is shown in Fig. 36. This buoy-based system was put on a lake of Jiangsu province of China, and it kept working for several months, as shown in Fig. 37.

## Conclusions

This chapter gives a brief introduction on the electrochemical microsensors and systems for the detections of several key indicators of water pollution, such as total phosphorus (TP), total nitrogen (TN), and heavy metal ions ( $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Hg}^{2+}$ ). The related sensitivity-enrichment methods, pretreatment processes are demonstrated to promote the performance of the sensors and systems. An integrated microfluidic system has been developed for the fully functional detection of TP, and a buoy-based automated multiparameters analytical system has been developed for the online monitoring of TP, TN, COD,  $\text{NH}_3\text{-N}$ , and several heavy metal ions ( $\text{Cu}^{2+}$ ,  $\text{Pb}^{2+}$ ,  $\text{Zn}^{2+}$ ). The developed microsensors and systems provide a useful approach and a meaningful reference for the development of water environment monitoring.

**Acknowledgments** This work is supported by the National Basic Research Program of China (973 Program, Grant No. 2015CB352100).

---

## References

- Adlnasab L, Ebrahimzadeh H (2013) A novel salt-controlled homogenous ionic liquid phase microextraction based on the salting out effect and optimization of the procedure using the experimental design methodology. *Anal Methods* 5:5165–5171
- Afkhami A, Khoshsafar H, Bagheri H et al (2014) Construction of a carbon ionic liquid paste electrode based on multi-walled carbon nanotubes-synthesized Schiff base composite for trace electrochemical detection of cadmium. *Mater Sci Eng C* 35:8–14
- Akyuz M, Ata S (2009) Determination of low level nitrite and nitrate in biological, food and environmental samples by gas chromatography-mass spectrometry and liquid chromatography with fluorescence detection. *Talanta* 79:900–904
- Al-Saidi HM, Emara AAA (2011) The recent developments in dispersive liquid-liquid microextraction for preconcentration and determination of inorganic analytes. *J Saudi Chem Soc* 18(6):745–761
- American Public Health Association (2000) Standard methods for the examination of water and wastewater. American Public Health Association, Washington, DC
- Amio AK, Agaosa YN (2008) 1-butyl-3-methylimidazolium hexafluorophosphate ionic liquid as a new solvent for the determination of lead (II) and cadmium (II) by anodic stripping voltammetry after extraction of the iodide complexes. *Anal Sci* 24(10):1363–1367
- Andruch V, Balogh IS, Kocúrová et al (2013) Five years of dispersive liquid-liquid microextraction. *Appl Spectrosc Rev* 48:37–41
- Ao XJ, Zhang XN (2009) Determining phosphate ions in precipitation with ion chromatography. *J Heilongjiang Environ* 33:37–39
- Aran HC, Salamon D, Rijnaarts T (2011) Porous photocatalytic membrane microreactor (P2M2): a new reactor concept for photochemistry. *J Photochem Photobiol Chem* 225(1):36–41
- Aravamudhan S, Bhansali S (2008) Development of micro-fluidic nitrate-selective sensor based on doped-polypyrrole nanowires. *Sensors Actuators B Chem* 132(2):623–630
- Attia AJ (2008) Photocatalytic iodometry over naked and sensitized zinc oxide. *Natl J Chem* 32:599–609
- Bagal-Kestwal D, Karve MS, Kakade B et al (2008) Invertase inhibition based electrochemical sensor for the detection of heavy metal ions in aqueous system: application of ultra-microelectrode to enhance sucrose biosensor's sensitivity. *Biosens Bioelectron* 24(4):657–664
- Baghdadi M, Shemirani F (2008) Cold-induced aggregation microextraction: a novel sample preparation technique based on ionic liquids. *Anal Chim Acta* 613:56–63

- Baghdadi M, Shemirani F (2009) In situ solvent formation microextraction based on ionic liquids: a novel sample preparation technique for determination of inorganic species in saline solutions. *Anal Chim Acta* 634(2):186–191
- Bagheri H, Afkhami A, Khoshsafar H et al (2015) Simultaneous electrochemical sensing of thallium, lead and mercury using a novel ionic liquid/graphene modified electrode. *Anal Chim Acta* 870:56–66
- Bai Y, Tong J, Bian C et al (2012) Micro cobalt electrodes for detection of total phosphorus in water. *Micro Nano Lett* 7(12):1176–1179
- Bai Y, Tong J, Wang J et al (2014a) Electrochemical microsensor based on gold nanoparticles modified electrode for total phosphorus determinations in water. *Nanobiotechnol IET* 8(1):31–36
- Bai Y, Tong J, Bian C et al (2014b) An electrochemical microsensor based on nano modified gold electrode for total phosphorus determination in water. *IET Nanobiotechnol* 8(1):31–36
- Bai Y, Tong J, Sun J et al (2014c) Microfluidic chip with interdigitated ultra-microelectrode array for total phosphorus detection. *Micro Nano Lett* 9(12):862–865
- Baldrianova L, Svancara I, Economou A et al (2006) Anodic stripping voltammetry at in situ bismuth-plated carbon and gold microdisc electrodes in variable electrolyte content unstirred solutions. *Anal Chim Acta* 580(1):24–31
- Berchmans S, Karthikeyan R, Gupta S et al (2011) Glassy carbon electrode modified with hybrid films containing inorganic molybdate anions trapped in organic matrices of chitosan and ionic liquid for the amperometric sensing of phosphate at neutral pH. *Sensors Actuators B Chem* 160(1):1224–1231
- Berduque A, Lanyon YH, Beni V et al (2007) Voltammetric characterisation of silicon-based microelectrode arrays and their application to mercury-free stripping voltammetry of copper ions. *Talanta* 71(3):1022–1030
- Bernalte E, Marín Sánchez C, Pinilla Gil E et al (2012) Gold nanoparticles-modified screen-printed carbon electrodes for anodic stripping voltammetric determination of mercury in ambient water samples. *Sensors Actuators B Chem* 161(1):669–674
- Berton P, Vera-candiotti L, Wuilloud RG (2013) A microextraction procedure based on an ionic liquid as an ion-pairing agent optimized using a design of experiments for chromium species separation and determination in water samples. *Anal Methods* 5:5065–5073
- Bonfil Y, Brand M (2000) Trace determination of mercury by anodic stripping voltammetry at the rotating gold electrode. *Anal Chim Acta* 424(1):65–76
- Bozzi A, Guasaquillo I, Kiwi J et al (2004) Accelerated removal of cyanides from industrial effluents by supported TiO<sub>2</sub> photo-catalysts. *Appl Catal B Environ* 51(3):203–211
- Brillas E, Bastida RM, Llosa E et al (1995) Electrochemical destruction of aniline and 4-chloroaniline for wastewater treatment using a carbon-PTFEO-fed cathode. *J Electrochem Soc* 142(6):1733–1741
- Bulatov AV, Tsapko AA, Moskvina LN (2009) Photometric cyclic-injection determination of phosphate and silicate ions simultaneously present in aqueous solutions. *J Anal Chem* 64(9):580–584
- Carpenter NG, Hodgson AWE, Pletcher D (1997) Microelectrode procedures for the determination of silicate and phosphate in waters—fundamental studies. *Electroanalysis* 9(17):1311–1317
- Chen DH, Ye X, Li K (2005) Oxidation of PCE with a UV LED photocatalytic reactor. *Chem Eng Technol* 28(1):95–97
- Clescerl LS, Greenberg AE, Eaton AD (eds) (1998) Standard methods for the examination of water and wastewater. American Public Health Association, American Water Works Association, Water Environment Federation, Washington, DC
- Cugnet C, Zaouak O, René A et al (2009) A novel microelectrode array combining screen-printing and femtosecond laser ablation technologies: development, characterization and application to cadmium detection. *Sensors Actuators B Chem* 143(1):158–163
- Czop E, Economou A, Bobrowski A (2011) A study of in situ plated tin-film electrodes for the determination of trace metals by means of square-wave anodic stripping voltammetry. *Electrochim Acta* 56(5):2206–2212

- Dai X, Compton RG (2005) Gold nanoparticle modified electrodes show a reduced interference by copper (II) in the detection of arsenic(III) using anodic stripping voltammetry. *Electroanalysis* 17(14):1325–1330
- Dai X, Nekrassova O, Hyde ME et al (2004) Anodic stripping voltammetry of arsenic(III) using gold nanoparticle-modified electrodes. *Anal Chem* 76(19):5924–5929
- Daneshvar N, Rabbani M, Modirshahla N et al (2005) Photooxidative degradation of acid red 27 in a tubular continuous-flow photoreactor: influence of operational parameters and mineralization products. *J Hazard Mater B* 118(1–3):155–160
- Daniel D, Grutz IGR (2007) Microfluidic cell with a TiO<sub>2</sub>-modified gold electrode irradiated by an UV-LED for in situ photocatalytic decomposition of organic matter and its potentiality for voltammetric analysis of metal ions. *Electrochem Commun* 9(3):522–528
- Davies TJ, Compton RG (2005) The cyclic and linear sweep voltammetry of regular and random arrays of microdisc electrodes: theory. *J Electroanal Chem* 585(1):63–82
- Dima GE, Beltramo GL, Koper MTM (2005) Nitrate reduction on single-crystal platinum electrodes. *Electrochim Acta* 50(21):4318–4326
- Dohyun K, Ira BG, Jack WJ (2007) Chronocoulometric determination of nitrate on silver electrode and sodium hydroxide electrolyte. *Analyst* 132:350–357
- Dorjpalam E, Takahashi M, Tokuda Y et al (2005) Controlling carrier density and its effect on I–V characteristics of the anatase–TiO<sub>2</sub> thin films prepared by a sputter deposition method. *Thin Solid Films* 483(1–2):147–151
- Ebrahimpour B, Yamini Y, Esrafil A (2012) Emulsification liquid phase microextraction followed by on-line phase separation coupled to high performance liquid chromatography. *Anal Chim Acta* 751:79–85
- EGA (2007) [http://www.nanoscience.co.jp/surface\\_analysis/pdf/icp-oes-ms-detection-limit-guide-ance-BR023.pdf](http://www.nanoscience.co.jp/surface_analysis/pdf/icp-oes-ms-detection-limit-guide-ance-BR023.pdf). Accessed 23 June 2017
- Elmolla ES, Chaudhuri M (2010) Photocatalytic degradation of amoxicillin, ampicillin and cloxacillin antibiotics in aqueous solution using UV/TiO<sub>2</sub> and UV/H<sub>2</sub>O<sub>2</sub>/TiO<sub>2</sub> photocatalysis. *Desalination* 252(1–3):46–52
- Endres F, MacFarlane D, Abbott A (2008) Electrodeposition from ionic liquid. *Trans Inst Met Finish* 86(4):182
- Fu B, Ure AM, West TS (1983) Column cementation on aluminium powder as a preconcentration technique for trace element determinations by spark-source mass spectrometry: part 1. Copper, lead, ruthenium and the noble metals. *Anal Chim Acta* 152:95–104
- Galloway JN, Townsend AR, Erisman JW et al (2008) Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science* 320(5878):889–892
- Gamboa JCM, Peña RC, Paixão TRLC et al (2009) A renewable copper electrode as an amperometric flow detector for nitrate determination in mineral water and soft drink samples. *Talanta* 80:581–585
- Gao S, Jin H, You J et al (2011) Ionic liquid-based homogeneous liquid–liquid microextraction for the determination of antibiotics in milk by high-performance liquid chromatography. *J Chromatogr A* 1218(41):7254–7263
- Gao S, Yang X, Yu W et al (2012) Ultrasound-assisted ionic liquid/ionic liquid-dispersive liquid–liquid microextraction for the determination of sulfonamides in infant formula milk powder using high-performance liquid chromatography. *Talanta* 99:875–882
- Garcia JC, Oliveira JL, Silva AEC et al (2007) Comparative study of the degradation of real textile effluents by photocatalytic reactions involving UV/TiO<sub>2</sub>/H<sub>2</sub>O<sub>2</sub> and UV/Fe<sup>2+</sup>/H<sub>2</sub>O<sub>2</sub> systems. *J Hazard Mater* 147(1–2):105–110
- Gardner RD, Zhou A, Zufelt NA (2009) Development of a microelectrode array sensing platform for combination electrochemical and spectrochemical aqueous ion testing. *Sensors Actuators B Chem* 136(1):177–185
- Gholizadeh A, Shahrokhian S, Zad AI et al (2012) Fabrication of sensitive glutamate biosensor based on vertically aligned CNT nanoelectrode array and investigating the effect of CNTs density on the electrode performance. *Anal Chem* 84(14):5932–5938
- Ghosha JP, Suib R, Langfordb CH et al (2009) A comparison of several nanoscale photocatalysts in the degradation of a common pollutant using LEDs and conventional UV light. *Water Res* 43(18):4499–4506

- Giacomino A, Abollino O, Malandrino M et al (2008) Parameters affecting the determination of mercury by anodic stripping voltammetry using a gold electrode. *Talanta* 75(1):266–273
- Gibbon-Walsh K, Salaün P, van den Berg CMG (2010) Arsenic speciation in natural waters by cathodic stripping voltammetry. *Anal Chim Acta* 662(1):1–8
- Godino N, Borrisé X, Muñoz FX et al (2009) Mass transport to nanoelectrode arrays and limitations of the diffusion domain approach: theory and experiment. *J Phys Chem C* 113(25):11119–11125
- Goldberger ML, Watson KM (2004) Collision theory. Courier Dover Publications, Mineola
- Guettaï N, Amar HA (2005) Photocatalytic oxidation of methyl orange in presence of titanium dioxide in aqueous suspension. Part II: kinetics study. *Desalination* 185(1–3):439–448
- Haeberle S, Zengerle R (2007) Microfluidic platforms for lab-on-a-chip applications. *Lab Chip* 7(9):1094–1110
- Helaleh MIH, Korenaga T (2000) Ion chromatographic method for simultaneous determination of nitrate and nitrite in human saliva. *J Chromatogr B* 744(2):433–437
- Herrmann JM (2010) Photocatalysis fundamentals revisited to avoid several misconceptions. *Appl Catal B Environ* 99(3–4):461–468
- Hezar T, Fajerweg K, Evrard D et al (2012) Gold nanoparticles electrodeposited on glassy carbon using cyclic voltammetry: application to Hg(II) trace analysis. *J Electroanal Chem* 664:46–52
- Hinkamp S, Schwedt G (1990) Determination of total phosphorus in waters with amperometric detection by coupling of flow-injection analysis with continuous microwave oven digestion. *Anal Chim Acta* 236:345–350
- Hitchman M, Tian F (2002) Studies of TiO<sub>2</sub> thin films prepared by chemical vapour deposition for photocatalytic and photoelectrocatalytic degradation of 4-chlorophenol. *J Electroanal Chem* 538–539:165–172
- Hu R, Luo X, Zheng H (2012) Design of a novel freeform lens for LED uniform illumination and conformal phosphor coating. *Opt Express* 20(13):13727–13737
- Hunt JM, Lewan MD, Hennes RJC (1991) Modeling oil generation with time-temperature index graphs based on the Arrhenius equation. *Am Assoc Pet Geol* 75(4):795–807
- Injang U, Noyrod P, Siangproh W et al (2010) Determination of trace heavy metals in herbs by sequential injection analysis-anodic stripping voltammetry using screen-printed carbon nanotubes electrodes. *Anal Chim Acta* 668(1):54–60
- Isoda T, Makimoto H, Imanag H et al (2007) Development of a source-drain electrode coated with an insulation layer for detecting concentration changes in a nitrate ion solution. *Sensors Actuators B Chem* 123(2):805–815
- Jakmuneea J, Junsomboon J (2008) Determination of cadmium, lead, copper and zinc in the acetic acid extract of glazed ceramic surfaces by anodic stripping voltammetric method. *Talanta* 77(1):172–175
- Jena BK, Raj CR (2008) Gold nanoelectrode ensembles for the simultaneous electrochemical detection of ultratrace arsenic, mercury, and copper. *Anal Chem* 80(13):4836–4844
- Jimidar M, Hartmann C, Cousement N et al (1995) Determination of nitrate and nitrite in vegetables by capillary electrophoresis with indirect detection. *J Chromatogr A* 706(1–2):479–492
- Jońca J, Fernández VL, Thouron D et al (2011) Phosphate determination in seawater: toward an autonomous electrochemical method. *Talanta* 87:161–167
- Jung W, Jang A, Bishop PL et al (2011) A polymer lab chip sensor with microfabricated planar silver electrode for continuous and on-site heavy metal measurement. *Sensors Actuators B Chem* 155(1):145–153
- Justyna J, Violeta LF, Danièle T et al (2011) Phosphate determination in seawater: toward an autonomous electrochemical method. *Talanta* 87:161–167
- Justyna J, William G, Carole B et al (2013) Reagentless and silicate interference free electrochemical phosphate determination in seawater. *Electrochim Acta* 88:165–169
- Khan S, Kazi TG, Soylyak M (2015) Ionic liquid-based ultrasound-assisted emulsification micro-extraction of cadmium in biological samples: optimization by a multivariate approach. *Anal Lett* 48(11):1751–1766
- Kim K, Kim KL, Shin KS (2012) Selective detection of aqueous nitrite ions by surface-enhanced Raman scattering of 4-aminobenzenethiol on Au. *Analyst* 137:3836–3840

- Kima SB, Hongb SC (2002) Kinetic study for photocatalytic degradation of volatile organic compounds in air using thin film TiO<sub>2</sub> photocatalyst. *Appl Catal B Environ* 35(4):305–315
- Kocúrová L, Balogh IS, Jana Š et al (2012) Recent advances in dispersive liquid–liquid micro-extraction using organic solvents lighter than water. A review. *Microchem J* 102:11–17
- Kodamatani H, Yamazaki S, Saito K et al (2009) Selective determination method for measurement of nitrite and nitrate in water samples using high-performance liquid chromatography with post-column photo-chemical reaction and chemiluminescence detection. *J Chromatogr A* 1216(15):3163–3167
- Kokkinos C, Economou A, Raptis I (2011) Disposable lithographically fabricated bismuth microelectrode arrays for stripping voltammetric detection of trace metals. *Electrochem Commun* 13(5):391–395
- Konstantinou IK, Albanis TA (2004) TiO<sub>2</sub>-assisted photocatalytic degradation of azo dyes in aqueous solution: kinetic and mechanistic investigations: a review. *Appl Catal B Environ* 49(1):1–14
- Lachheb H, Puzenat E, Houas A et al (2002) Photocatalytic degradation of various types of dyes (Alizarin S, Crocein Orange G, Methyl Red, Congo Red, Methylene Blue) in water by UV-irradiated titania. *Appl Catal B Environ* 39(1):75–90
- Laidler KJ (1984) The development of the Arrhenius equation. *J Chem Educ* 61(6):494–498
- Lau OW, Ho SY (1993) Simultaneous determination of traces of iron, cobalt, nickel, copper, mercury and lead in water by energy-dispersive x-ray fluorescence spectrometry after preconcentration as their piperazino-1,4-bis(dithiocarbamate) complexes. *Anal Chim Acta* 280(2):269–277
- Lee JH, Lee WH, Bishop PL et al (2009) A cobalt-coated needle-type microelectrode array sensor for in situ monitoring of phosphate. *J Micromech Microeng* 19(2):025022
- Lee DM, Yun HJ, Yu S et al (2012) Design of an efficient photocatalytic reactor for the decomposition of gaseous organic contaminants in air. *Chem Eng J* 187:203–209
- Leopold K, Foulkes M, Worsfold PJ (2009) Preconcentration techniques for the determination of mercury species in natural waters. *Trends Anal Chem* 28(4):426–435
- Li YQ, Delsing ACA, With D et al (2005) Luminescence properties of Eu<sup>2+</sup>-activated alkaline-earth silicon-oxynitride MSi<sub>2</sub>O<sub>2</sub>-δN<sub>2</sub>+ 2/3δ (M= Ca, Sr, Ba): a promising class of novel LED conversion phosphors. *Chem Mater* 17(12):3242–3248
- Li Y, Sun J, Bian C et al (2011a) A microfluidic sensor chip with renewable in-situ copper modified microelectrode for continuous monitoring of nitrate. Paper presented at Transducers'11, Beijing, 5–9 June 2011
- Li F, Xue M, Ma X (2011b) Facile patterning of reduced Graphene oxide film into microelectrode array for highly sensitive sensing. *Anal Chem* 83(16):6426–6430
- Li Z, Xia S, Bian C et al (2015) Salt-induced ionic liquid dispersive liquid–liquid microextraction and filter separation. *Anal Methods* 8(5):1096–1102
- Li Z, Xia S, Wang J et al (2016) Determination of trace mercury in water based on N-octylpyridinium ionic liquids preconcentration and stripping voltammetry. *J Hazard Mater* 301:206–213
- Lin M, Cho MS, Choe WS et al (2009) Electrochemical analysis of copper ion using a Gly–Gly–His tripeptide modified poly(3-thiopheneacetic acid) biosensor. *Biosens Bioelectron* 25(1):28–33
- Lindstrom H, Wootton R, Iles A (2007) High surface area titania photocatalytic microfluidic reactors. *AIChE J* 53(3):695–702
- Liu AC, Chen DC, Lin CC et al (1999) Application of cysteine monolayers for electrochemical determination of sub-ppb copper(II). *Anal Chem* 71(8):1549–1552
- Lu PL, Huang KS, Jiang SJ (1993) Determination of traces of copper, cadmium and lead in biological and environmental samples by flow-injection isotope dilution inductively coupled plasma mass spectrometry. *Anal Chim Acta* 284(1):181–188
- Lu G, Wu XG, Lan YH et al (1999) Studies on 1:12 phosphomolybdic heteropoly anion film modified carbon paste electrode. *Talanta* 49:511–515
- Mace KA, Duce RA (2002) The use of UV photo-oxidation for the determination of total nitrogen in rainwater and water-extracted atmospheric aerosol. *Atmos Environ* 36(39–40):5937–5946

- Maher W, Woo L (1998) Procedures for the storage and digestion of natural waters for the determination of filterable reactive phosphorus, total filterable phosphorus and total phosphorus. *Anal Chim Acta* 375(1–2):5–47
- Maher W, Krikowa F, Wruck D et al (2002) Determination of total phosphorus and nitrogen in turbid waters by oxidation with alkaline potassium peroxodisulfate and low pressure microwave digestion, autoclave heating or the use of closed vessels in a hot water bath: comparison with Kjeldahl digestion. *Anal Chim Acta* 463(2):283–293
- Martinis EM, Bertón P, Olsina RA et al (2009) Trace mercury determination in drinking and natural water samples by room temperature ionic liquid based-preconcentration and flow injection-cold vapor atomic absorption spectrometry. *J Hazard Mater* 167(1–3):475–481
- Mavrodineanu R, Schultz JI, Menis O (1973) Accuracy in spectrophotometry and luminescence measurements: proceedings. U.S. National Bureau of Standards, Washington, DC, p P2
- Meruva RK, Meyerhoff ME (1996) Mixed potential response mechanism of cobalt electrodes toward inorganic phosphate. *Anal Chem* 68(13):2022–2026
- Ministry of Environmental Protection of the People's Republic of China (2002) Environmental quality standards for surface water
- Ministry of Environmental protection of the People's Republic of China (2012) Environmental quality report of 2012
- Moiseev MA, Doskolovich LL, Kazanskiy NL (2011) Design of high-efficient freeform LED lens for illumination of elongated rectangular regions. *Opt Express* 19(103):A225–A233
- Moujahid W, Eichelmann-Daly P, Strutwolf J et al (2011) Microelectrochemical systems on silicon chips for the detection of pollutants in seawater. *Electroanalysis* 23(1):147–155
- Noroozifar M, Khorasani-Motlagh M, Taheri A et al (2007) Application of manganese(IV) dioxide microcolumn for determination and speciation of nitrite and nitrate using a flow injection analysis-flame atomic absorption spectrometry system. *Talanta* 71:359–364
- O'Dell JW (ed) (1993) Determination of phosphorus by semi-automated colorimetry, environmental monitoring systems laboratory. Office of Research and Development, USEPA, Cincinnati
- Oms MT, Cerda A, Cerda V (2003) Sequential injection system for on-line analysis of total nitrogen with UV-mineralization. *Talanta* 59(2):319–326
- Opallo M, Lesniewski A (2011) A review on electrodes modified with ionic liquids. *J Electroanal Chem* 656(1–2):2–16
- Orozco J, Fernández-Sánchez C, Jiménez-Jorquera C (2008) Underpotential deposition–anodic stripping voltammetric detection of copper at gold nanoparticle-modified ultramicroelectrode arrays. *Environ Sci Technol* 42(13):4877–4882
- Papaiconomou N, Lee J, Salminen J et al (2008) Selective extraction of copper, mercury, silver, and palladium ions from water using hydrophobic ionic liquids. *Ind Eng Chem Res* 47(15):5080–5086
- Praline J, Guennoc AM, Limousin N et al (2007) ALS and mercury intoxication: a relationship? *Neurol Neurosurg* 109(10):880–883
- Qin J (2013) A high efficiency microfluidic-based photocatalytic microreactor using electrospun nanofibrous TiO<sub>2</sub> as photocatalyst. *Nanoscale* 5:4687–4690
- Raimbault P, Slawyk G, Coste B et al (1990) Feasibility of using an automated colorimetric procedure for the determination of seawater nitrate in the 0 to 100 nM range: examples from field and culture. *Mar Biol* 104(2):347–351
- Rodrigues JA, Rodrigues CM, Almeida PJ et al (2011) Increased sensitivity of anodic stripping voltammetry at the hanging mercury drop electrode by ultracathodic deposition. *Anal Chim Acta* 701(2):152–156
- Roig B, Gonzalez C, Thomas O et al (1999) Measurement of dissolved total nitrogen in wastewater by UV photooxidation with peroxodisulphate. *Anal Chim Acta* 389(1–3):267–274
- Salaüin P, van den Berg CMG (2006) Voltammetric detection of mercury and copper in seawater using a gold microwire electrode. *Anal Chem* 78(14):5052–5060
- Salaüin P, Planer-Friedrich B, van den Berg CMG (2007) Inorganic arsenic speciation in water and seawater by anodic stripping voltammetry with a gold microelectrode. *Anal Chim Acta* 585(2):312–322



- Sandison ME, Anicet N, Glidle A et al (2002) Optimization of the geometry and porosity of microelectrode arrays for sensor design. *Anal Chem* 74(22):5717–5725
- Seebunrueng K, Santaladchaiyakit Y, Srijaranai S (2014) Vortex-assisted low density solvent based demulsified dispersive liquid–liquid microextraction and high-performance liquid chromatography for the determination of organophosphorus pesticides in water samples. *Chemosphere* 103:51–58
- Sempere A, Oliver J, Ramos C (1993) Simple determination of nitrate in soils by second-derivative spectroscopy. *J Soil Sci* 44(4):633–639
- Shen XZ, Liu ZC, Xie SM et al (2009) Degradation of nitrobenzene using titania photocatalyst co-doped with nitrogen and cerium under visible light illumination. *J Hazard Mater* 162(2–3):1193–1198
- Shyla B, Nagendrappa G (2011) A simple spectrophotometric method for the determination of phosphate in soil, detergents, water, bone and food samples through the formation of phosphomolybdate complex followed by its reduction with thiourea. *Spectrochim Acta A Mol Biomol Spectrosc* 78(1):497–502
- Siboni MS, Samadi MT, Yang JK et al (2012) Photocatalytic removal of Cr(VI) and Ni(II) by UV/TiO<sub>2</sub>: kinetic study. *Desalin Water Treat* 40(1–3):77–83
- Silva SD, Shan D, Cosnier S (2004) Improvement of biosensor performances for nitrate determination using a new hydrophilic poly (pyrrole-viologen) film. *Sensors Actuators B Chem* 103(1–2):397–402
- Silva I, Araujo WR, Paixão TRLC et al (2013) Direct nitrate sensing in water using an array of copper microelectrodes from flat flexible cables. *Sensors Actuators B Chem* 188:94–98
- Slavec M, Hocevar SB, Baldrianova L et al (2010) Antimony film microelectrode for anodic stripping measurement of cadmium(II), lead(II) and copper(II). *Electroanalysis* 22(14):1617–1622
- Song Y, Wu L, Li N et al (2015) Utilization of a novel microwave-assisted homogeneous ionic liquid microextraction method for the determination of Sudan dyes in red wines. *Talanta* 135:163–169
- Stanisz E, Werner J, Zgoła-Grześkowiak A (2014a) Liquid-phase microextraction techniques based on ionic liquids for preconcentration and determination of metals. *TrAC Trends Anal Chem* 61:54–66
- Stanisz E, Werner J, Matusiewicz H (2014b) Task specific ionic liquid-coated PTFE tube for solid-phase microextraction prior to chemical and photo-induced mercury cold vapour generation. *Microchem J* 114:229–237
- Sten OE (1998) The phosphate sensor. *Biosens Bioelectron* 13:981–994
- Stortini AM, Moretto LM, Mardeganb A et al (2015) Arrays of copper nanowire electrodes: preparation, characterization and application as nitrate sensor. *Sensors Actuators B* 207(A):186–192
- Suárez R, Horstkotte B, Cerdà V (2014) In-syringe magnetic stirring-assisted dispersive liquid–liquid microextraction for automation and downscaling of methylene blue active substances assay. *Talanta* 130:555–560
- Sun P, Armstrong DW et al (2010) Ionic liquids in analytical chemistry. *Anal Chim Acta* 661(1):1–16
- Svobodova-Tesarova E, Baldrianova L, Stoces M et al (2011) Antimony powder-modified carbon paste electrodes for electrochemical stripping determination of trace heavy metals. *Electrochim Acta* 56(19):6673–6677
- Taylor P, Clevenger WL, Smith BW et al (2006) Trace determination of mercury: a review. *Crit Rev Anal Chem* 27(1):1–26
- Tong J, Dong T, Bian C et al (2015) An integrated photocatalytic microfluidic platform enabling total phosphorus digestion. *J Micromech Microeng* 25(2):025006
- Torma F, Kádár M, Tóth K et al (2008) Nafion<sup>®</sup>/2,2'-bipyridyl-modified bismuth film electrode for anodic stripping voltammetry. *Anal Chim Acta* 619(2):173–182
- Tsuji E, Hirata N, Aoki Y et al (2012) Preparation of non-annealed anatase TiO<sub>2</sub> film on ITO substrate by anodizing in hot phosphate/glycerol electrolyte for dye-sensitized solar cells. *Mater Lett* 91:39–41

- Udnan Y, McKelvie ID, Grace MR et al (2005) Evaluation of on-line preconcentration and flow-injection amperometry for phosphate determination in fresh and marine waters. *Talanta* 66:461–466
- Uğurlu M, Karaoğlu MH (2009) Removal of AOX, total nitrogen and chlorinated lignin from bleached Krsft mill effluents by UV oxidation in the presence of hydrogen peroxide utilizing TiO<sub>2</sub> as photocatalyst. *Environ Sci Pollut Res* 16(3):265–273
- UNEP (United Nations Environment Programme) Collaborating Centre on Energy and Environment (2006) Environmental pollution and its effects. [http://www.uceee.org/environmental\\_Pollution.html](http://www.uceee.org/environmental_Pollution.html). Accessed 23 June 2017
- Urbanová V, Vytřas K, Kuhn A (2010) Macroporous antimony film electrodes for stripping analysis of trace heavy metals. *Electrochem Commun* 12(1):114–117
- Vijian PN, Sadana RS (1980) Determination of lead in drinking waters by hydride generation and atomic-absorption spectroscopy, and three other methods. *Talanta* 27(4):321–326
- Wang WY, Ku Y (2006) Photocatalytic degradation of reactive red 22 in aqueous solution by UV-LED radiation. *Water Res* 40(12):2249–2258
- Wang Y, Qu JH, Liu HJ (2006a) Preparation and electrochemical properties of the Pd-modified Cu electrode for nitrate reduction in water. *Chin Chem Lett* 17(1):61–64
- Wang L, Bai J, Huang P et al (2006b) Self-assembly of gold nanoparticles for the voltammetric sensing of epinephrine. *Electrochem Commun* 8(6):1035–1040
- Wang J, Bian C, Tong J et al (2012a) L-aspartic acid/L-cysteine/gold nanoparticles modified microelectrode for simultaneous detection of copper and lead. *Thin Solid Films* 520(21):6658–6663
- Wang J, Bian C, Tong J et al (2012b) Simultaneous detection of copper, lead and zinc on tin film/gold nanoparticles/gold microelectrode by square wave stripping voltammetry. *Electroanalysis* 24(8):1783–1790
- Wang JF, Bian C, Tong J et al (2013) Microsensor chip integrated with gold nanoparticles-modified ultramicroelectrode array for improved electroanalytical measurement of copper ions. *Electroanalysis* 25(7):1713–1721
- Wang Z, Zhang L, Li N et al (2014) Onic liquid-based matrix solid-phase dispersion coupled with homogeneous liquid–liquid microextraction of synthetic dyes in condiments. *J Chromatogr A* 1348:52–62
- Ward-Jones S, Banks CE, Simm AO et al (2005) An in situ copper plated boron-doped diamond microelectrode array for the sensitive electrochemical detection of nitrate. *Electroanalysis* 17:1806–1815
- Wei G, Yang Z, Chen C (2003) Room temperature ionic liquid as a novel medium for liquid/liquid extraction of metal ions. *Anal Chim Acta* 488(2):183–192
- Wikipedia (2017) Transducers. <https://en.wikipedia.org/wiki/Transducer>. Accessed 23 June 2017
- Wilson B, Gandhi J, Zhang CL (2011) Analysis of inorganic nitrogen and related anions in high salinity water using ion chromatography with tandem UV and conductivity detectors. *J Chromatogr Sci* 49(8):596–602
- Woo L, Maher W (1995) Determination of phosphorus in turbid waters using alkaline potassium peroxodisulphate digestion. *Anal Chim Acta* 315(1–2):123–135
- Wu H, Guo J, Du L et al (2013) A rapid shaking-based ionic liquid dispersive liquid phase microextraction for the simultaneous determination of six synthetic food colourants in soft drinks, sugar- and gelatin-based confectionery by high-performance liquid chromatography. *Food Chem* 141(1):182–186
- Xi JT, Zhang QF, Xie SH (2011) Fabrication of TiO<sub>2</sub> aggregates by electrospraying and their application in dye-sensitized solar cells. *Nanosci Nanotechnol Lett* 3(5):690–696
- Xiao D, Yuan HY, Li J et al (1995) Surface-modified cobalt-based sensor as a phosphate-sensitive electrode. *J Anal Chem* 67(2):288–291
- Xie FZ, Lin XC, Wu XP et al (2008) Solid phase extraction of lead (II), copper (II), cadmium (II) and nickel (II) using gallic acid-modified silica gel prior to determination by flame atomic absorption spectrometry. *Talanta* 74(4):836–843

- Xu H, Xing S, Zeng L et al (2009) Microwave-enhanced voltammetric detection of copper(II) at gold nanoparticles-modified platinum microelectrodes. *J Electroanal Chem* 625(1):53–59
- Xu X, Su R, Zhao X et al (2011) Ionic liquid-based microwave-assisted dispersive liquid–liquid microextraction and derivatization of sulfonamides in river water, honey, milk, and animal plasma. *Anal Chim Acta* 707(1–2):92–99
- Yamamoto Y, Nishino Y, Ueda K (1985) Determination of trace amounts of copper, lead and zinc in cements by X-ray fluorescence spectrometry after precipitation separation with hexamethylenammonium hexamethylenedithiocarbamate. *Talanta* 32(8):662–664
- Yang J, Chen J, Zhou Y et al (2011) A nano-copper electrochemical sensor for sensitive detection of chemical oxygen demand. *Sensors Actuators B* 153(1):78–82
- Yokoyama Y, Danno T, Haginoya M et al (2009) Simultaneous determination of silicate and phosphate in environmental waters using pre-column derivatization ion-pair liquid chromatography. *Talanta* 79(2):308–313
- Yu LY, Zhang Q, Xu Q et al (2015) Electrochemical detection of nitrate in PM2.5 with a copper-modified carbon fiber micro-disk electrode. *Talanta* 143:245–253
- Zhang XL, Tian Y (2006) Study on sensor of nitrate ions based on polypyrrole nanowires modified electrode. *Chin J Sensors Actuators* 19:309–312
- Zhang L, Jiang X, Wang E et al (2005) Attachment of gold nanoparticles to glassy carbon electrode and its application for the direct electrochemistry and electrocatalytic behavior of hemoglobin. *Biosens Bioelectron* 21(2):337–345
- Zhao RS, Wang X, Zhang LL et al (2011) Ionic liquid/ionic liquid dispersive liquid–liquid microextraction, a new sample enrichment procedure for the determination of hexabromocyclododecane diastereomers in environmental water samples. *Anal Methods* 3:831–836
- Zhou Q, Bai H, Xie G et al (2007) Temperature-controlled ionic liquid dispersive liquid phase micro-extraction. *J Chromatogr A* 1177(1):43–49
- Zou ZW, Han J, Jang A et al (2007) A disposable on-chip phosphate sensor with planar cobalt microelectrodes on polymer substrate. *J Biosens Bioelectron* 22:1902–1907

---

**Part V**

**Microactuators**



# Electrostatic Comb-Driven Actuator for MEMS Relay/Switch

Jiahao Zhao and Yongfeng Gao

## Contents

Introduction .....	908
Literature Review of Electrostatic Comb-Drive MEMS Relays .....	909
Theoretical Analysis of Design of In-Plane MEMS Electrostatic Actuator for Switches .....	911
Optimization Algorithm and Realization of Segmented Driving Electrode .....	916
Optimization Results Analysis .....	920
Experimental Verification of Driving Electrode Optimization .....	923
Structure and Machining of Electrostatic Actuator for MEMS Switch .....	923
Test Results of Driving Electrode Driving Force .....	925
Conclusion .....	927
References .....	927

## Abstract

A method of optimizing the shape of the driving electrode for switching application is proposed. This method increases the driving force, and decreases the driving voltage of the in-plane electrostatic MEMS switches as well. In this method, optimized comb shape is adopted so that better performance is obtained without the deterioration of other aspects of performance: the area of the device and the requirement of the process are similar to the traditional in-plane devices. Compared with the traditional comb tooth electrode, the experimental results show that the pull-down voltage of the optimized driving electrode is 39% lower than that of the traditional comb tooth electrode, and the driving performance is improved obviously. At a line width of 6  $\mu\text{m}$ , a MEMS switch with a critical pull-down voltage of 14 V is implemented.

---

J. Zhao (✉) · Y. Gao  
Tsinghua University, Beijing, China  
e-mail: [falxon@mail.tsinghua.edu.cn](mailto:falxon@mail.tsinghua.edu.cn)

### Keywords

MEMS switch · In-plane · Comb shape · High driving force · Low pull-down voltage

## Introduction

With the development of micro-electro-mechanical system (MEMS), MEMS technology has played an increasingly significant role in many fields. MEMS devices can divide into MEMS sensors and MEMS actuators. MEMS relay/switch is a kind of typical MEMS actuators. MEMS relay is processed by MEMS technology, comparing to traditional electromagnetic relay, it has fast response, low power consumption and miniaturization. Compared with the electronic switch or solid-state relay, MEMS relay has better signal switching quality because it switches signals by physical contact and separation. Therefore, MEMS relay has attractive application prospects and huge potential market in aerospace, communication, test and measurement, automatic control fields and so on.

The utility of MEMS relay need reliable MEMS drive technology, all these technologies have different principles and schemes, but all of them have their own advantages and disadvantages. At present, the main drive principle has following several kinds: electrostatic, electromagnetic, piezoelectric and electro-thermal type. The electromagnetic drive method has the advantage of low drive voltage, however it has high power consumption and needs external magnetic field by adding a permanent magnet processing coil, however, the way will increase the process complexity. The electro-thermal MEMS relay has low drive voltage but high power consumption. The piezoelectric MEMS relay has excellent response speed and reasonable power consumption, but it has weak driving force and complex process. The driving principle of electrostatic is the earliest proposed principle and also the most widely researched MEMS driving principle. Electrostatic MEMS relay has fast response speed, simple structure and driving method, low power consumption. In summary, because of its outstanding advantages, the electrostatic drive is the most widely used driving method, and also the most promising way to achieve industrialization.

Electrostatic MEMS relays/switches are divided into two different structures: the off-surface switches and in-plane switches. The comparison between them is shown in Table 1.

In-plane electrostatic drive MEMS switching process has simple process and design freedom, it can realize industrialization easily. However, the in-plane

**Table 1** Comparison of off-surface and in-plane switches

Type	Advantages	Disadvantages
Off-surface	Low driving voltage	Process is complex
In-plane	Simple process and design freedom	Relatively high driving voltage

electrostatic drive MEMS has relatively high driving voltage, optimizing the shape of the driving electrode is a way to improve driving performance. Electrostatic comb-drive MEMS relays occupy an important position in in-plane drive MEMS switches. Optimizing the electrode shape of comb finger can reduce the driving voltage and improve its driving performance.

---

## Literature Review of Electrostatic Comb-Drive MEMS Relays

As described above, electrostatic drive technology has fast response speed, simple structure and driving method, low power consumption, but it is difficult to balance the driving force and driving distance. The driving distance of the in-plane electrostatic MEMS switch is defined by the photolithography pattern, and it is difficult to achieve a very small line width, so the driving force is relatively small and the driving voltage is relatively high. This problem exists not only in switching devices, but also in resonant devices. In order to reduce the driving voltage, a number of methods have been proposed in the study of MEMS electrostatic actuation. Optimizing the electrode shape to build different types of comb fingers is an effective method to reduce the driving voltage.

In 1992, the Tokyo Institute of IBM proposed a submicron gap comb-drive microactuator design. They designed the step comb finger and realized a submicron gap between the opposed comb electrodes in the operating state, which reduced the driving voltage greatly (Hirano et al. 1992). There was a noteworthy technical detail, the comb gap after releasing of the etch is large. The researchers inserted the comb fingers into a smaller gap through a *Post-release Positioning* method, which pre-biased the comb driver by an auxiliary electrostatic drive and permanently fixed the bias using the welding action of the short-circuit current. The folded beam width was 2  $\mu\text{m}$ , beam length was 200  $\mu\text{m}$ , submicron working gap minimum reached 0.2  $\mu\text{m}$ . It used only 12 pairs of comb fingers to achieve 11.1 V driving voltage at 3.6  $\mu\text{m}$  driving distance. The method of reducing the driving voltage in this study was to optimize the driving electrode shape and reduce the return stiffness, however its *Post-release Positioning* process was complex.

In 1996, the University of Twente researchers developed a low-voltage comb-drive electrostatic actuator, the device had a pull-down voltage of 20 V, a pull-down critical distance of 40  $\mu\text{m}$ , and 136 pairs of comb finger with gap distance of 2.2  $\mu\text{m}$ . The width of the supporting folded beam was 4.8  $\mu\text{m}$  and the length of single folded beam was 500  $\mu\text{m}$ . This study was focused on the driving applications of large displacement, its low drive voltage is derived from low stiffness of spring beam.

In 1997, Researchers at the Samsung Research and Development Center and the Korea Institute of Science and Technology (KAIST) developed a comb-drive MEMS gyroscope. They also adopted the step comb design. The method to reduce driving voltage of this study was to optimize the electrode shape too.

In 1998, researchers at Cornell University reported that, changing the shape of contact can optimize the driving force characteristics, including linearity (Ye et al. 1998). In the same year, researchers from Griffith University in Australia established their own theoretical model for *Angled Comb Finger*. They made a detailed theoretical analysis and carried out some experiments to verify it. This study proposed a force density as an indicator of the comb-drive performance (Rosa and Harrison 1998a, b). These two studies are the earliest research on the quantitative research of driving electrode shape optimization in the literatures.

In 2003, Brian D. Jensen and others, at University of Michigan, issued a paper for analysis and experimental verification for different singular comb finger. This research attained the analytic relationship between the comb-shape function and driving force function under the certain assumptions. Experiments showed that the theory and experiment were in good agreement with the hypothesis (Jensen et al. 2003). In resonant applications, the drive bias can be made to function as a modulation system stiffness by different comb finger shapes. The analytical relationship between the driving electrode shape and driving force under medium and long stroke is given under certain conditions, which is significant for the design of driving electrode shape. The analytic functions such as rational fraction (including fractional, cubic and quadratic), linear function, piecewise function are used as the mathematical description of the shape of the driving electrode, and the analytical relation is verified.

In 2009, Researchers at the University of Twente in the Netherlands and researchers at the IBM Institute used a stepped finger shape to improve comb-drive performance at large displacements (Engelen et al. 2009). This study is for long stroke-driven applications.

In 2009, Researchers at the National University of Seoul in Seoul (NUS) developed an in-plane RF MEMS switch with 15 V driving force using 1,200 pairs of comb fingers, whose process gap is 2.5  $\mu\text{m}$ . The stroke of this finger shape, called *double stop*, is very short compared with the previous stepped finger shape, which matches the operation characteristics of the switch. This study was focused on the optimization of short-stroke electrode shapes.

Summarizing the above literature, the current method to reduce the electrostatic drive voltage of the in-plane structure mainly includes increasing the number of driving electrodes (comb finger number), reducing the stiffness of the return structure, reducing the driving gap and optimizing the electrode shape. Reducing the drive distance will be limited by process capability and yield. The process capability determines the limit of the drive distance, i.e., the smaller the distance, the lower the yield. Increasing the number of drive electrodes will undoubtedly increase the area of a single device, leading to an increase of cost. Reducing the stiffness of the return structure reduces the restoring force and stability of the actuator, especially in applications requiring vibration and shock resistance, meanwhile, too small of the return stiffness in the switching application will also bring adhesion problems. Optimizing the finger shape to increase the driving force, which will not increase the device area, is the best choice to reduce the driving voltage of comb-drive actuator for MEMS relays/switches.



At present, the electrode optimization methods can be divided into two categories: qualitative optimization and quantitative optimization.

Qualitative optimization is based on the characteristics of electrostatic force and the needs of applications, focusing on optimizing the form of the electrode, its role is similar to the selection of shape representation function in quantitative optimization. This optimization method in the choice of the shape of the drive electrode is often arbitrary, so there is often a lot of room for improvement on the results of qualitative optimization. Quantitative optimization use a certain mathematical model to describe the shape of the drive electrode, and then optimize the driving performance by optimizing the parameters in the mathematical model. Therefore, the quantitative optimization effect is better than qualitative optimization.

However, the quantitative optimization still has some limitations. The existing method is based on the long stroke design conditions, that is, the electrostatic drive stroke is much larger than the driving gap. Under such conditions, the relative driving electrode shape can satisfy the approximate condition of the parallel plate capacitance model. And, the existing methods are limited by the type of analytic functions, only linear equation and rational fractional function can describe the shape of the driving electrodes accurately. However, the driving electrode shape described in these two forms is probably a very small part of all feasible driving shapes, and its optimization effect is limited by the type of the selected function.

The existing quantitative optimization method in the design of the beginning is for long stroke applications, such as MEMS resonator or MEMS displacement platform. In these applications, the limitations described above are resolved by specific application requirements and design conditions. The MEMS switch has a short stroke, which is often equal to or slightly larger than the driving gap. The basis of the existing quantitative optimization solution is unsupported in the MEMS switch. Furthermore, the shape of the electrode of the MEMS switch has a larger curvature, using a linear function or a rational fractional function to describe the MEMS electrode shape has a lot of limitations. Considering the above reasons, the existing driving electrode optimization method cannot be directly applied to the design of the driving electrode of the MEMS switch.

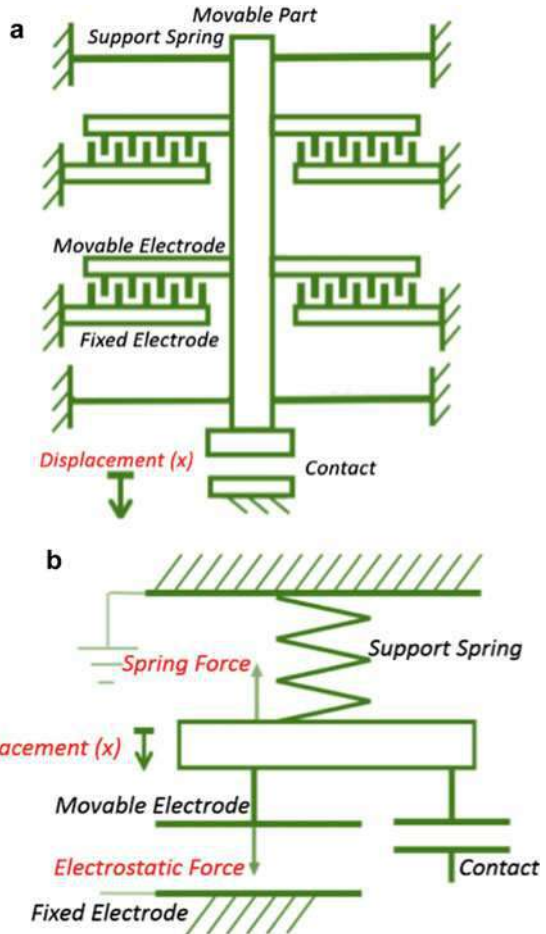
Therefore, the study of electrostatic driving electrode shape for MEMS switching devices has theoretical significance for the design of electrostatic MEMS switches. At the same time, it is important to find the optimal design of electrostatic driving electrodes for MEMS switch under given conditions for improving the device driving performance.

---

## **Theoretical Analysis of Design of In-Plane MEMS Electrostatic Actuator for Switches**

Electrostatic MEMS switches can be simplified as shown in Fig. 1, the mechanical – electrostatic field coupling model. The working principle and process of electrostatic MEMS switch can be described as below. The movable electrode of the switch, which is connected with the movable contact, together constitute the movable

**Fig. 1** Schematic view of electrostatic MEMS switch (a) and mechanics – electrostatic field coupling model (b)



structure. The movable structure is regarded as the source. The switch drive electrode is a spring structure connected to the device substrate, and is seen as a gate. A pair of opposing capacitor electrodes are formed before the movable electrode. When a bias voltage is applied between the movable electrode (source electrode) and the driving electrode (gate electrode), an electrostatic field is generated due to the charging of the capacitor, thereby pulling the entire movable electrode. The structure overcomes the spring return force towards the drive electrode and the stationary contact (which is regarded as the drain) is connected to the moveable contact. At this time, if the bias voltage between the drive electrode and moveable electrode is removed, the original charge of the capacitor discharge occurs, the electrostatic force disappears, and the moveable contact separates from the static contact, thus the switch is disconnected.

In the quasi-static process, if the electrostatic MEMS switch is to be turned on, the electrostatic force  $F_{es}$  must be greater than the recovery force  $F_{re}$  in the entire trip. As

the electrostatic force and the recovery force change with the displacement of the movable part of  $x$ , the conditions for the closure of the switch mathematically are expressed as

$$F_{es}(x) \geq F_{re}(x) (0 \leq x \leq s) \quad (1)$$

Where  $s$  represents the travel of the movable contact to the stationary contact.

The return force  $F_{re}$  can be simplified as a linear spring return force, where, set the spring stiffness  $k$ , the movement direction of the displacement  $x$ , the return force can be simplified as:

$$F_{re}(x) = kx \quad (2)$$

While the electrostatic driving force can be expressed as [58]:

$$F_{es}(x) = \frac{N}{2} \frac{\partial C}{\partial x} V^2 \quad (3)$$

Where  $C$  is the capacitance of the driving electrode unit,  $N$  is the number of driving electrode units, and  $V$  is the driving voltage. Substituting Eqs. 1 and 2 into Eq. 3 gives:

$$\frac{N}{2} \frac{\partial C}{\partial x} V^2 \geq kx, 0 \leq x \leq s \quad (4)$$

Therefore, to reduce the driving voltage  $V$ , one can reduce the spring stiffness  $k$ , increase the number of electrode units  $N$ , or improve the value of  $\partial C/\partial x$ .

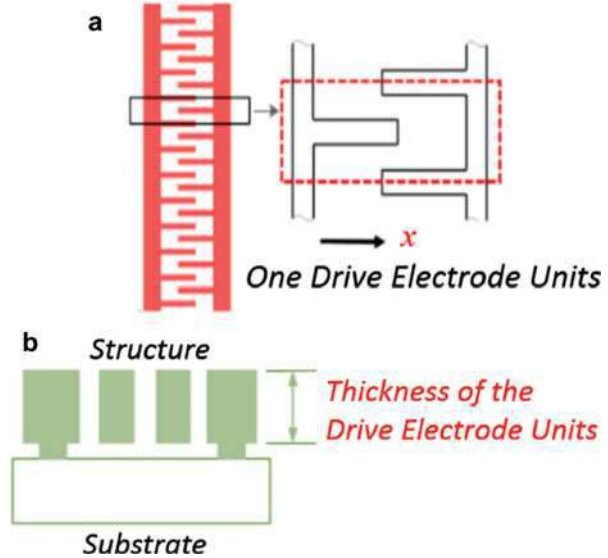
The first two ways have obvious side effects: reducing the spring stiffness  $k$  will lead to low spring return force, will increase the risk of adhesion failure; increasing the number of electrode units will inevitably improve the quality of mobile devices and increase the device area, resulting in deterioration of device stability and cost.

Therefore, the solution with the most potential is to improve the value of  $\partial C/\partial x$ , which is the ratio between capacitance and the displacement of the electrode, and it is controlled by (1) shape of the driving electrode and (2) the thickness of the electrode. In general, the thickness of the electrode structure and the thickness of the spring is the same. The spring stiffness  $k$  and the in-plane switch  $\partial C/\partial x$  are proportional to the thickness of the electrode. Thus varying the thickness of the structure has no effect on reducing the driving voltage. Changing the shape of the driving electrode does not have the above-mentioned side effects. For the in-plane MEMS switch, the driving electrode shape is determined by the photolithographic pattern.

The drive electrode unit is shown in Fig. 2. Let it be assumed that the unit thickness capacitance of the driving electrode unit is  $C'$  and the thickness of the electrode unit is  $h$ .

$$\frac{\partial C}{\partial x} = \frac{\partial C'}{\partial x} h \quad (5)$$

**Fig. 2** Schematic view of the drive electrode units; (a) in-plane; (b) cross-section



Substituting Eq. 5 into Eq. 3 and letting N be 1, one can get the electrostatic driving force of driving electrode unit:

$$F_{es0}(x) = \frac{1}{2} \frac{\partial C'}{\partial x} hV^2 \tag{6}$$

On the other hand, the electrostatic field  $U$  in the driving electrode unit satisfies the Laplace equation, as shown in Fig. 3.

Thus its driving force can be expressed integral form:

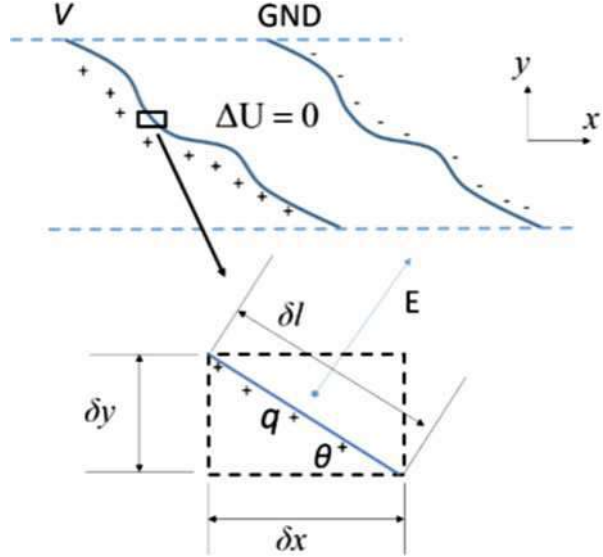
$$F_{es0}(x) = \int \frac{1}{2} \varepsilon E^2 \sin \theta h \delta l = \frac{h}{2} \int \varepsilon E^2 \delta y \tag{7}$$

Where  $E$  is the electric field strength at the driving electrode, and  $\varepsilon$  is the dielectric constant. Since the electrostatic field is often in the air or in the vacuum, the value of the vacuum permittivity can be taken. The  $y$ -axis represents the direction perpendicular to the direction of motion  $x$  direction. Let Eqs. 7 and 6 be equal, one can get:

$$\frac{\partial C'}{\partial x} = \frac{1}{V^2} \int \varepsilon E^2 \delta y = \frac{1}{V^2} \int \varepsilon (\nabla U|_{U=V})^2 \delta y \tag{8}$$

Equation 8 is the expression of  $\partial C'/\partial x$ . It can be seen that there are two principles to increase the driving force by raising the differential term: first, to increase the electric field strength  $E$  at the electrode boundary under the same driving voltage  $V$ ; second, to avoid the projection of electrode boundary on the  $y$ -axis being zero.

**Fig. 3** The electrostatic field and driving force of the drive electrode units (Gao 2015)



The principle of the first point is that, at the same voltage, not only to improve the peak of the electric field strength  $E$ , but also to improve the distribution of field strength; and the principle of the second point is to avoid the shape of the electrode paralleling to the direction of the motion side, which is reason of the tendency to optimize the rectangular electrodes into inclined electrodes [60–64].

For a summary: the optimization of the control variable is the electrode shape, the objective function is the driving force curve, and the optimization object is to make the target function as large as possible. The process of finding the objective function from the control variable is as follows:

1. In the case of a certain displacement, the electrode shape is the boundary of the electric field. Use the displacement to solve the Laplace equation of electrostatic field;
2. at the electrode boundary, use the square of the electric field intensity along the  $y$ -axis integration to obtain the driving force  $F_{es}$  under the current displacement;
3. Change the displacement, repeat the above two steps to get the driving force curve  $F_{es}(x)$ .

This optimization problem is mathematically more complex, and the mapping of the control variable to the objective function is a function-to-function mapping instead of a scalar-to-scalar one. Finding the objective function of the analytical method is difficult. On the other hand, considering the numerical solution of the Laplace equation has good stability and convergence, the numerical solution method is more likely to calculate the objective function. Therefore, the numerical method for the objective function calculation is more feasible.

## Optimization Algorithm and Realization of Segmented Driving Electrode

In the previous section, the control variables, objective functions and algorithms of the shape optimization problem of the electrostatic driving electrode have been clarified. However, there are still some problems in how to choose the optimization method:

Without any restriction, the control variable to be optimized is an abstract function without any specific form. This unrestricted optimization is feasible as a mathematical problem, but not feasible in practical engineering design. Therefore, the first problem is how to constrain the abstract function to a concrete, parameterized form of function.

In previous work on electrode shape quantitative optimization, researchers tend to express electrode shapes in a single elementary functional analytic form [60–63]. However, the shape of the curve is limited by the choice of the form of the function. The qualitative optimization of the shape of the electrode is used in the idea: different positions in the electrode, using different electrode shape, one can take into account the different driving stages of the driving force needs. This idea suggests that the shape of the driving electrode can be expressed in the form of a polyline. This expression has the following advantages:

1. The mathematical expression is simple, each segment is a straight line;
2. Not a single, specific mathematical expression to describe the shape of the entire drive electrode, the expression is not limited by the function selection. Theoretically, as long as the number of segments is enough, this form can be accurate enough for any elementary Function form for approximation;
3. The expression of the form is simple. Thus the non-linear graphics design is converted into multi-segment approximation of the fitting process before the layout. The conversion error is relatively easier to be reduced.

For example, if the drive electrode unit is divided into  $N$  segments, each segment can be represented by a two-dimensional vector  $(x_n, y_n)$ , and  $2N$  variables are needed to be used to describe the shape of the control electrode. As a result, this optimization problem is transformed into a  $2N$ -dimensional function optimization problem. Considering that the optimization problem has multidimensional parameters and the control variables are not mapped to the objective function, a genetic algorithm is used to optimize the solution.

Genetic algorithm is a heuristic search algorithm that simulates the natural selection process, which is usually an effective solution to optimization problems and search problems [65]. Genetic algorithm is a kind of evolutionary algorithm, inspired by the natural evolution of the optimization problem solutions, such as genetic, mutation, selection, cross-over and so on.

Genetic algorithms was begun in the 1950s by biologists in computer simulation of biological genetic processes [66, 67]. In the 1960s, their embryonic form was used by researchers as an algorithm for solving problems [68]. John Hollander of the

University of Michigan and his team put forward the normalized genetic algorithm, the genetic algorithm to promote the credit. By the 1980s, the genetic algorithm has gradually entered the practical and played an important role in a lot of fields such as the engineering optimization and design, business decision-making and production scheduling.

According to the general process of genetic algorithm, when solving the optimization problem, first select a group of candidate solutions; according to adaptability, calculate the fitness of these candidate solutions; according to the fitness, retain the optimal solution. After that, a new candidate solution is generated and then the above process is repeated until the termination condition is satisfied. The process of generating new candidate solutions is to mimic the genetic process of organisms, by coding each solution “genetically,” and to randomize the encoding of the exchange between different solutions (“genetic”) and random encoding changes (“mutations”), to generate new “gene” coding, which is the genetic source of the naming algorithm. And the termination condition may be that the number of iterations reaches the upper limit or the adaptability of the candidate unmarshal meets the requirement. The following describes the use of genetic algorithms to describe the multi-segment line to optimize the shape of the driving electrode flow.

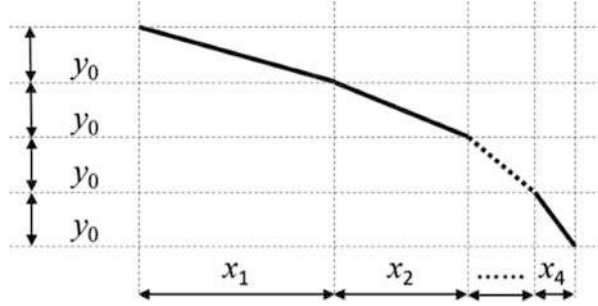
### (1) Coding

In order to be able to compare the objective function of different designs, the width of the drive unit should be consistent. So the design of the domain should be selected the same drive unit width  $W$ , the size of the drive unit in the  $y$ -axis direction. Taking into account the second principle of the optimization method in the previous section, the  $x$ -axis driving force is not generated for sub-line segments whose  $y$ -axis projection is zero. Let  $y_0$  be the  $y$  component of all sub-segment vectors, the line segment can be expressed only with the  $x$  component, the control variable value can be reduced to the number of segments  $N$ , and has the following relationship:

$$W = y_0 N \quad (9)$$

The next step is to determine  $y_0$ . When  $W$  is a constant,  $y_0$  will be smaller when more segments are adapted. The electrode shape will be expressed more accurately and detailed. But on the other hand, more segments will lead to increment of the amount of control variables, consumption of optimization process computing resource, and computing time. Therefore, it is necessary to determine  $y_0$  in a balanced manner. In this paper, the width of the driving unit is about several microns. The result of pre-calculation shows that  $y_0$  takes  $1 \mu\text{m}$  to meet the requirements. Then for a drive unit of width  $N \text{ m}$ , the shape can be represented by  $N$ -dimensional vectors  $(x_1, x_2, x_3 \dots x_n)$ , the encoding of the driving electrode unit is completed. The coding is shown in Fig. 4. As the process of MEMS processing, lithography, etching and other steps for the graphic's too small corners of the arc shape will produce deformation. Consider the needs of experimental verification, to minimize the design results and the difference between the results of the processing. In this paper, the

**Fig. 4** Schematic view of the polyline shaped driving electrode (Gao 2015)



radius of curvature is about 1–1.5  $\mu\text{m}$ . Therefore, after finishing the coding, we need to check the shape and avoid the design of large deviation.

(2) Selecting the Initial Candidate Unmarshal

In general, there are no clear criteria for the selection of the initial candidate unmarshal. The experimental results show that, in the problem of shape optimization of the driving electrode, the better initial candidate de-grouping should include various “gene” codes, whose inclination angles are small to large. In other words, the initial candidate de-grouping should have good enough “gene” diversity. If the initial candidate unmarshaled “genes” are monotone, the candidate solutions are easy to fall into the local optimal solution.

(3) Solving the Adaptability of the Unmarshals

In the driving electrode shape optimization problem, the adaptability is the driving force or driving force curve at a certain displacement of the electrode. As explained in the previous section, there are three steps from driving the electrode shape to solving the driving force:

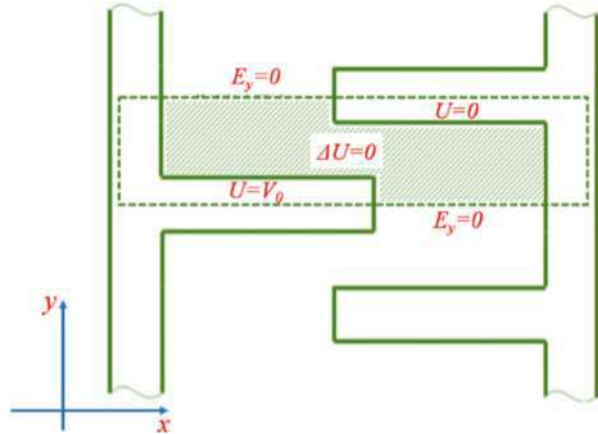
The first step is to determine the boundary condition of the electrostatic field according to the shape of the driving force and to solve the Laplace equation of the electrostatic field. The boundary of the electrostatic field is made up of four sides, two opposing electrode shape sides and two opposing symmetry axes. One side of the electrode condition is the potential of 0, one side of the electrode condition is the potential for the driving voltage, symmetrical axes have symmetry conditions, the boundary condition of the electric field component is zero. The boundary conditions are shown in Fig. 5.

To determine the boundary of the electrostatic field, one is to determine the shape of the opposing electrode, according to the principle of symmetry, relative driving electrodes should have the same shape. In addition, the electrostatic field boundary condition also determines the initial distance of the opposing electrode.

The first condition is that the initial minimum distance between the driving electrodes should not be less than the minimum line width to be machined. The line width generally refers to the width of the thinnest line pattern in the graph. The driving electrode optimization is related to the gap between the electrodes that are relative to each other.



**Fig. 5** Schematic view of the boundary conditions



However, for irregularly shaped electrodes whose shape is to be designed, the “line width” can be redefined as the minimum distance between any two points on the counter electrode, since the shape is more regular and value. The second condition is that the short-circuit distance between the opposing electrodes should not be less than the design requirement. According to these two conditions, the minimum value of the initial distance of the driving electrode is determined, whichever is the larger. Electrode short-circuit distance as a parameter involved in a wide range of its determination often needs to be combined with other conditions given, such as the working environment of the device shock and vibration conditions, restoring force requirements, and in the device design process, it may be adjusted according to the design results and design conditions. The focus of this paper is not on the determination of design conditions, so the conditions for the minimum short-circuit distance are considered here.

The next work is to solve the electrostatic field Laplace equation. The stability and convergence of the Laplace equation are good. Commercial computing softwares, such as COMSOL Multiphysics, etc., can be a good solution to complete the work. On the other hand, one can also develop custom solver, which is based on the finite difference method.

The second step is to compute the driving force based on the solution of the electric field. If one is using commercial software for electric field solution, he can use commercial software coming with the post-processing function for driving force calculation. If one is writing a custom solver, he can compute the driving force according to Eq. 7.

The third step is to change the displacement. Repeat the first step and the second step, get the corresponding displacement of the driving force, one can calculate the driving force curve.

#### (4) Generating New Candidate Unmarshales

After calculating the driving force curve of this generation, M drive electrode units with the lowest critical voltage drop can be selected as seed for the next generation

candidate de-grouping. In this paper, two strategies of cloning and genetic variation are adopted to generate the next generation candidate set.

“Clone” strategy is the seed of the previous generation, directly becoming a candidate for the next generation of solutions. “Genetic variation” refers to: the selection of the previous generation of two seeds, each control variable randomly selected two seeds corresponding to the location of the code as their own code and within a certain range of random variation to generate the next generation of candidate solutions. “Clone” strategy can guarantee the stability of optimal solution. The adaptability (driving force or driving voltage) of the entire candidate de-grouping (drive electrode cell group) is monotonically optimized between generational generations without the occurrence of adaptive degradation. “Genetic variation” can be based on the characteristics of the seed within a certain range of random search. Suppose that the number of the next generation candidate solutions generated by the above method is  $K$ , then (Eq. 10) holds. Where  $M$  is the number of seed solutions selected by the previous generation, the first term on the right of the equation is the number of the next generation candidate generated by “cloning,” and the second term is the number of the next generation candidate generated by “genetic variation.”

$$K = M + M(M - 1)/2 \quad (10)$$

At this point, a new generation of candidate unmarshal is generated. However, the emphasis of this paper is not on genetic algorithm, so here we only choose one of the methods to solve the problem of driving electrode optimization.

#### (5) Cycle Solution

The above steps (1) to (4) are repeated until the termination condition is satisfied. At this time, the final generation of seeds corresponding to the shape of the drive electrode is the result of optimal design. In the actual calculation, the solutions in the candidate solutions are gradually convergent as the number of cycles increases, and the driving force curve will increase more and more slowly. Therefore, these two factors and the number of cycles can be considered as cyclic termination conditions.

#### (6) Algorithm to Achieve

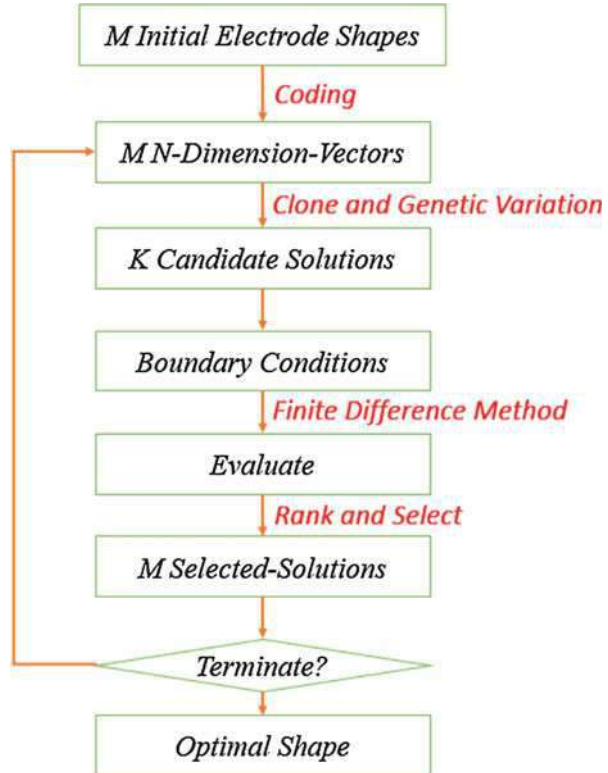
In this work, custom procedures are used to complete the drive electrode optimization process, the program flow chart is shown in Fig. 6.

## Optimization Results Analysis

### The Constraint Value

Constraints used in the optimization process are listed in Table 2. Processing line width of  $6 \mu\text{m}$  is because for the current MEMS lithography and etching technology, this is a more relaxed processing requirement. The current processing level in this line width conditions can achieve better yield. The drive short stroke is the empirical

**Fig. 6** Flow chart of the custom electrode shape optimization process



value taken under the conditions of mechanical stability and restoring force. The width of the drive electrode is based on the pre-calculated results, balance optimization of computing resources and shape optimization results after the value.

According to the optimization method described in the previous section, we obtain the optimization results of the driving electrode unit under the current constraint conditions, as shown in Fig. 7.

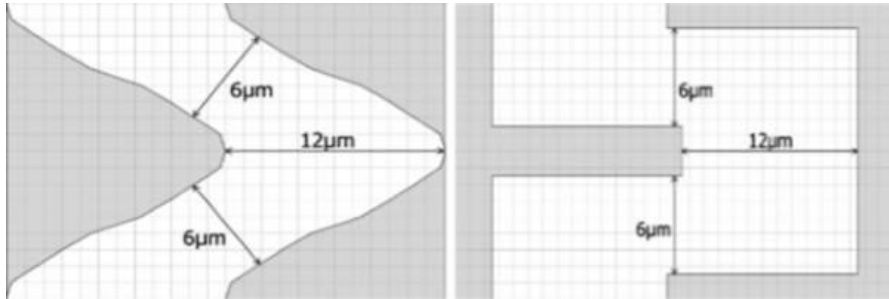
To compare the effect of driving electrode shape optimization, the optimized design is compared with the driving force curve of a single drive unit of the conventional design, as shown in Fig. 8.

It can be seen that the optimum design driving force is 17% higher than the conventional comb tooth when the displacement is 0  $\mu\text{m}$ . As the displacement increases, the driving force of the traditional comb tooth rises very slowly, and the driving force of the optimization design rises at a very fast speed. When the displacement is 6  $\mu\text{m}$ , the driving force of optimized design is 200% higher than that of traditional comb.

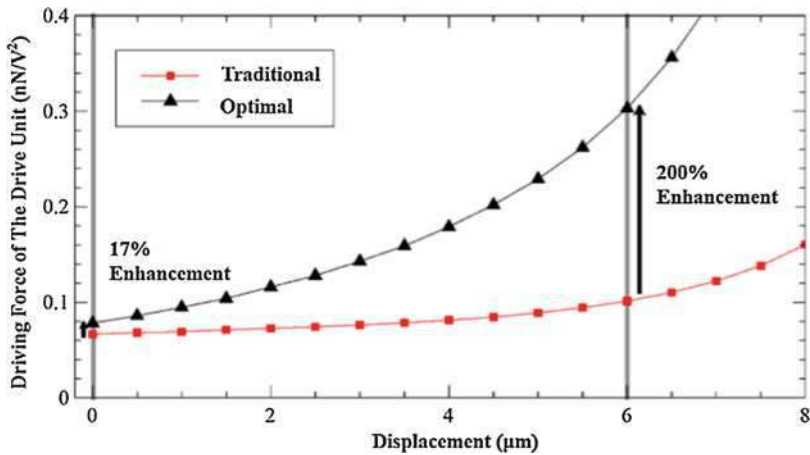
In order to compare the optimal design and the driving performance of the traditional comb more clearly, take the drive voltage of 1 V, the two field strength on the electrode boundary y-dimensional one-dimensional surface and the plane of the two-dimensional surface are plotted. The electric field when the movable

**Table 2** Constraint value table of electrode shape optimization solving

Constraint	Value
Line width	6 $\mu\text{m}$
Travel	12 $\mu\text{m}$
Drive electrode unit width	9 $\mu\text{m}$

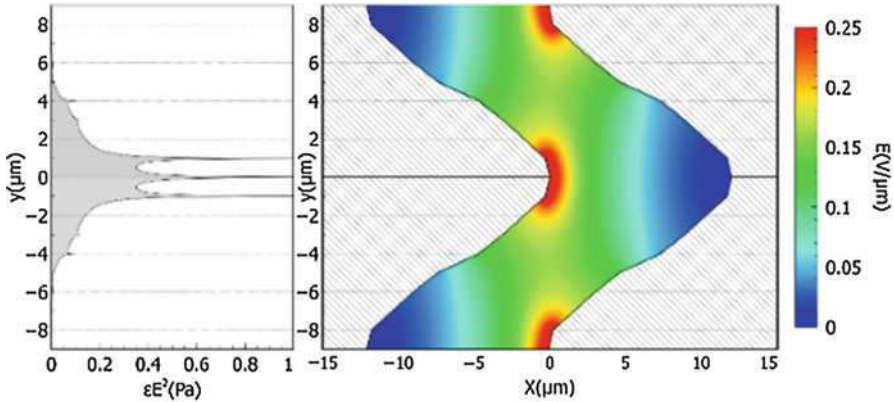


**Fig. 7** Comparison of the shapes of the driving electrodes before and after optimization (Gao 2015)

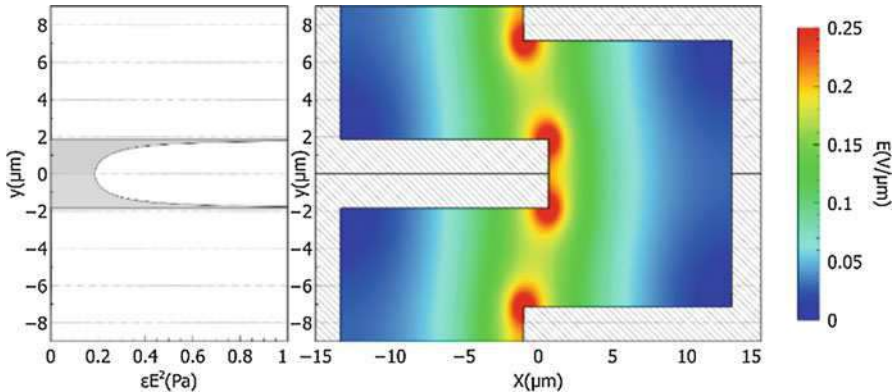


**Fig. 8** Comparison between driving force of the traditional and optimal electrode design

electrode displacement is 0  $\mu\text{m}$  is shown in Figs. 9 and 10. Although the maximum and the distribution of the two-dimensional electric field strength are similar, the square of the field strength of the electrode boundary has a larger distribution in the y direction. When the displacement of the movable electrode is 6  $\mu\text{m}$ , as shown in Figs. 11 and 12, in the optimal design, almost all the field width distribution, for the traditional comb teeth, high field strength part is still concentrated in the tooth tip. This shows that the optimization design follows the optimization principle of



**Fig. 9** Simulation of the two-dimensional distributions of y-axis electrostatic field intensity at displacement = 0 μm of the OPTIMAL design (Gao 2015)



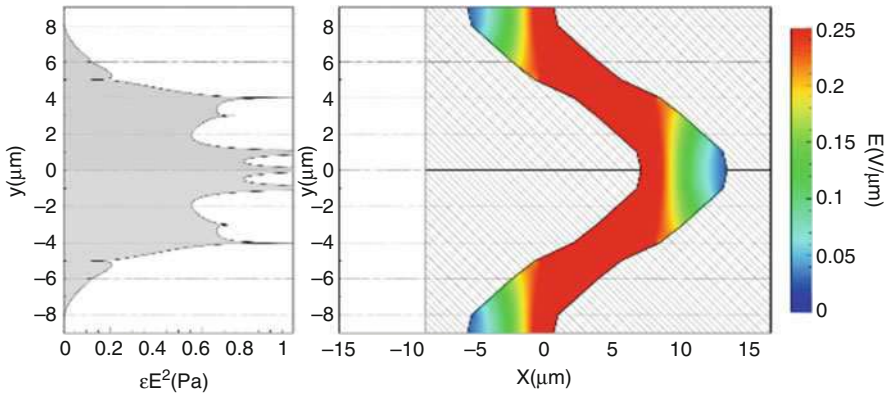
**Fig. 10** Simulation of the two-dimensional distributions of y-axis electrostatic field intensity at displacement = 0 μm of the TRODITIONAL design (Gao 2015)

increasing the distribution of the electric field on the electrode boundary to the y-axis.

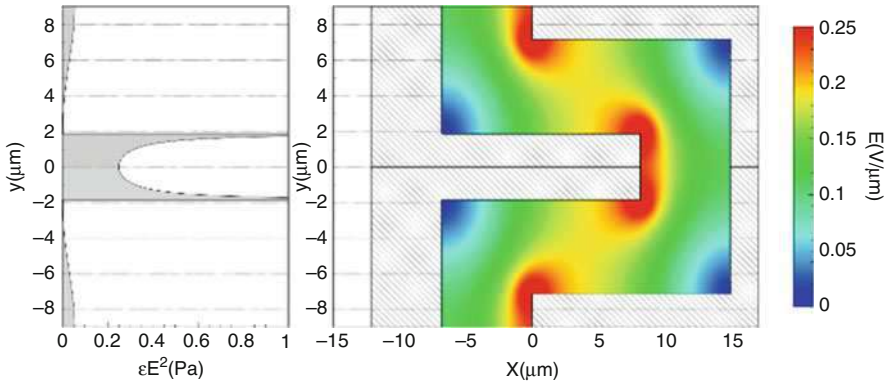
## Experimental Verification of Driving Electrode Optimization

### Structure and Machining of Electrostatic Actuator for MEMS Switch

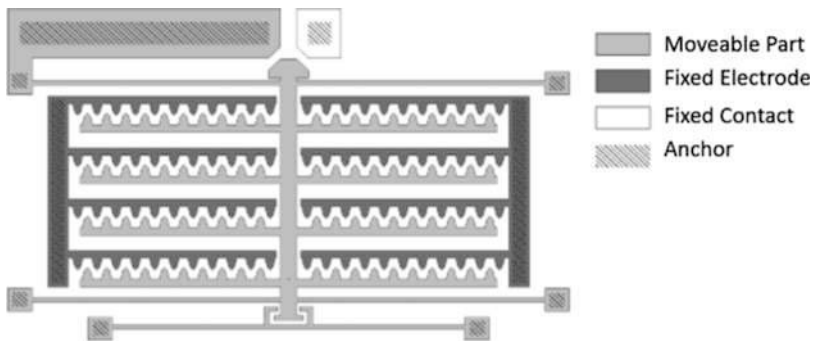
The structure of the MEMS driver is shown in Fig. 13. In the design, four slender fixed support beams are used as the spring to provide restoring force, the springs' roots are fixed on the most peripheral area of the device through the anchor. In the middle of the mobile structure, array drive electrode units are arranged, the upper



**Fig. 11** Simulation of the two-dimensional distributions of y-axis electrostatic field intensity at displacement = 6 μm of the OPTIMAL design (Gao 2015)

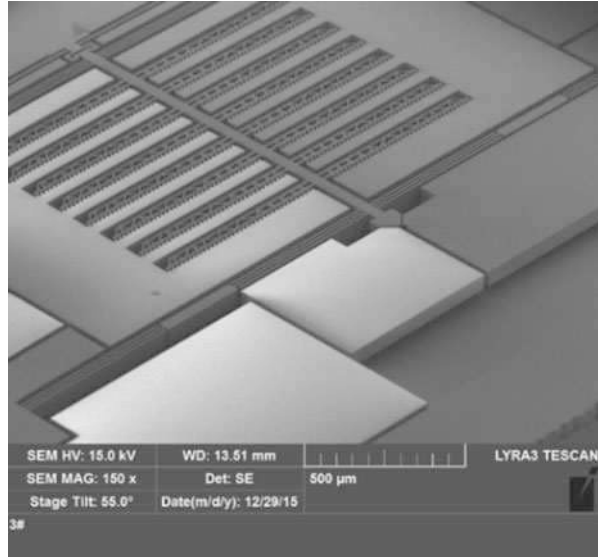


**Fig. 12** Simulation of the two-dimensional distributions of y-axis electrostatic field intensity at displacement = 6 μm of the TRODITIONAL design (Gao 2015)

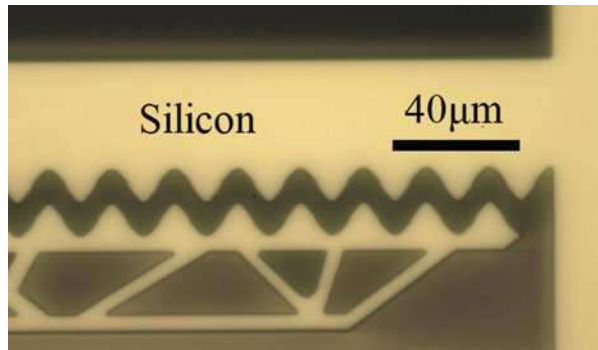


**Fig. 13** MEMS electrostatic actuator structure diagram (Gao 2015)

**Fig.14** A SEM photograph of the resulting device (Gao 2015)



**Fig. 15** A microscope photograph of the drive electrode unit (Gao 2015)



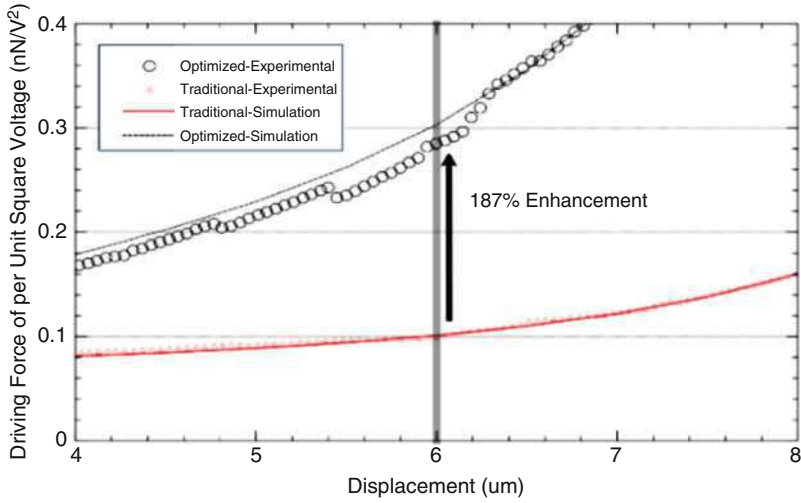
part is dynamic contact and static contact. When the device is in operation, the movable structure moves under the action of the electrostatic force, and the middle part moves upwards, so that the dynamic contact comes into contact with the static contact.

A SEM photograph of the resulting device is shown in Fig. 14, and a microscope photograph of the drive electrode unit is shown in Fig. 15.

### Test Results of Driving Electrode Driving Force

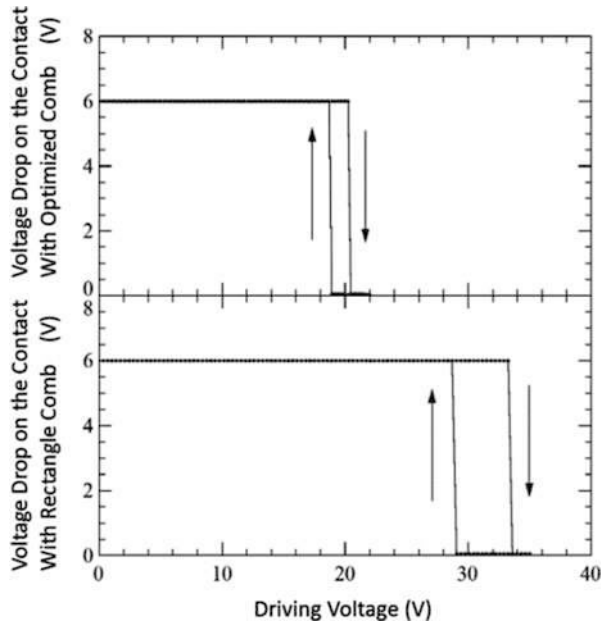
In this paper, the driving force curve is measured for the optimized MEMS static driver and the MEMS driver with traditional comb finger. For comparison purposes, the measurement results are converted into the driving force of a single drive unit at square





**Fig. 16** Experimental comparison between optimal design and traditional comb-drive force curve (Gao 2015)

**Fig. 17** Pull-down driving force experimental comparison between optimal design and traditional comb-drive (Gao 2015)



voltage, as shown in Fig. 16. Due to the limitation of the measurement method, the driving force data of which the driving displacement is too small to be measured. Compared with the simulation results, it can be seen that the measured driving force is increased by 187% at the displacement of 6  $\mu\text{m}$  in the effective displacement section, which is basically the same as 200% increase of the simulation.



To further compare the pull-down effects of different drive electrodes, the voltage sweep method was used to measure the pull-down voltage. The test results are shown in Fig. 17.

The comparison shows that the comb finger shape optimization method proposed in this paper can improve the driving force and reduce the driving voltage in the electrostatic MEMS switch.

---

## Conclusion

This paper summarizes and compares the in-plane MEMS electrostatic driving technology. Aiming at the characteristics of MEMS switches, an optimization method of in-plane MEMS electrostatic comb-drive finger shape is proposed. Taking the process line width and stroke as the boundary conditions, numerical solution of the electrostatic field equation was used to solve the optimization problem. The genetic algorithm was used to optimize the finger shape. Under the conditions of line width of 6  $\mu\text{m}$  and stroke of 12  $\mu\text{m}$ , the above optimization problems are solved and the results are compared with the traditional rectangular electrodes in simulation and experiment. The experimental results show that the driving force is increased by 187% at the displacement of 6  $\mu\text{m}$ , which verifies the validity and feasibility of the proposed optimization method.

---

## References

- Engelen JBC, Lantz MA, Rothuizen HE et al (2009) Improved performance of large stroke comb-drive actuators by using a stepped finger shape: solid-state sensors, actuators and microsystems conference, Transducers 2009. International, 2009[C]. IEEE, 2009
- Gao Y (2015) Research on the in-plane electrostatic MEMS DC Switch [D]. Tsinghua University
- Hirano T, Furuhashi T, Gabriel KJ et al (1992) Design, fabrication, and operation of submicron gap comb-drive microactuators. *J Microelectromech Syst* 1(1):52–59
- Jensen BD, Mutlu S, Miller S et al (2003) Shaped comb fingers for tailored electromechanical restoring force. *J Microelectromech Syst* 12(3):373–383
- Kang S, Kim HC, Chun K (2009) A low-loss, single-pole, four-throw RF MEMS switch driven by a double stop comb drive. *J Micromech Microeng* 19:35011
- Legtenberg R, Groeneveld AW, Elwenspoek M (1996) Comb-drive actuators for large displacements. *J Micromech Microeng* 6(3):320–329
- Park KY, Oh YS (1998) Laterally oscillated and force-balanced micro vibratory rate gyroscope supported by fish-hook-shaped springs. *Sensors Actuators A-Physical* 64(1):69–76
- Rosa MA, Harrison HB (1998a) Enhanced electrostatic force generation capability of angled comb finger design used in electrostatic comb-drive actuators. *Electron Lett* 34(18):1787–1788
- Rosa MA, Harrison HB (1998b) Improved operation of micromechanical comb-drive actuators through the use of a new angled comb finger design. *J Intell Mater Syst Struct* 9(4)
- Ye W, Mukherjee S, Macdonald NC (1998) Optimal shape design of an electrostatic comb drive in microelectromechanical systems. *J Microelectromech Syst* 7(1):16–26



# MEMS Actuators Driven by Lorentz Force

Jinling Yang and Jinying Zhang

## Contents

Overview .....	930
Microactuators Driven by Lorentz Force .....	931
Key Principle .....	931
Key Research Findings .....	934
Potential Applications .....	943
References .....	944

## Abstract

This chapter presents MEMS actuators driven by Lorentz force. This kind of actuators linearly depends on the current perpendicular to the magnetic field, requires no magnetic materials, and has no magnetic hysteresis effect. These advantages result in the actuators of simple structure, linear motion, fast response, reasonable power consumption, and ideal for large stroke application. A novel actuator with folded beams structure is developed for a large lateral stroke. A displacement of more than 55  $\mu\text{m}$  was achieved with a magnetic field of 0.14 T and the driving current of 8 mA. The actuator can generate a large displacement by a low driving voltage and can be easily integrated with CMOS circuits. Lorentz force is proportional to the magnetic field and the driving current, which results in an easy control of the lateral displacement by the driving current. The simple structure and fabrication process of the actuator lead to a high fabrication yield and good stability.

J. Yang (✉) · J. Zhang

Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

State Key Laboratory of Transducer Technology, Shanghai, China

School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing, China

e-mail: [jlyang@semi.ac.cn](mailto:jlyang@semi.ac.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,

[https://doi.org/10.1007/978-981-10-5945-2\\_28](https://doi.org/10.1007/978-981-10-5945-2_28)

929

---

**Keywords**MEMS actuator · Lorentz force · Large stroke · Large displacement · High yield

---

**Overview**

Microactuators are important building blocks for many MEMS devices, which generate forces or displacements to realize scanning, tuning, manipulating, or delivering function. In the last decades, a large number of actuators have been developed for various applications (Judy et al. 1995; Jeong et al. 2004; Grade et al. 2003; Ko et al. 2002; Qui et al. 2003; Hwang et al. 2003) and they are usually driven by electrostatic, thermal, piezoelectric, and electromagnetic methods, each of which has its advantage and restriction. For example, the thermal actuators can provide a large force and stroke but suffer from high power consumption and long response time. The material used in the piezoelectric actuators is usually not compatible with IC technology. A magnetic field and magnetic material are needed for the electromagnetic actuators. Microelectrostatic actuators are more popular than others due to their high compatibility with mature microfabrication processes, low power consumption, simple structure, and quick response, but require high operating voltage and show nonlinear behavior.

In many applications, such as optical switches, variable optical attenuators (VOA), and microscanners, the MEMS actuators are highly desired to provide large displacements of several tens of micrometers with a small voltage below 3.3 or 5 V supplied by CMOS circuits (Jeong et al. 2004). A variety of MEMS actuators have been proposed to achieve a large stroke. Electrostatic (Grade et al. 2003), electromagnetic (Ko et al. 2002), thermal actuation methods (Qui et al. 2003), and a chevron-type compliant structure (Hwang et al. 2003) have provided large displacements. However, a large driving voltage beyond 5 V is needed to achieve a large stroke, and thus the practical application of the actuators is limited. Lorentz force resulting from the interaction of an electric field and a magnetic field has been widely employed to excite MEMS devices (Sun et al. 2012; Lee et al. 2011; Lin et al. 2013). It linearly depends on the current perpendicular to the magnetic field, requires no magnetic materials, and has no magnetic hysteresis effect. These advantages result in the actuators of simple structure, linear motion, fast response, reasonable power consumption, and ideal for large stroke application.

In the following sections, MEMS actuators driven by Lorentz force are introduced. Section “[Microactuators Driven by Lorentz Force](#)” introduces typical MEMS actuators driven by Lorentz force, such as the microactuators with vertical architecture or curved beam springs. Several crucial parameters for achieving large displacement under a low driving voltage are discussed. In section “[Key Principle](#),” the designing principle for MEMS actuators driven by Lorentz force with large lateral stroke is presented. The key research findings in recent studies are demonstrated in section “[Key Research Findings](#).” Finally, the potential application is discussed for the MEMS actuators driven by Lorentz force.

## Microactuators Driven by Lorentz Force

Microactuators with large displacement under low driving voltage are highly desired in optical communication, imaging, switches, and variable optical attenuators (VOA), etc. Compared to electrostatic or piezoelectric forces, electromagnetic driving based on Lorentz force has clear advantages.

Ko et al. developed a vertical leaf spring (Ko et al. 2002), which has a high aspect ratio with a width of 1.2  $\mu\text{m}$ , thickness of 16  $\mu\text{m}$ , and length of 920  $\mu\text{m}$ . The leaf spring consisting of polysilicon and aluminum layers has resistance of 5  $\Omega$  and can achieve a large deflection of 60  $\mu\text{m}$  with a driving current of 566 mA; however, a high driven voltage of 28 V is required.

A dual-axis actuator with curved-beam springs consisting of silicon and aluminum was reported by Fang et al. (Lee et al. 2011). This structure could be used as a micropositioning stage and can perform two-axis in-plane motions. The output displacement can reach  $\pm 13 \mu\text{m}$  with input currents of  $\pm 150 \text{ mA}$ . Later on, a one-axis Lorentz force actuator was developed by Fang et al. (Lin et al. 2013). The motion-decoupled spring connected by two serpentine beams on the same side performed a large displacement in the desired direction; the deviation of the displacement in the unwanted direction was restricted. As a consequence, an input current of 50 mA leads to  $x$ -axis and  $y$ -axis displacement of 40  $\mu\text{m}$  and 0.9  $\mu\text{m}$ , respectively.

These results indicate that tens and even hundreds of milliamp current is required to achieve a displacement of several tens of micrometers. In practical applications, microactuator array is usually employed, thus the required driving current could be quite large and would lead to high power consumption and severe thermal dissipation problems, etc. For realizing large displacement under a low driving voltage, more elaborating designs are needed, for example, material properties and structural parameters should be refined for a moderate spring constant and good stress matching between the structural layers. The following section focuses on the designing principle for a MEMS actuator with large lateral stroke driven by Lorentz force.

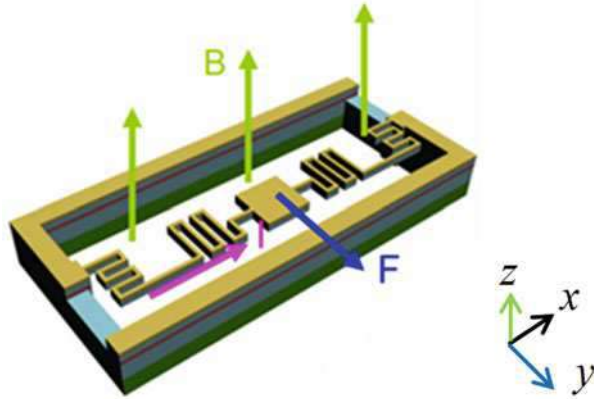
---

### Key Principle

The working principle of the presented MEMS actuator is schematically shown in Fig. 1. Various reported configurations of beams are based on the same principle (Ko et al. 2002; Lee et al. 2011; Lin et al. 2013; Lv et al. 2015a, b). The actuator consists of folded beams and a moving plate. When applying a current  $I$  through the beams and a static magnetic field  $B$  perpendicular to the beam plane, Lorentz force  $F$ , perpendicular to the direction of the current and the magnetic field, will drive the actuator into displacement in the  $xy$  plane.

The lateral stroke of the actuator linearly depends on the current in terms of Eq. 1, Lorentz force can be expressed as

**Fig. 1** Schematic drawing of the MEMS actuator



$$|\vec{F}| = |I \vec{L} \times \vec{B}| = BIL \cdot \sin \theta \tag{1}$$

Where  $B$  is the magnetic flux density of the magnetic field,  $I$  is the current flowing through the beam,  $L$  is the beam length, and  $\theta$  is the angle between the current and the magnetic field.

As shown in Fig. 1, the actuator beam is located in  $xy$  plane and the magnetic field is perpendicular to the folded beam. When a current flows through the folded beam, a driving force  $F$  along the  $y$  direction is generated, as expressed by

$$F = BIL_e \tag{2}$$

where the effective length  $L_e$  is the length of the beam contributing to the lateral motion.

The actuator can be considered as a spring with a spring constant  $k$  (Senturia 2001), it deforms under an external force  $F$ , meanwhile, an elastic restoring force  $F_e$  can be generated

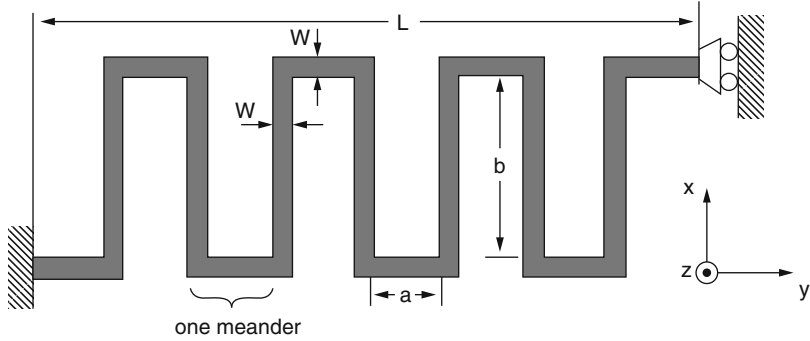
$$F_e = ky \tag{3}$$

where  $y$  is the displacement of the beam. At an equilibrium state, the driving force  $F$  is equal to the elastic restoring force  $F_e$ . The displacement can be calculated by

$$y = \frac{F_d}{k} = \frac{BIL_e}{k} \tag{4}$$

To achieve a large displacement,  $B$ ,  $I$ ,  $L_e$ , and  $k$  should be optimized.  $k$  is determined by the mechanical properties of the material and the structure of the actuator (Fedder 1994).

Figure 2 is the schematic drawing of the folded beam, its spring constant can be calculated (Fedder 1994)



**Fig. 2** Schematic drawing of folded beam

$$k_x \approx \frac{48EI_z}{(a+w)^2(a+w+b+w)n^3} \quad (5)$$

$$k_y \approx \frac{48EI_z}{(b+w)^2[3(a+w)+b+w]n^3} \quad (6)$$

where  $E$  is the Young's modulus,  $I_z$  is the bending moment of inertia about the  $z$ -axis,  $n$  is the folded times,  $w$  is the width of the beam,  $a$  and  $b$  are the row and line pitch, respectively.

For beams connected in series, the spring constant can be calculated as (Fedder 1994)

$$\frac{1}{k} = \frac{1}{k_1} + \frac{1}{k_2} \quad (7)$$

For parallel connection, the spring constant can be calculated as

$$k = k_1 + k_2 \quad (8)$$

For the folded beam as shown in Fig. 3, its spring constant can be calculated by

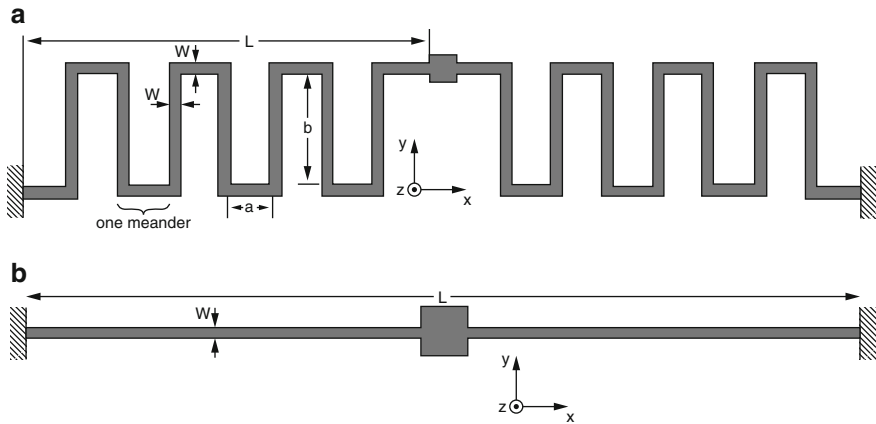
$$k_x \approx \frac{96EI_{z,b}}{b^2(3a+b)n^3} \quad (9)$$

where  $E$  is the Young's modulus,  $I_z$  is the bending moment of inertia about the  $z$ -axis,  $n$  is the folded times,  $a$  and  $b$  are the row and line pitch, respectively (Fedder 1994).

For the double-ended beam with  $b = 0$ , the spring constant can be computed

$$k = \frac{32EWH^3}{L^3} \quad (10)$$

where  $W$  and  $H$  are the width and thickness of the actuator, respectively.



**Fig. 3** Schematic drawing of the actuator, (a) folded beam and (b) double-ended beam

Substituting Eq. 9 into Eq. 3, the displacement can be expressed as

$$y = \frac{BILa^2n^3}{96EI_{z,b}}(a + b) \quad (11)$$

The displacement linearly increases with  $B$ ,  $I$ , or  $b$ .

## Key Research Findings

Single crystal silicon is often used as the beam material due to its excellent mechanical properties and robust lifetime, metal materials are required to transmit the current. Thus, composite films consisting of silicon and metal are often adopted in the actuator structure to generate the Lorentz force.

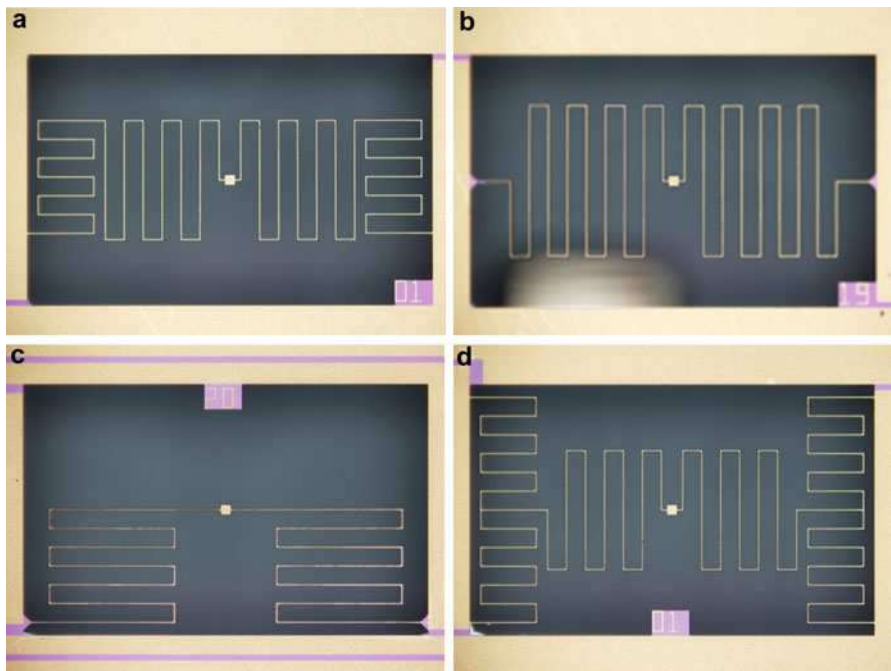
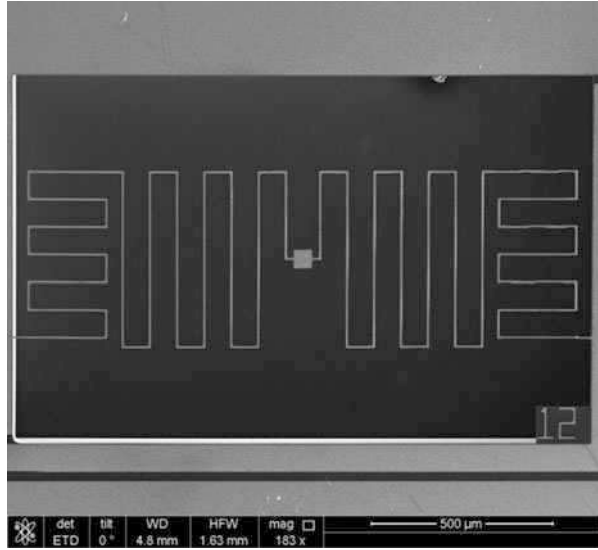
A novel MEMS actuator driven by Lorentz force for a large lateral stroke was recently developed under a low driving voltage (Lv et al. 2015a, b). The actuator showed excellent linear characterization and can be integrated with CMOS circuits.

The fabrication process of the actuator started with SOI wafer with 5  $\mu\text{m}$  thick (100)-oriented top Si layer. A metal layer Cr and Au with thicknesses of 30 and 300 nm was deposited onto the top Si layer by electron beam evaporation and was patterned. The top Si layer was etched by RIE to produce the folded beam. The Si substrate was etched from the backside by DRIE, and the beam was released in hydrofluoric acid (Lv et al. 2015a).

Figure 4 shows the scanning electron microscope (SEM) photograph of a typical actuator. Figure 5 shows the optical photographs of different actuators.

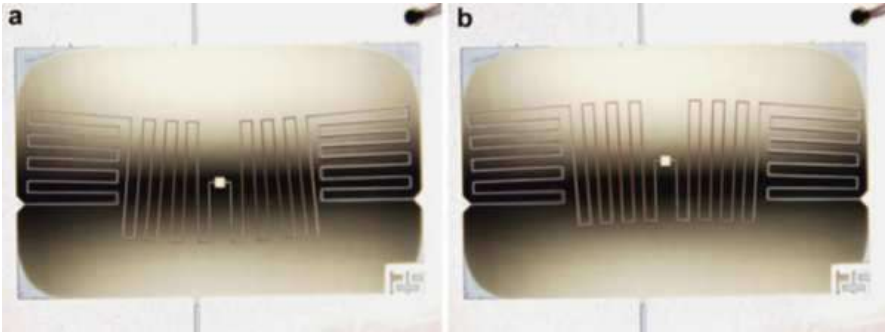
The MEMS actuator was integrated with PCB circuit via wire bonding and was placed on a permanent magnet, which can provide a static magnetic field for the MEMS actuator. The magnetic flux density imposed on the MEMS actuator is about

**Fig. 4** SEM photographs of the fabricated actuator



**Fig. 5** Optical photographs of actuators, (a)–(d) for type 1–4, respectively





**Fig. 6** The optical photographs of the deformed actuators

0.3 T. By applying a driving current, the displacement magnitudes of the MEMS actuator were obtained. The displacement is measured by an optical microscope and the resistance of the MEMS actuator was monitored by a multimeter. Figure 6 shows the optical photographs of the deformed MEMS actuator in the opposite direction.

The relationship between the displacement of the actuator and the flowing current of 0–8 mA in a magnetic flux density of 0.3 T are depicted in Fig. 7. Very good linear dependence was observed for different structures in Fig. 5. Type 1 actuator has the folded beam in both  $x$  and  $y$  direction and thus has the smallest stiffness among the four kinds of structures, which results in the largest displacement. Types 2 and 3 have the folded beams in one direction, possess a relatively large stiffness. Type 4 actuator has a symmetrical structure, more anchors and constraints lead to the smallest displacement among the four actuators. Type 1 actuator will be used as example in the following discussion.

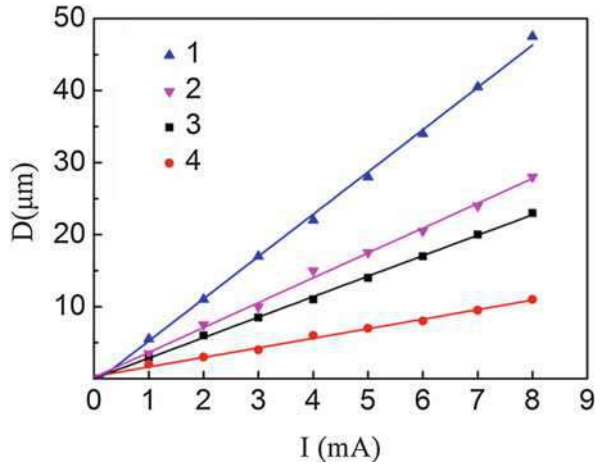
A lateral stroke of more than 47  $\mu\text{m}$  was obtained for type 1 actuator driven by a current of 8 mA, which coincided with the calculation and simulation results. The displacement is proportional to the applied current. The displacements switch to the opposite direction under a reversed current and linearly increase with the current, as shown in Fig. 8.

Figure 9 gives the dependence of the displacement on the magnetic flux density. It can be seen that the displacement is proportional to the magnetic flux density, which agrees with the theoretical analysis. By increasing the magnetic flux density, a much larger lateral stroke can be achieved.

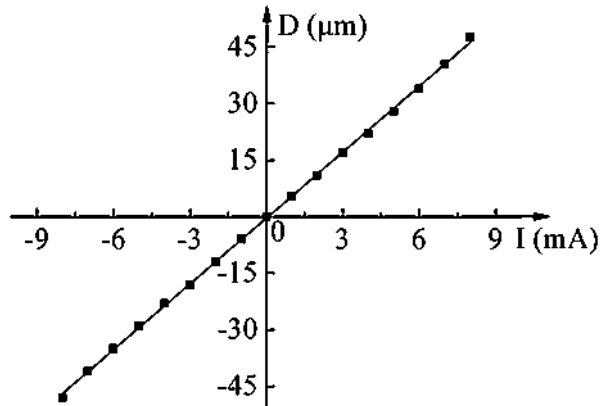
To have a deep understanding on the stability of the actuator, cycling test was carried out. The actuator was tested by monitoring the displacements of the actuator while the driving current goes from 0 up to 8 mA and down from 8 to 0 mA, repeatedly. The cycled displacements versus the driving current are shown in Fig. 10. The actuator has excellent stability and could precisely be controlled. The slight scattering of the displacements at one current were caused by the displacement measurement error with the microscope.

Although the folded beam is very flexible, it is strong and stable enough to be immune from environmental noise. And the actuator has small mass (about 0.6 mg).

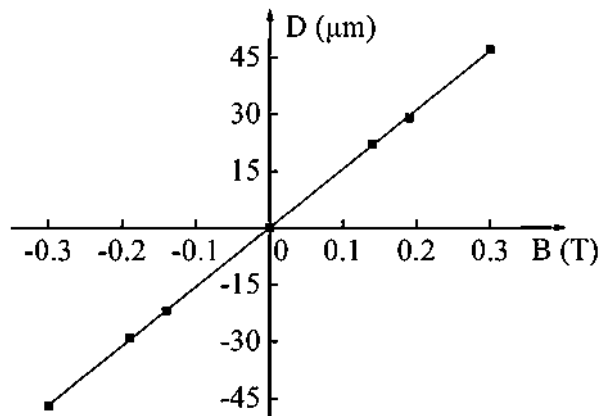
**Fig. 7** Linear dependence of displacement on applied current from 0 to 8 mA with a magnetic flux density of 0.3 T



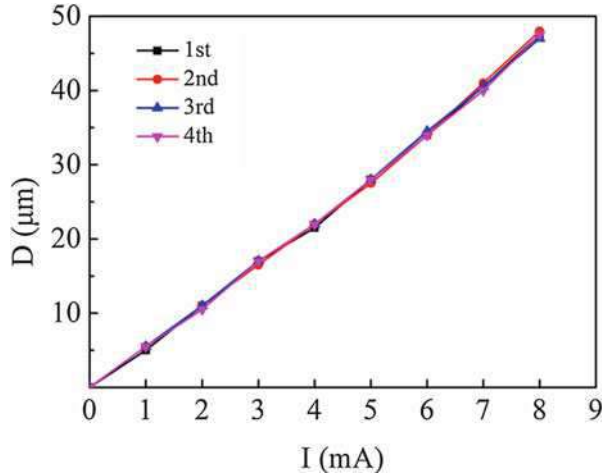
**Fig. 8** Dependence of displacement on applied current from  $-8$  to 8 mA with a magnetic flux density of 0.3 T



**Fig. 9** Dependence of displacement on magnetic flux density from 0 to 0.3 T with a current of 8 mA



**Fig. 10** Cycling test for the actuator driven by a current from 0 to 8 mA with a magnetic flux density of 0.3 T



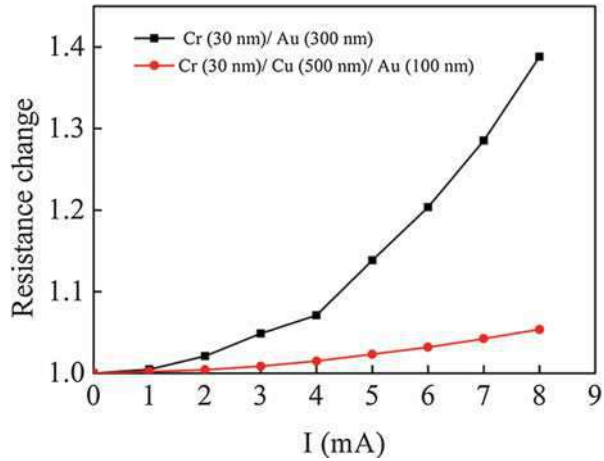
Its resonance frequency is much higher than that of the environment and no mechanical coupling to the environment occurs (Lv et al. 2015b). Therefore, the environment noise has no observable effect on the actuator. During testing, the actuator was placed on a table without any mechanical isolation and steadily displaced with the driving Lorentz force. As shown in Fig. 10, the displacements remained unchanged with the same driving force during the cycling test.

The metal layers used in the MEMS actuator were Cr and Au with thicknesses of 30 nm and 300 nm, respectively. The resistance of the MEMS actuator measured by a multimeter is 340  $\Omega$ , and significantly increased with the current due to Joule heat generated by the current passing through the folded beam, for example, the resistance of the actuator increases to 465  $\Omega$ , about 1.4 times of the original value when the current goes up to 8 mA, as shown in Fig. 11.

To integrate with a CMOS circuits, the driving voltage of the actuator should be lower than 3 V. However, the resistance increase would raise the driving voltage beyond this limit. To reduce the resistance changes with the applied current, a thicker metal multilayer of Cr, Cu, and Au with the respective thickness of 30, 600, and 100 nm was used, the Joule heat generated by the driving current decreases, the resistance change with the current is dramatically reduced, as shown in Fig. 11, the resistance of the actuator at 8 mA was only 1.05 times of the original value (smaller than 300  $\Omega$ ). Therefore, the driving voltage is smaller than 2.4 V, below the limiting value for CMOS circuits even at the maximum driving current of 8 mA. Generally, a lateral displacement of more than 47  $\mu\text{m}$  was achieved with a magnetic flux density of 0.3 T and a current of 8 mA. A linear dependence of the displacement on the current was achieved for the actuator. Much larger displacement can be obtained by increasing the magnetic field and the current.

In spite of the apparent advantages of these actuators, there are still several limitations for the design. For instance, a complex architecture is needed in order to achieve low  $k$ . And the stress matching between the Si and metal layers are not

**Fig. 11** Dependence of the resistance on the applied current from 0 to 8 mA



satisfactory and results in reduced fabrication yield. Moreover, SOI wafers are needed to make homogenous beams, thus increases the fabrication costs.

Compared to single crystal Si, the  $\text{SiO}_2$  film has much lower Young's modulus and larger compressive stress. Therefore, the  $\text{SiO}_2$  film has potential to produce the bilayer beam with small  $k$  and to achieve good stress matching with metal layer. Keeping this in mind, an improved novel MEMS actuator was developed (Fedder 1994). A straightforward structure consisting of  $\text{SiO}_2$  and aluminum has several advantages, such as plain folded beam structure, large displacement under a low driving current/voltage, superb stress matching between the two layers, and high fabrication yield (nearly 100%). These advantages ensure excellent reliability and potential application for large array systems.

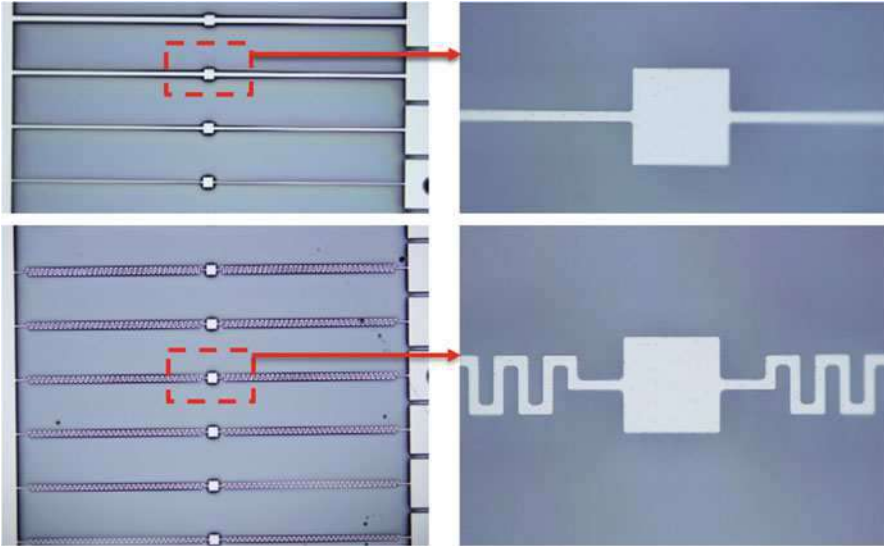
The fabrication process started with a (100)-oriented Si wafer. Thermal oxidation was done to produce  $0.9 \mu\text{m}$  thick silicon dioxide, then Al layer was deposited by electron beam evaporation and was patterned by dry etching. The reactive ion etching (RIE) was employed to remove the Al and  $\text{SiO}_2$  layer. Subsequently, the silicon substrate was etched from the backside by deep reactive ion etching (DRIE) until  $50 \mu\text{m}$  thick Si remaining. Finally, the silicon substrate was isotropically etched from the front side with the top Al layer as the etching mask and the actuator were released (Lv et al. 2015b).

Figure 12 shows the optical photographs of different  $\text{SiO}_2/\text{Al}$  actuators. The simple fabrication process results in a high yield (nearly 100%).

A smaller magnetic flux density of about 0.14 T is imposed on the MEMS actuator. A digital current source was used to provide the driving current.

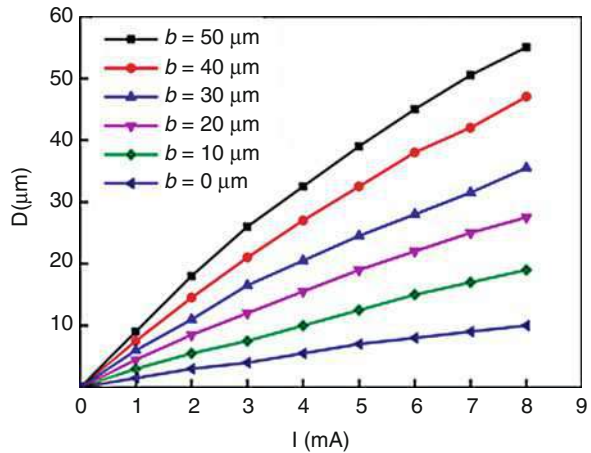
The relationship between the displacement of the actuator and the flowing current of 0–8 mA with a magnetic flux density of 0.14 T is depicted in Fig. 13. The displacement increased with the current.

The displacement of the actuator at 8 mA and a magnetic flux density of 0.14 T changed linearly with  $b$ , as shown in Fig. 14, which agrees well with the theoretical results.



**Fig. 12** Optical photographs of different actuators

**Fig. 13** Dependence of displacement on applied current from 0 to 8 mA with a magnetic flux density of 0.14 T

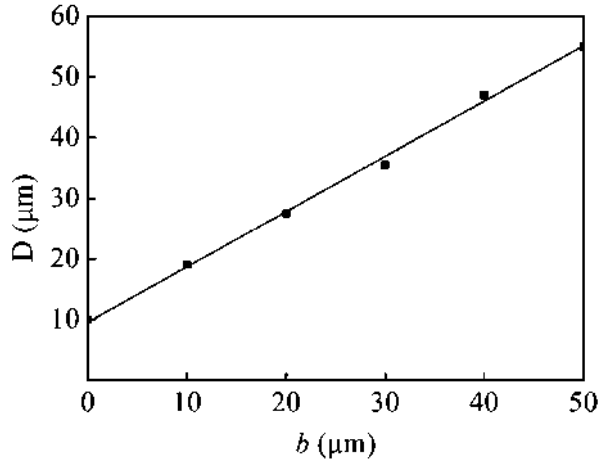


A lateral stroke of more than 55 μm was obtained for the actuator with  $b = 50 \mu\text{m}$  driven by a current of 8 mA, which coincided with the calculation and simulation results. The displacement is proportional to the applied current and switch to the opposite direction under a reversed current, as shown in Fig. 15.

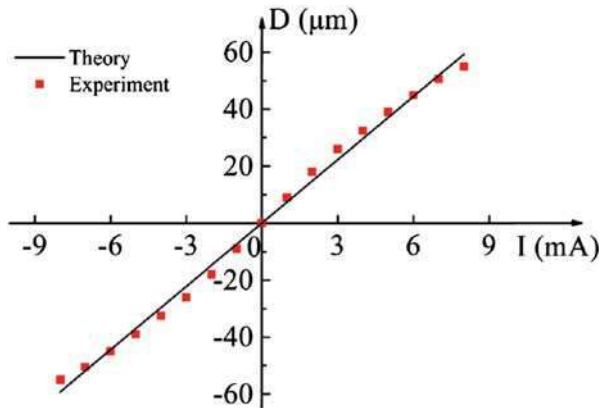
The cycled displacements versus the driving current are shown in Fig. 16. The actuator has excellent stability and could be precisely controlled.

When the current flows through the beam containing the metal layer, heating effect is distinct. The actuator beam with 1 μm thick Al originally has a low

**Fig. 14** Dependence of displacement on  $b$  with an applied current of 8 mA and a magnetic flux density of 0.14 T



**Fig. 15** Dependence of displacement on applied current from  $-8$  to 8 mA with a magnetic flux density of 0.14 T

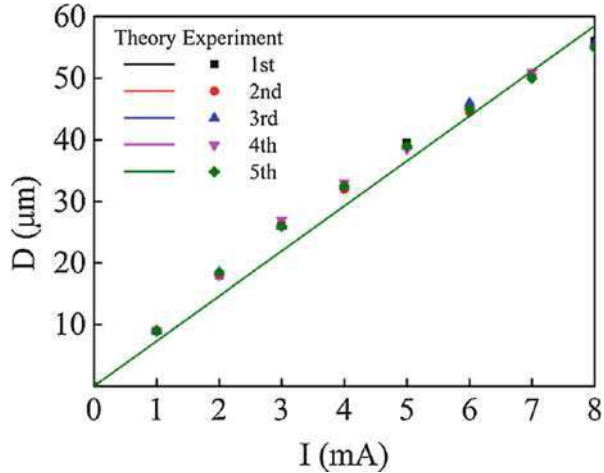


resistance, about  $85 \Omega$ . The resistances of metals increase with temperature. When applying a current on the actuator, Joule heating takes place, changes the resistance and elastic stiffness of the actuator, and finally influences the current dependence of the displacement.

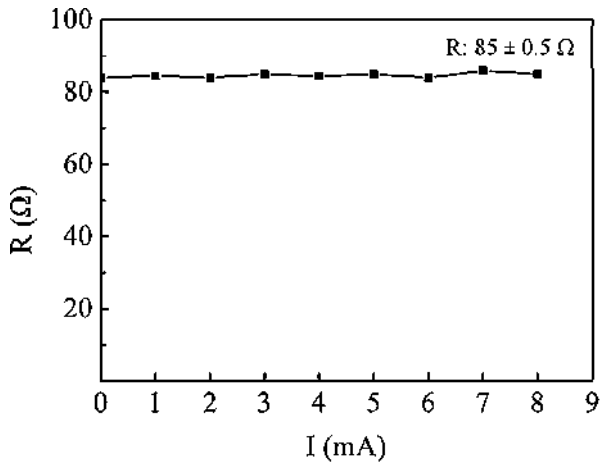
In order to clarify the effect of Joule heating on the displacement of the actuator, the current dependence of the resistance and the time dependence of the resistance and displacement were studied, as shown in Figs. 17, 18 and 19. The resistances of the actuator driven by a current of  $0 = 8$  mA remain unchanged with slight drifts ( $85 \pm 0.5 \Omega$ ), as shown in Fig. 17. Moreover, while driven with a current of 8 mA for 2 h, the resistances of the actuator still keep unchanged ( $85 \pm 0.1 \Omega$ ), as shown in Fig. 18.

When the temperature rises, there will be thermal stress resulting from the different thermal expansion coefficients of Al and  $\text{SiO}_2$ . The thermal stress may

**Fig. 16** Cycling test for the actuator driven by a current of 0–8 mA under a magnetic flux density of 0.14 T



**Fig. 17** Dependence of the resistance on the applied current from 0 to 8 mA

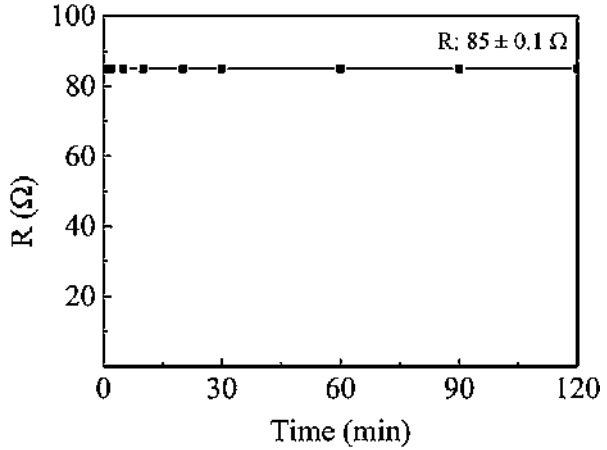


lead to stiffness change of the actuator and finally the displacement change with the temperature. However, the measured displacements of the actuator driven by a current of 8 mA for 2 h keep constant ( $55 \pm 0.5 \mu\text{m}$ ), as shown in Fig. 19.

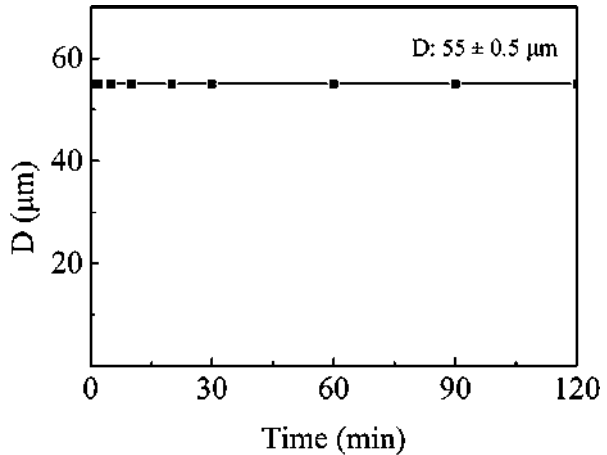
The generated thermal power by the maximum driving current of 8 mA is 5.44 mW. The abovementioned results indicate that thermal dissipation in the suspended bimetallic structure of Al and  $\text{SiO}_2$  balances Joule heat, which results in no clear temperature rising.

Generally, by optimizing the structure and geometry of the actuator, a lateral displacement of more than 55  $\mu\text{m}$  was achieved with a magnetic flux density of

**Fig. 18** Dependence of the resistance on the time with the current of 8 mA



**Fig. 19** Dependence of the displacement on the time with the current of 8 mA



0.14 T and a current of 8 mA. The high fabrication yield and good stability ensure potential application of the actuator in a large array system.

## Potential Applications

Much effort has been made to achieve large displacement under low driving voltages. The MEMS actuator driven by Lorentz force can be applied as optical switches, VOA, microscanners, and so on. The developed MEMS actuator shows a linear dependence of the lateral displacement on the driving current and has high fabrication yield and excellent stability. These promising performances pave the way to the large array system in the future.



## References

- Fedder GK (1994) Simulation of microelectromechanical systems, University of California. *Ther Ber* 32(4):293
- Grade JD, Jerman H, Kenny TW (2003) Design of large deflection electrostatic actuators. *J Microelectromech Syst* 8:2–9
- Hwang IH, Shim YS, Lee JH (2003) Modeling and experimental characterization of the chevron-type bi-stable microactuator. *J Micromech Microeng* 13:948–954
- Jeong SH, Jong SK, Jan GK (2004) Structural optimization of a large-displacement electromagnetic Lorentz force microactuator for optical switching applications. *J Micromech Microeng* 14:1585–1596
- Judy JW, Muller RS, Zappe H (1995) Magnetic microactuation of polysilicon flexure structures. *J Microelectromech Syst* 4:162–169
- Ko JS, Lee ML, Lee DS, Choi CA, Kim YT (2002) Development and application of laterally driven electromagnetic microactuator. *Appl Phys Lett* 81:547–549
- Lee FY, Tang TL, Fang W (2011) Development of a novel dual-axis large-displacement microstage using Lorentz force actuators and curved-beam springs. *Process Eng* 25:689–692
- Lin SJ, Lee CC, Sung WL et al (2013) Displacement enhancement of 1-axis Lorentz force actuator. *Trans Eurosens XXVII*:1591–1594
- Lv XD, Wei WW, Mao X, Yang JL, Yang FH (2015a) A novel MEMS actuator with large lateral stroke driven by Lorentz force. *J Micromech Microeng* 25:025009
- Lv XD, Wei WW, Mao X, Chen Y, Yang JL, Yang FH (2015b) A novel MEMS electromagnetic actuator with large displacement. *Sensors Actuators A* 221:22–28
- Qui J, Lang J H, Slocum AH., Strumpler R (2003) A high-current electrothermal bistable MEMS relay. *IEEE MEMS* 64–67
- Senturia SD (2001) *Microsystem design 3*. Kluwer, Boston
- Sun CM, Wu CL, Wang C et al (2012) Implementation of complementary metal-oxide-semiconductor microelectromechanical systems Lorentz force two Axis angular actuator. *Jpn J Appl Phys* 51(6S):06FL09



# A MEMS Inertial Switch Based on Nonsilicon Surface Micromachining Technology

Zhuoqing Yang, Guifu Ding, Yan Wang, and Xiaolin Zhao

## Contents

Introduction .....	946
Basic Physical Model and Working Principle .....	947
Basic Category and Progress of MEMS Inertial Switch .....	950
MEMS Inertial Switch Based on Non-silicon Surface Micromachining .....	954
Vertically Driven Inertial Microswitch .....	954
Laterally Driven Inertial Microswitch .....	971
Multidirectional Sensitive Inertial Microswitch .....	980
Application Example Demonstration .....	988
Conclusions .....	990
References .....	992

## Abstract

The inertial switch as a kind of passive electric device is also called shock sensor (G-sensor), acceleration switch, vibration threshold sensor, or G-switch. With the developments of semiconductor and integrated circuit technologies, especially, the MEMS-based inertial switches have been attracting much attention due to many advantages such as small size, lower costs, and large volume production. And they are widely used in many applications such as accessories, toys, the transportation of special goods, automotive electronics, remote monitoring (RMON), Internet of Things (IoT) fields, etc. In this chapter, the basic physical model and the working principle of MEMS-based inertial switch is presented firstly. Then the latest progress of the MEMS inertial switch is introduced. Subsequently, the MEMS inertial switch based on non-silicon surface

Z. Yang (✉) · G. Ding · Y. Wang · X. Zhao

National Key Laboratory of Science and Technology on Micro/Nano Fabrication, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University (SJTU), Shanghai, China

e-mail: [yzhuoqing@sjtu.edu.cn](mailto:yzhuoqing@sjtu.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,  
[https://doi.org/10.1007/978-981-10-5945-2\\_31](https://doi.org/10.1007/978-981-10-5945-2_31)

945

micromachining technology is described in detail, including its design, simulation, fabrication, and characterization. In addition, the inertial microswitches with different sensitive directions are proposed and fabricated, including the triaxial inertial switch and omnidirectional sensitive ones. Finally, a simple application example of the fabricated MEMS inertial switch is also performed for potential vibration monitoring module and system applications.

---

**Keywords**

MEMS · Inertial microswitch · Surface micromachining · Non-silicon · Electroplating

---

## Introduction

The inertial switch is also called shock sensor (G-sensor), acceleration switch, vibration threshold sensor, or G-switch. The traditional inertial switch is intended to disable an electric fuel pump in automotive applications. This functionality is required in some vehicle racing applications, since an electric fuel pump may otherwise continue operating after a collision or rollover ([https://en.wikipedia.org/wiki/Inertial\\_switch](https://en.wikipedia.org/wiki/Inertial_switch). Accessed 19 Jun 2017). With the developments of semiconductor and integrated circuit technologies, especially, the arisen microelectromechanical systems (MEMS) technology, the MEMS-based inertial switches have been attracting much attention due to many advantages such as small size, lower costs, and large volume production (Ma et al. 2003). They are widely used in many applications such as accessories, toys, the transportation of special goods, automotive electronics, remote monitoring (RMON), etc. (Ongkodjojo and Tay 2006; Zimmermann et al. 1995; Matsunaga et al. 2002; Michaelis et al. 2000). In the coming years, the rapid growth of Internet of Things (IoT) will lead to a very large demand of MEMS inertial switches which are used to monitoring in the IoT. So the inertial microswitch as a passive sensor device has attracted a great deal of interests. It can be used to monitor the shock level from the environment in order to protect the device against the damage from overload shock in the IoT sensing stage. Figure 1 demonstrates how to use the inertial switch sensor to monitor whether the special goods had been subjected to an unbearable overload shock or not during its transportation in the IoT system. Once the shock from the environment exceeded the allowed threshold level, the inertial switch would put out a pulse signal, which then was sent to information control center by radio-frequency component. As a result, the real-time monitoring of the whole logistics procedure could be realized conveniently. In addition, compared with the accelerometer, the inertial switch has some advantages, such as simpler structure and interface circuit, less power consumption, and lower manufacturing cost. Standard accelerometers always contain a constant parasitic power draw even no acceleration or impact exists. In some unattended and implantable sensors in remote areas or long-life time systems where the power supply is limited, inertial switches with small energy consumption can maintain more endurance before the power would be charged or recharged (Currano et al. 2010).



**Fig. 1** Application of MEMS inertial switch in monitoring the transportation of goods

In this chapter, six sections are organized: the second section will give the basic physical model of the MEMS-based inertial switch and its working principle. Then the latest progress of the MEMS inertial switch will be introduced in the third section. The MEMS inertial switch based on non-silicon surface micromachining will be described in detail in the fourth section, including its design, simulation, fabrication, and characterization. The inertial microswitches with different sensitive directions will be also discussed in this section. In the fifth section, a simple application example of the fabricated MEMS inertial switch will be presented for potential vibration monitoring module. Finally, a main summary and further work about the MEMS inertial switch is given.

---

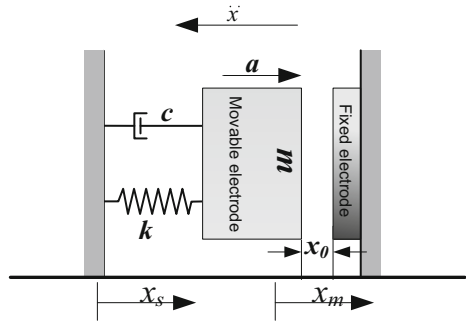
## Basic Physical Model and Working Principle

The inertial switch can realize the closing action by driving an effective mass under the applied inertial acceleration. It is essentially both a threshold sensor and an actuator. The typical inertial switch is a system which consists of two parts: a suspended proof mass as the movable electrode and a fixed component as the stationary electrode. The basic physical model of an inertial switch is usually a “spring-proof mass-damping” system, as shown in Fig. 2, that is, a seismic mass with a finite mass ( $m$ ) and a suspension spring whose mass can be ignored and elastic coefficient is  $k$  and the damping coefficient of system is  $c$ . The equation of motion of the proof mass can be expressed as

$$m\ddot{x}_s + c(\dot{x}_m - \dot{x}_s) + k(x_m - x_s) = 0 \quad (1)$$

where  $x_s$  and  $x_m$  are the displacements of the substrate and the proof mass, respectively, and  $\ddot{x}_s$  is the subjected acceleration of the overall system. When a high enough

**Fig. 2** Basic physical model of an inertial microswitch



acceleration amplitude was applied onto the switch system along its sensing direction, the proof mass electrode would rapidly move forward to the stationary electrode and touch it, switching the external circuit on. Subsequently, the proof mass electrode would be quickly drawn back by spring, switching the circuit off. Considering the substrate as a non-inertial frame, the shock acceleration applied to the proof mass is  $a(t) = -\ddot{x}_s$ , Eq. (1) of motion was simplified to

$$m\ddot{x} + c\dot{x} + kx = ma(t) \tag{2}$$

where  $x$  is the relative displacement between proof mass and substrate, and  $x = x_m + x_s$ . Consider that the acceleration  $a(t)$  applied to the device in practical work is a half-sine wave with amplitude  $a_0$  and impulse width  $t_0$ , as shown in Fig. 3. When we ignore the structural damping, Eq. (2) can be rewritten as

$$m \frac{d^2x}{dt^2} + kx + ma_0 \sin \frac{\pi}{t_0} t = 0 \tag{3}$$

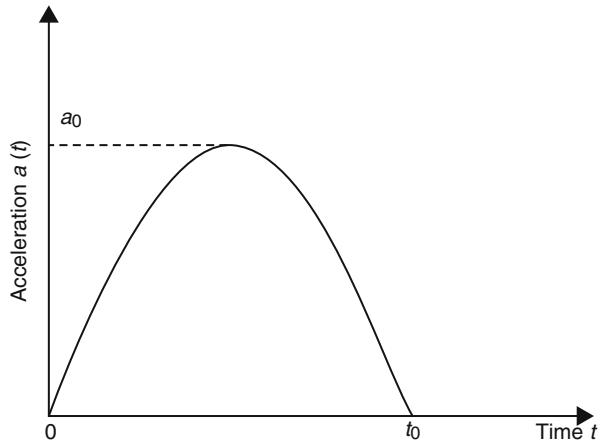
By solving the differential equation (3), we have

$$x = -\frac{a_0}{\frac{k}{m} - \left(\frac{\pi}{t_0}\right)^2} \left[ \sin \frac{\pi}{t_0} t - \frac{\pi}{\sqrt{\frac{k}{m}} t_0} \sin \sqrt{\frac{k}{m}} t \right] \tag{4}$$

The differential is

$$\begin{aligned} \dot{x} &= -\frac{a_0}{\frac{k}{m} \frac{t_0}{\pi} - \frac{\pi}{t_0}} \left( \cos \frac{\pi}{t_0} t - \cos \sqrt{\frac{k}{m}} t \right) \\ &= -\frac{2a_0}{\frac{k}{m} \frac{t_0}{\pi} - \frac{\pi}{t_0}} \sin \left[ \left( \frac{\pi}{t_0} + \sqrt{\frac{k}{m}} \right) \frac{t}{2} \right] \sin \left[ \left( \frac{\pi}{t_0} - \sqrt{\frac{k}{m}} \right) \frac{t}{2} \right] \end{aligned} \tag{5}$$

**Fig. 3** Schematic of half-sine wave shock acceleration applied to the inertial switch



By solving  $x' = 0$ , we have the extreme points

$$t = \frac{2n\pi}{\frac{\pi}{t_0} \pm \sqrt{\frac{k}{m}}} \quad (n = 0, 1, 2, \dots) \tag{6}$$

when  $\frac{\pi}{t_0} < \sqrt{\frac{k}{m}} < \frac{3\pi}{t_0}$ , in the interval of  $0 \leq t \leq t_0$ ,  $x$  reaches only one maximum and one minimum, where  $x_{\max} = 0$  and  $x_{\min}$  is negative. As a result, the max displacement is the absolute value of  $x_{\min}$ , i.e.,

$$|x|_{\max} = |x_{\min}| = \frac{a_0}{\frac{k}{m} - \left(\frac{\pi}{t_0}\right)^2} \left( \sin \frac{2\pi}{1 + \sqrt{\frac{k}{m}} \frac{t_0}{\pi}} - \frac{\frac{\pi}{t_0}}{\sqrt{\frac{k}{m}}} \sin \frac{2\pi\sqrt{\frac{k}{m}}}{\frac{\pi}{t_0} + \sqrt{\frac{k}{m}}} \right) \tag{7}$$

For a contact of the two electrodes to be made (i.e., to switch on),  $x$  must be greater than or equal to  $x_0$ . The minimum acceleration required to activate the switch is defined to be the corresponding “threshold acceleration”  $a_{\text{th}}$ . By solving  $|x|_{\max} = x_0$ , we have

$$a_{\text{th}} = x_0 \frac{\frac{k}{m} - \left(\frac{\pi}{t_0}\right)^2}{\left( \sin \frac{2\pi}{1 + \sqrt{\frac{k}{m}} \frac{t_0}{\pi}} - \frac{\frac{\pi}{t_0}}{\sqrt{\frac{k}{m}}} \sin \frac{2\pi\sqrt{\frac{k}{m}}}{\frac{\pi}{t_0} + \sqrt{\frac{k}{m}}} \right)} \tag{8}$$

which shows the relationship between the threshold acceleration  $a_{th}$  and the intrinsic frequency  $\omega_0 = \sqrt{k/m}$  in the inertial switch.

---

## Basic Category and Progress of MEMS Inertial Switch

Since 1972 Frobenius et al. from Westinghouse Research Laboratories proposed a microminiature threshold accelerometer, it indicates the emergence of a new type of acceleration detection device (Frobenius et al. 1972). This threshold accelerometer device gradually developed into the so-called inertial switch today. The authors designed the device with a wide threshold acceleration detection range from a few tens of g up to maybe 100,000 g (g means the gravity,  $1\text{ g} = 9.8\text{ m/s}^2$ ), and its final fabricated device is shown in Frobenius et al. (1972). It is a kind of simple metal cantilever beam structure as the threshold sensing element, which realizes a wide range threshold acceleration detection utilizing the designed different long beam with different thresholds. The authors thought that the device could have many applications, for example, as impact acceleration sensors in safety automobiles or centrifugal force sensors in rotary machines since it is small and can be made in batches at low cost. As relatively rough as the structural design and fabrication may seem today, such similar design concept has been affecting the follow-up researchers that even the micro-fabrication technology has a great progress. Such threshold accelerometers with cantilever beams structure have been widely studied, and the corresponding properties have been gradually improved in many works (Petersen et al. 1982; Go et al. 1996; Tønnesen et al. 1997; Selvakumar et al. 1996, 2001; Sun et al. 1997; Younis et al. 2007a, b).

At the same time, with the development of some new micro-fabrication technologies, many kinds of inertial switches have been proposed. Different functional mechanism and new materials are introduced into the design for improving the performance properties of the inertial switches. Here some typical kinds of inertial switches based on MEMS technology are described as a short progress summary for readers to understand easily.

*Inertial Switch Using Squeeze Film Effect to Enhance Contact* Matsunaga et al. from Akebono Brake Industry Co. Ltd., Japan, reported an acceleration switch (inertial switch or G-switch) with extended holding time using squeeze film effect for side airbag systems in 1999 (Matsunaga et al. 1999). And the related design and characterization have been studied in details in their subsequent works (Minami et al. 1999; Matsunaga and Esashi 2002). The device has a glass-silicon-glass structure, and the silicon has a contact part, inside beams, a surrounding seismic mass, and outside beams. When the inertial switch is subjected to an applied acceleration in the sensitive direction, the surrounding mass moves toward the glass. The squeeze film damping force between the surrounding mass and the glass will make the movable electrode and the stationary electrode to have a longer contact time (holding time). As a result, the contact effect between the two electrodes in the inertial switch has

been enhanced. The test result of the fabricated device shows that the contact time is up to 6 ms depending on the acceleration amplitude. However, on the one hand, the larger contact area is needed in the inertial switch in order to realize the squeeze film damping effect. As a result, the dimensions of proof mass and contact plate are larger, which is not conducive to the miniaturization of the inertial switch device. On the other hand, in those inertial switches with high threshold acceleration levels, the damping pull-in force would become weaken compared to the elastic restoring force of the springs. Therefore, a larger damping pull-in force would be required to realize the contact-enhanced function in the inertial switch.

*Inertial Switch with Latching Structure* If a latching structure was used to lock the movable electrode and the stationary electrode, the stable contact and close state could be realized in the inertial switch effectively. Luke J. Currano et al. from the US Army Research Laboratory have designed and fabricated an MEMS inertial switch with latching structure (Whitley et al. 2005, 2012; Wittwer et al. 2008; Currano et al. 2008, 2010; Loke et al. 1991). When the movable electrode impacts and contacts with the stationary electrode under the applied acceleration shock, the two electrodes are locked by the latching structure. This design makes the movable electrode not to return to its initial position after the applied acceleration shock withdraws and equivalently has an infinite long contact time. However, this device is only one-shot. In order to use this device repeatedly, the authors added an unlocking structure that driven the latching electrodes to be open by using the thermal reset actuators. The operation of the device can be divided into four distinct phases: (i) a phase before the mass contacts the latch, (ii) a phase where the two are in contact and the mass is pushing past the latch, (iii) a phase where the mass has moved all the way beyond the latch, and (iv) a phase where the mass has settled back against the latch into its resting state. Although this inertial switch makes sure the contact reliability and prolongs the contact time effectively, the reset process is complicated, the production cost is increased, and the application is also limited.

*Inertial Switch with CNTs for Prolonging Contact Time* J. Lee et al. introduced carbon nanotubes (CNTs) into the electrode surfaces of the inertial switch by thermal chemical vapor deposition process (Choi et al. 2011; Lee et al. 2011, 2012). The deformable CNT-contact electrodes are utilized for inertial switch to extend the contact time and therefore to obtain reliable and stable output signals. When the movable electrode collides with the stationary electrode, the CNT-contact electrodes are elastically deformed, significantly prolonging the contact time. The measured contact time of the inertial switch with the integrated deformable CNT-contact electrodes is 114  $\mu\text{s}$ , while the identical inertial switch without CNT-contact electrodes is 7.5  $\mu\text{s}$ . Although the contact effect has been improved effectively in this design, the fabrication process of the inertial switch with aligned and deformable CNTs is obviously complicated.

*Inertial Switch Using Liquid-Metal Droplet as Movable Electrode* J. Kim et al. designed an inertial switch that uses a microscale liquid-metal droplet in a



microstructured channel as the active movable electrode to sense the applied acceleration (Yoo and Kim 2009; Park et al. 2010; Yoo et al. 2011). A liquid-metal droplet is placed in the selectively modified V-shape channel as an initial off state. Once the applied acceleration is more than the threshold level of the inertial switch, the droplet passes through the neck and switches on the sensing electrodes. Using this design, the inertial switch can realize a good contact effect and stable close state. However, the used liquid-metal mercury is highly toxic, which brings a certain risk during the device fabrication. Moreover, the recycle of the switch device is difficult after failure. Once the recycle of those abandoned inertial switches is omitted, there will be a lot of bad effects on the environment.

*Inertial Switch with Threshold Adjusting by Electrostatic Force* An inertial switch using electrostatic pull-in force to enhance the contact effect was reported in 2007 (Jia et al. 2007). By using electrostatic force combined with the dynamic shocking force, the switch features a threshold adjustable capability from 1000 to 5000 g range. The inertial switch can still keep a close state utilizing the electrostatic force between the movable and the stationary electrode after the applied shock acceleration disappears. The schematic of designed cantilever-type inertial switch with electrostatic force and corresponding SEM of the fabricated device array can be found in Jia et al. (2007). The authors state that the electrostatic force can compensate for the threshold inaccuracy from fabrication tolerance and improve the contact effect. However, to the inertial switch with low threshold level, the effect of the introduced electrostatic force on the threshold precision would become obvious and non-ignorable in the present design.

*Inertial Switch with Bistable Postbuckling Structures* J. Zhao et al. design an inertial switch based on the postbuckling theory of large deflection beams for airbag restraint systems (Zhao et al. 2006, 2007, 2010). In the design, the proof mass as the movable electrode is supported by two oblique beams. When the inertial switch is subjected to the applied acceleration shock along the vertical direction, the proof mass will move toward the bottom stationary electrode and contact with each other. As a result, the two oblique beams will have buckling deformation, and a long contact time can be obtained during the postbuckling state in the inertial switch. To the designed inertial switch with threshold level of 20 g, it has  $-4$  ms response time, and the contact time between the movable electrode and the stationary electrode is up to  $-10$  ms. Theoretically, this design of inertial switch would have an infinite long contact time if it had not a reset operation. But in practice, the stiff impact between the proof mass and the stationary electrode may lead to the bounce phenomenon, which is especially obvious in those inertial switches with high threshold level. So in this kind of inertial switch, the safety performance will be an important issue.

*Inertial Switch with Contactless Mechanism* In the traditional design concept of the inertial switch, the contact between movable and stationary electrodes is usually utilized to realize the switch-on state. A. Ongkodjojo et al. propose a novel design concept of inertial switch based on the contactless mechanism triggering. It

essentially integrates a bimorph actuator and a field effect transistor (FET) (Ongkodjojo and Tay 2006). The inertial switch is proposed with a minimum acceleration of 4.8 g for the detection of person fall activities in the healthcare applications. The cantilever bimorph with gold layer is used as the movable gate electrode to form the FET. When the inertial switch senses an applied shock acceleration, which is higher than the threshold level, it deflects toward the gate insulating layer. The deformation will make the current flow between the source (S) and the drain (D), which turns on the FET for the inertial switch triggering. Unfortunately, the authors didn't show the fabrication and experiment measurement of the proposed device, only the design concept and simulation optimization method was given.

*Inertial Switch with CMOS Circuit* With the development of the integrated circuit (IC) technologies, it offers the possibility of integrating mechanical sensor structures on the ASIC layer based on the post-CMOS process. Researchers from University of Bremen and Infineon Technologies reported an inertial switch fabricated with an electroplating technology which can be integrated on top of a pre-processed CMOS signal processing circuit (Ongkodjojo and Tay 2006; Wycisk et al. 1999, 2000; Michaelis et al. 1999). The movable complicated microstructures can be vertically integrated with bottom signal processing circuit, where the switch-on signal from the inertial switch is put into the circuit and then the triggering command will be put out after completing the signal processing in the CMOS circuit.

*Inertial Switch with Triaxial Sensitivity* In order to meet the requirement from the application environment, some inertial switches with multidirectional sensitivity have been proposed in these years. Recently, L. J. Currano et al. from the US Army Research Laboratory again designed and fabricated a triaxial inertial switch with an annular mass and a centrally clamped suspension composed of four spiral springs (Currano et al. 2013). Six contact electrodes separately distribute around the annular mass, which are used to detect the accelerations from three main axial directions ( $\pm x$ ,  $\pm y$ ,  $\pm z$ ). The contact time of the inertial switch at 50 g acceleration shock is about 200  $\mu\text{s}$  in z-axis direction. The z-axis threshold level is significantly different with x- and y-axis due to the difficulties in design and optimization. In addition, some bounces still occur in the inertial switch when the electrodes contact with each other.

*Summary* With the developments of microelectronics, sensor network, MEMS, and especially IoT technologies, more and more sensors and actuators with ultralow power consumption are widely needed. This exactly brings a large space for development of the inertial microswitch that has almost no power at the normal state. Even at the moment of pulse outputting (switch-on), the inertial microswitch only consumes extremely low power. So this passive MEMS inertial switch becomes a good choice to those low-power microsystem applications. There have been many novel designs and fabrication technologies proposed by researchers from different fields as abovementioned. The further more researches will focus on how to improve

the performances of the inertial microswitch, especially dynamic response and axial sensitivity, axis disturbances, low-cost batch fabrication, ultralow power consumption, long-life endurance, systematic integration, and better adaptability in the complicated environment.

In brief, the basic development trends of the inertial microswitch can be generally divided into the following several points: (1) The sensitive direction gradually trending multidirectional sensitivity instead of single directional sensitivity, (2) Focus on how to improve the axial sensitivity of the single-axis microswitch and lower the axial disturbances in the multi-axis microswitch, (3) Trend to develop the inertial microswitch from with single threshold to with multithresholds or adjusting thresholds, (4) Develop the low threshold level inertial microswitch with high precision, (4) Design and fabricate the inertial microswitch with good overall performance for universal applications, and (5) Trend to integrate the microswitch with CMOS circuits for systematic integration and application.

---

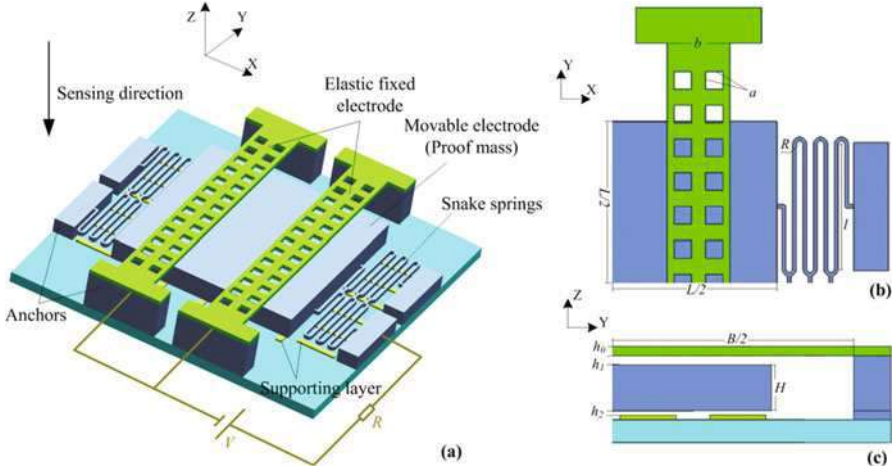
## **MEMS Inertial Switch Based on Non-silicon Surface Micromachining**

In this section, the MEMS inertial microswitch based on non-silicon surface micromachining technology will be proposed and fabricated. Different from the traditional bulk-silicon etching process, here the low-temperature photoresist modeled multi-layer metal electroplating technology is used to complete the inertial microswitch prototypes on a quartz or glass wafer substrate. Some typical novel inertial microswitch designs including vertically-driven and laterally-driven configurations; single-axis, multi-axis, and omnidirectional sensitive properties; and CNTs/Cu composite flexible for contact-enhanced will be given and described generally. The detailed simulation, optimization, and measurement of every inertial microswitch can be also found in our related reports.

### **Vertically Driven Inertial Microswitch**

#### **With Compliant Bridge-Type Beam as Stationary Electrode**

*Device design:* The switch was designed as a single-step acceleration device with regard to the surface micromachining technology. It consists of three main parts: a thick suspended proof mass electrode with conjoined snake springs (the sensing structure), the parallel beams with holes above the proof mass as a bridge-type beam compliant fixed electrode, and a short pillar array as the supporting layer. Anchors fixed on the substrate suspend the springs and the parallel beams. Figure 4 shows the sketch and the parameters of the designed inertial microswitch (including an external circuit). There is 20  $\mu\text{m}$  distance ( $h_1$ ) between the fixed electrode and the proof mass, keeping the circuit off in a static state. When sufficient acceleration amplitude is applied in the sensing direction, the proof mass rapidly moves toward the bridge-type beam stationary electrode and touches it, switching the external circuit



**Fig. 4** Schematic of the designed inertial switch with compliant beam as stationary electrode. (a) 3D sketch, (b) top view, and (c) side view of a quarter of the inertial switch

on. Furthermore, the proof mass, drawn back by snake springs, returns to the off position, switching off the external circuit.

The designed switch shown in Fig. 4 has several advantages. Firstly, the thick proof mass is enough to cause a large inertial force under the applied acceleration, improving the sensitivity of the switch. Secondly, the springs are designed in two conjoined forms instead of four separate ones to facilitate the coordination movement of the proof mass and springs. Thirdly, the 3–4 μm gap ( $h_2$ ) between the proof mass and the supporting layer limits the proof mass rebound after impacting the stationary electrode, protecting the switch from intensive shock damage. Moreover, a series of holes in the elastic beams can decrease the stiffness of the fixed electrode and increase the contact time between the two electrodes. In addition, these holes also provide convenience for the final release of the device.

*Finite element (FE) simulation:* A simple physical model for the inertial switch is shown in Fig. 2. When acceleration is applied to a spring-mass inertial system in the sensing direction,  $\ddot{x}$ , with an applied force,  $F$ , the resulting motion of the seismic mass can be expressed with the dynamic equilibrium equation as

$$m\ddot{x} + c\dot{x} + kx = F \tag{9}$$

where  $m$  is the seismic mass,  $c$  presents the damping coefficient,  $x$  is the moving distance, and  $k$  is the effective spring constant. And this formula is substantially the same as Eq. (1). In a subsequent experimental test of the fabricated inertial switch, a potential (−3 V) will be applied between the movable plate and the fixed electrode. Thus, the applied force ( $F$ ) in Eq. (9) will include an electrostatic force between the two plates, which is expressed as

$$F_{electrostatic} = \frac{\epsilon_0 A}{2(d_0 - x)^2} V^2 \quad (10)$$

where  $d_0$  is the initial gap between the two electrodes, i.e.,  $h_1$ ,  $x$  is the moving distance,  $V$  is the applied voltage on the two electrodes,  $A$  is the opposite area between the two electrodes, and  $\epsilon_0$  is the dielectric constant of air. The dimensions of the designed inertial switch are  $L = 900$ ,  $l = 280$ ,  $a = 50$ ,  $b = 160$ ,  $h_0 = 20$ ,  $h_1 = 20$ ,  $H = 100$ ,  $R = 15$ , and  $B = 550 \mu\text{m}$ . The thickness of the conjoined snake spring is  $10 \mu\text{m}$ . The electrostatic force  $F_{electrostatic}$  can be calculated by substituting correlative values into Eq. (10). Figure 15 shows the electrostatic force and the elastic force in the motion of the proof mass in the inertial switch. The elastic force,  $F_{elastic}$ , from the deformation of the spring can be calculated as

$$F_{elastic} = kx \quad (11)$$

where the effective spring constant  $k$  can be obtained by FE simulation in ANSYS. It can be seen from Fig. 5 that the elastic force is much higher than the electrostatic force on the seismic mass until the distance between the two electrodes is less than  $0.1 \mu\text{m}$  (point of intersection  $A$ ). Therefore, we will neglect the effect of the electrostatic force on properties of the inertial switch.

The FEM was used to simulate and then analyze the dynamic contact process in the deigned inertial switch. A quarter of the switch structure was modeled after considering its symmetry. ANSYS software was used to construct the FE model in a manner consistent with the geometric parameters and dimensions of the device. SOLID45, TARGE170, and CONTA174 element types were chosen for the dynamic contact analysis, and the SWEEP method was used to mesh the model. Literature indicates that electroplated nickel (Ni) has a wide range of mechanical properties, so here the bulk material was set as metal nickel temporarily. Of course, it's better to use the experimental measured value as the material properties. The undersides of the anchors and the supporting layer were constrained in all degrees of freedom (DOF). The symmetry boundary constraint was applied to the model symmetry planes, where Fig. 6 shows the FE model with two contact pairs. When the inertial switch closed, an adhesion force between touching two electrodes will occur in the process of dynamic impact. However, this adhesion behavior is affected from various factors including roughness, structural pattern, contact separation, etc. (Jones et al. 2003; Tas et al. 1996; Rogers and Phinney 2001). It's very complicated and difficult to evaluate by a single aspect. Here we will temporarily exclude the effect of the adhesion in the simulation.

A static FEM simulation was done firstly to obtain the total spring constant for vertical deflection, which is  $25 \text{ N}\cdot\text{m}^{-1}$ . The resonance frequency of the first mode of the spring-mass system is  $850 \text{ Hz}$ , which can be extracted by a FEM modal analysis. In the same way, the obtained fundamental resonance frequency and the spring constant of the bridge-type compliant fixed electrode are  $71,428 \text{ Hz}$  and  $1190 \text{ N}\cdot\text{m}^{-1}$ , respectively. The damping was not temporarily taken into account for simplifying the analysis in the simulation. Ideal half-sine wave (seen in Fig. 3)

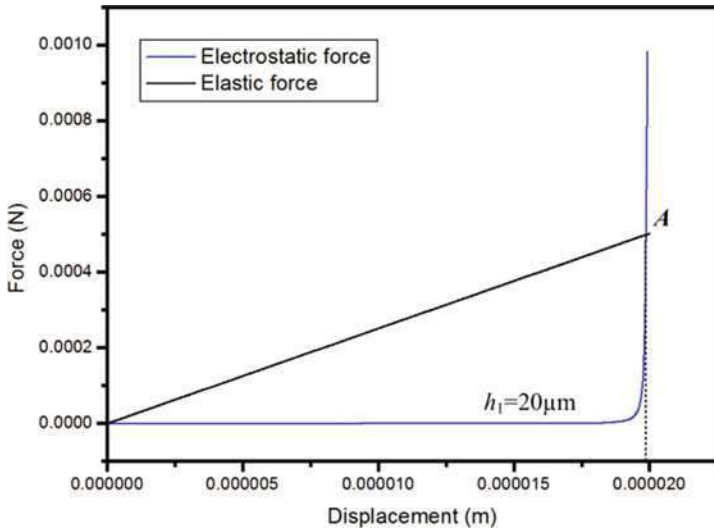


Fig. 5 Electrostatic force between electrodes and elastic force from spring in the inertial switch

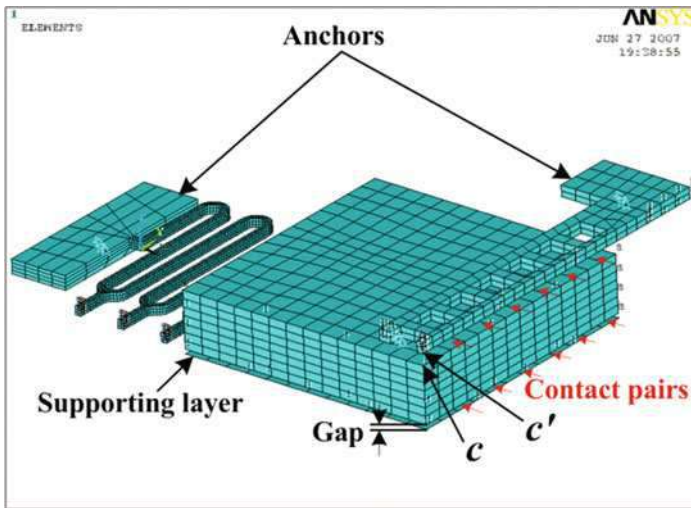
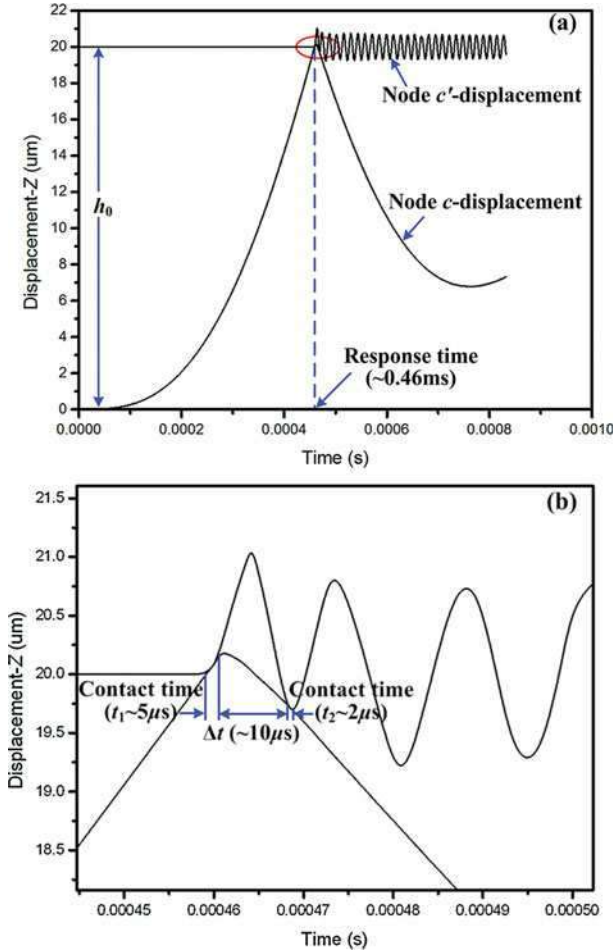


Fig. 6 FE model for dynamic simulation in inertial switch with compliant bridge-type beam

shock acceleration with 55 g amplitude was implemented in the transient contact simulation along the negative  $z$ -axis direction. The dynamic response curves of the microswitch under the applied shock acceleration are shown in Fig. 7. Figure 7a shows the displacement-time curves of nodes  $c$  and  $c'$  in the  $z$ -direction. It can be seen that the inertial switch is closed when 55 g acceleration was applied, and two

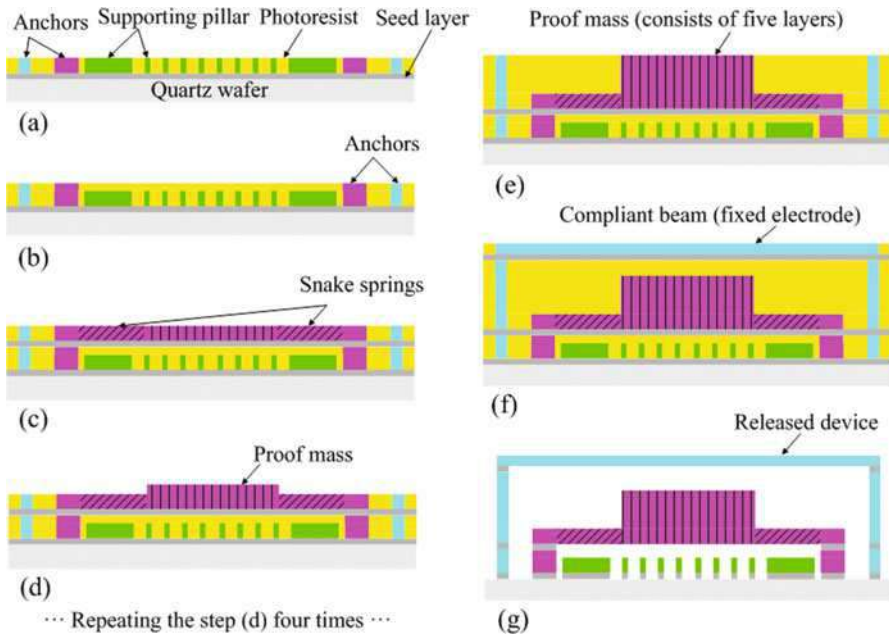


**Fig. 7** Simulated results of the inertial switch under 55 g half-sine wave shock acceleration. (a) Dynamic response displacement-time curves of nodes *c* and *c'*, (b) close-up of the red ellipse plot, simulated contact times and their delay

transitory contacts (bounce phenomenon) occurred during the closing, encircled by the red ellipse. Furthermore, the delay until the two electrodes contact each other, i.e., the response time of the inertial switch, is approximately 0.46 ms. Figure 7b amplifies the curve within the red ellipse, showing that the two transitory contact times ( $t_1, t_2$ ) are about 5 and 2  $\mu s$ , respectively. The delay ( $\Delta t$ ) between the two transitory contacts is about 10  $\mu s$ .

*Micro-fabrication:* The designed inertial microswitch was fabricated on an insulated quartz (or glass) wafer by low-cost metal electroplating technology. The compositions of nickel electrolyte are Ni  $[\text{NH}_2\text{SO}_3]_2(600 \text{ g/L})$ ,  $\text{H}_3\text{BO}_3(25 \text{ g/L})$ ,



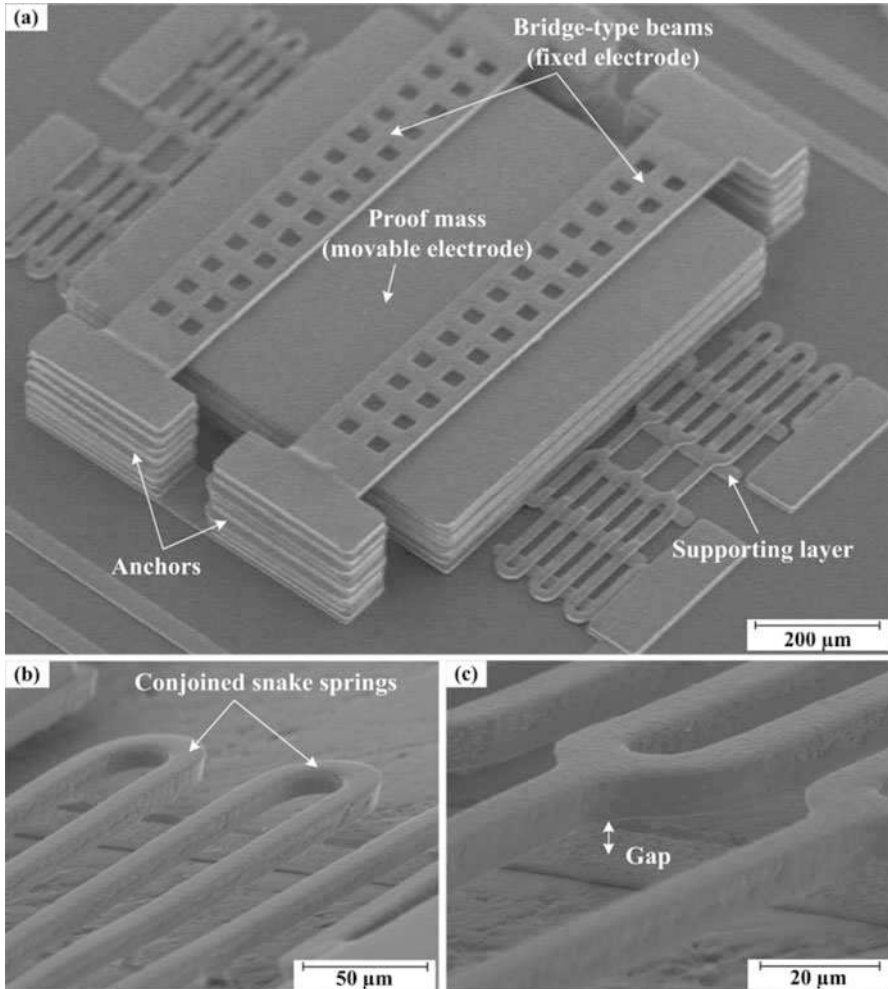


**Fig. 8** Main micro-fabrication process of inertial switch with bridge-type stationary electrode. (a) Chromium/copper (Cr/Cu) was sputtered as the seed layer for the plating of the device structure, then photoresist spin coating and photolithography were completed. The supporting pillar layer and the basic anchors close to the substrate were electroplated in nickel (Ni) as the structural material. (b) Next the anchors for suspending the springs and the proof mass were electroplated in Ni. (c) The snake springs structures were electroplated after secondly sputtering the Cr/Cu seed layer. (d) A 20  $\mu\text{m}$  layer of the proof mass was electroplated. (e) The thick proof mass consisting of five layers was completed by repeating the step (d). (f) Next the bridge-type beams as the compliant stationary electrode and the last anchors for suspending them were electroplated. (g) Finally, sacrificial layers were dissolved in 15% NaOH solution and then rinsed in alcohol to obtain the released inertial switches. Rinsed samples were dried on a hot plate in air

and  $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$  (10 g/L). The electroplating solution conditions are pH 4.0, temperature of 40  $^\circ\text{C}$ , and current density of 2.0  $\text{A}/\text{dm}^2$ . The volume of the proof mass is controlled by the electroplating time and monitored by a stylus profiler (Dektak 6 M, Veeco, USA). The main fabrication process steps of the inertial switch are described and sketched in Fig. 8a–g as follows:

The scanning electron micrograph (SEM) of the fabricated inertial switch prototype is shown in Fig. 9a. The structures, such as the snake springs, stack proof mass and supporting pillar layer, etc. can be clearly seen in Fig. 9b. In the close-up SEM, the proof mass and the attached snake springs are suspended as designed. The gap between the supporting layer and the proof mass is also shown Fig. 9c. No stiction phenomenon in the release processes of microswitches was observed, although there was a very small gap between the proof mass and the supporting layer. This result is from a short pillar array distribution designed into the supporting layer as shown in

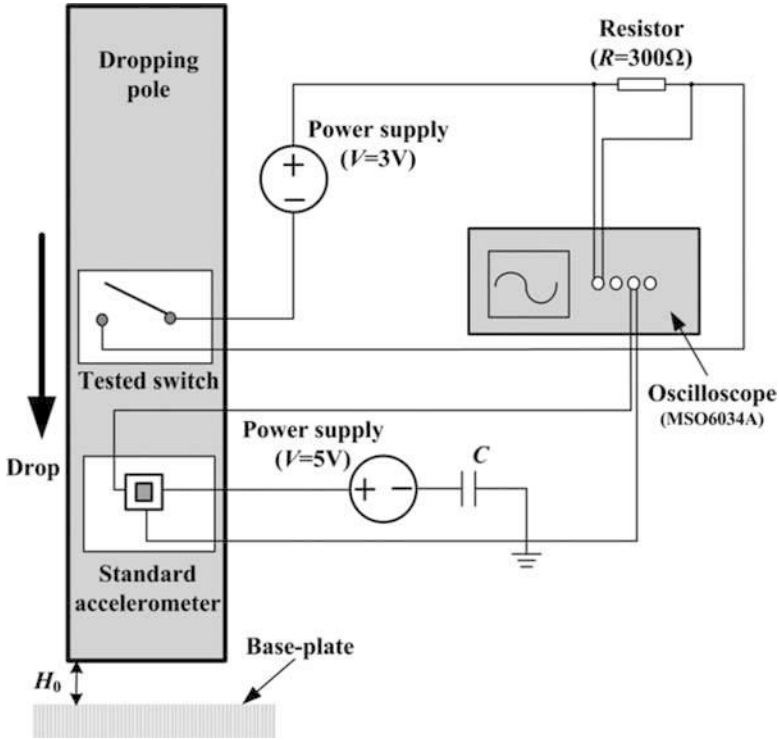




**Fig. 9** (a) SEM of the fabricated inertial switch prototype, (b) suspension conjoined snake springs, (c) the gap between the supporting pillar layer and snake springs

Fig. 4. The inertial switch was sealed using an all-metal cap that was fabricated by UV-LIGA technology. The dimension of the packaged single device by epoxy adhesive bonding is about  $3.2 \times 2.1 \times 1.3 \text{ mm}^3$ , which is smaller than many other reported devices (Ma et al. 2003, 2004; Michaelis et al. 2000; Wycisk et al. 2000).

*Experimental test and discussion:* The packaged inertial switch was tested by the standard drop hammer system setup as shown in Fig. 10. The tested microswitch and a calibrated standard high-g accelerometer (by Analog Devices Inc.) were both fixed onto the dropping steel pole, making sure their sensing directions were perpendicular to the surface of the base plate. During the experiment the steel



**Fig. 10** Sketch of the test setup for the inertial switch device

pole was dropped from a known height and collided with a heavy buffer base plate to generate a half-sine wave shock acceleration applied to the tested switch. In order to avoid the effect of the delay of the accelerometer on the test response time of the inertial switch, an external trigger signal from the pole can be obtained through the base plate to record the time when the impact occurs. The g-level of the switch closure was evaluated by reading the peak value of putout signal from the standard accelerometer. With this setup, the generated shock acceleration signal was collected by the standard accelerometer, which has a sensitivity of 8 mV/g. When both the accelerometer and the tested switch on the pole were simultaneously freely released from different predetermined heights  $H_0$ , the output signals under the different accelerations were detected by a multichannel oscilloscope (MSO6034A, Agilent, USA) shown in Fig. 10. Therefore, the response time is defined as the delay between the impact trigger signal of the steel pole and the close signal of the tested switch, indicating how long it takes for the movable electrode to travel to the stationary electrode.

The test setup was calibrated with the reference standard accelerometer and the g-level set to 60 g. The maximum offset ( $V_{\max}$ ) of the output voltage is  $-480$  mV, as shown in Fig. 11a. It can be seen that there are two main pulse signal outputs from

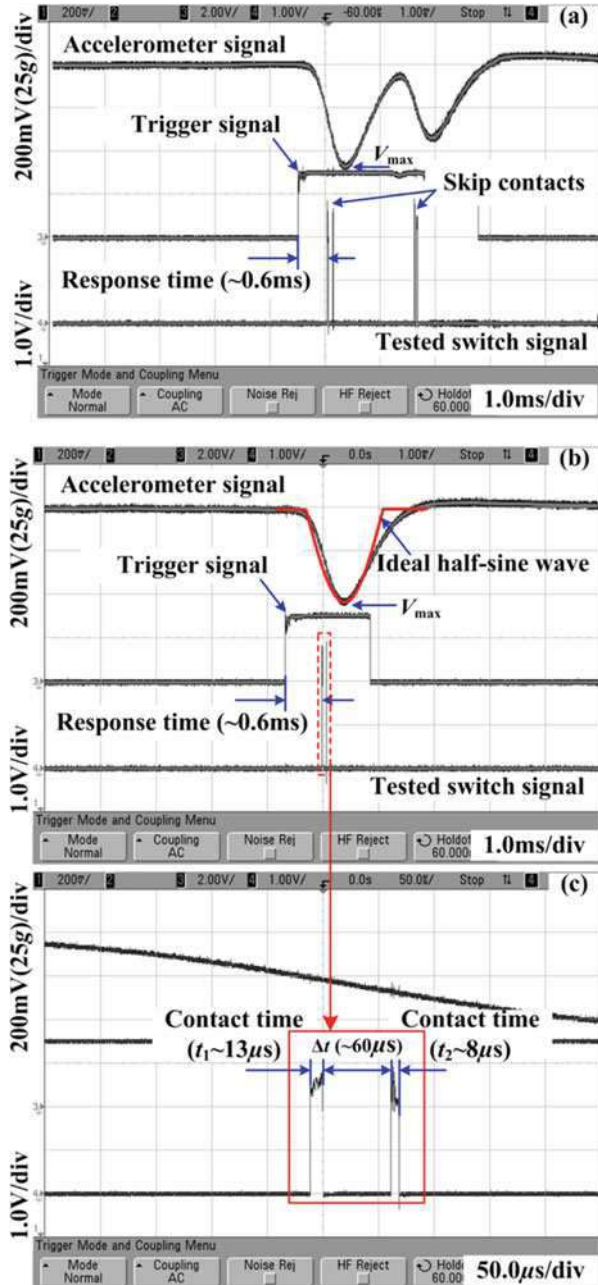
the tested switch, indicating that it had closed two times in response to two successive shock accelerations ( $-60$  and  $-45$  g) in the test. It is worth noting that each main pulse signal actually consists of two secondary pulses. This indicates that the two electrodes of the tested inertial switch transiently contact twice during each closing, which is the bounce (skip contact) phenomenon. In the drop hammer tests of other packaged inertial switches in the same batch, the skip contact phenomenon can also be observed. In order to analyze this phenomenon in detail, a smaller acceleration of  $55$  g was applied to the tested switch by adjusting the height,  $H_0$ . In this case, the  $V_{\max}$  is about  $440$  mV, and there is only one main pulse signal from the switch, shown in Fig. 11b. Furthermore, it can be seen that the main pulse signal still contains two secondary pulses. The response time of the microswitch is approximately  $0.6$  ms as shown in Fig. 11a–b. Two transitory contact times ( $t_1$ ,  $t_2$ ) of the microswitch in the bounce process are  $-13$  and  $8$   $\mu$ s, respectively, and the delay ( $\Delta t$ ) between the two transitory contacts is  $-60$   $\mu$ s as shown in Fig. 11c.

The above measured results have an agreement with the simulation shown in Fig. 7. It is recognized that there is a deviation between the simulated and experimental results (for both the response and contact times). This deviation is mainly because the half-sine wave shock acceleration applied to the switch in the test is not ideal like the simulation as shown in Fig. 11b. The value for bulk Ni was used instead of that for electroplated Ni in the above transient simulation analysis. The electroplated Ni generally has a smaller Young's modulus, leading to a low system spring constant ( $k$ ) in the tested microswitch. Therefore, its vibration frequency ( $\sqrt{k/m}$ ) will be lower, resulting in a longer response time in the test than in the simulation. In addition, the adhesion force between touching electrodes will also enhance their contact during practical work of the switch, which may be a reason that the test contact time is longer than simulated one.

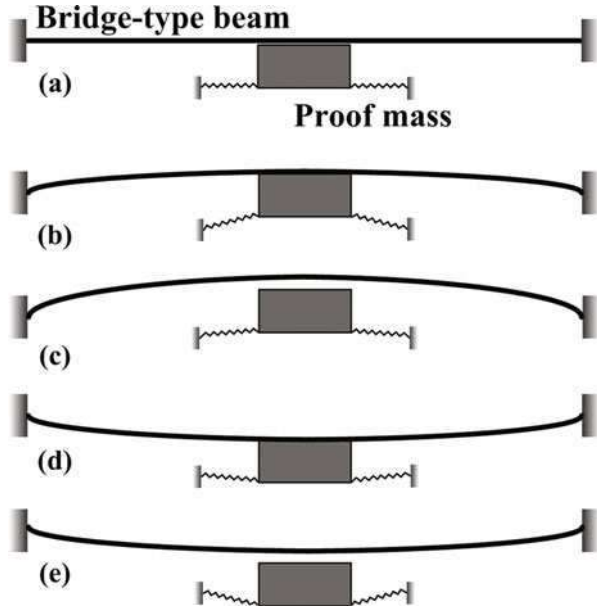
In the simulation we also simulated a situation where a proof mass impacted a rigid fixed electrode instead of a compliant one, which is the case in conventional inertial switch. As a result, the proof mass bounced intensively after impacting the rigid fixed electrode. The contact time during the intensive impact is extremely short (about  $1$   $\mu$ s), even less than the second transitory contact time ( $t_2$ ) in the bounce process. The simulation result indicates that the inertial switch with a bridge-type compliant fixed electrode has an improved contact effect compared with the conventional case with a rigid fixed electrode.

The FE simulation clearly illustrates the dynamic contact process, and the bounce contact phenomenon has been proven in the experiment of the present inertial switch with a bridge-type compliant fixed electrode. The dynamic contact between the two electrodes of the switch is sketched and explained in Fig. 12. Figure 12a, b shows that the proof mass moves toward and starts to impact the bridge-type beam. As a result, the first transitory contact occurs and the contact time is  $t_1$ . In Fig. 12c, the beam then moves up after being impacted, while the proof mass rebounds down. Subsequently, both elements separate from each other. The bridge-type compliant beam begins to vibrate up and down, shown in Fig. 12d. The second transitory

**Fig. 11** Measured results of the inertial switch by the drop hammer system. (a) Bounce (skip contact) phenomenon in the switch at shock accelerations, (b) response time of the switch at 55 g, (c) two contact times and delay in the switch during bounce



**Fig. 12** Schematic diagram of both the impact and the contact between the proof mass (movable electrode) and bridge-type beam (compliant fixed electrode) in the inertial switch: (a) → (b) The proof mass just starting to impact the compliant beam and the first transitory contact occurs; (c) the proof mass rebounds after impacting the beam, and they separate from each other (i.e., the delay  $\Delta t$ ); (d) vibrated beam catches up to the rebounding proof mass, and the second transitory contact occurs; (e) the proof mass and the beam again separate from each other because of their different movement velocities

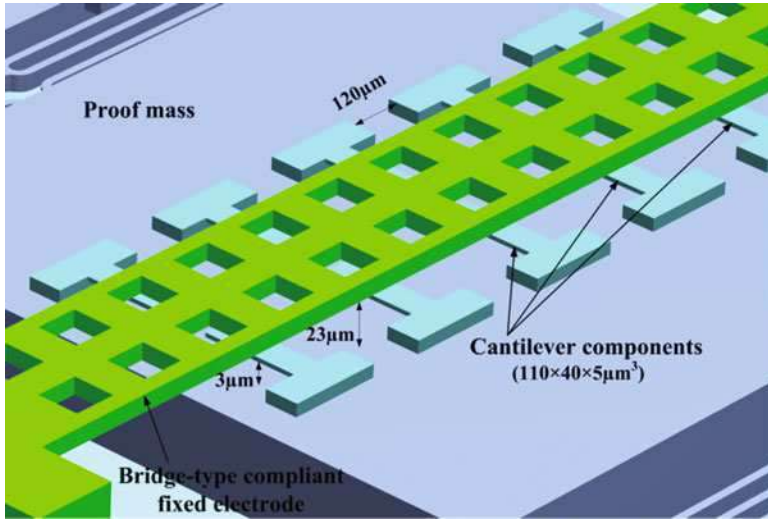


contact occurs when the beam moves down and catches up to the rebounding proof mass with contact time  $t_2$ . The delay between the two contacts is  $\Delta t$ . In Fig. 12e, the bridge-type beam rebounds up under forced vibration, while the proof mass continues moving down. Finally, the beam and the proof mass again separate from each other.

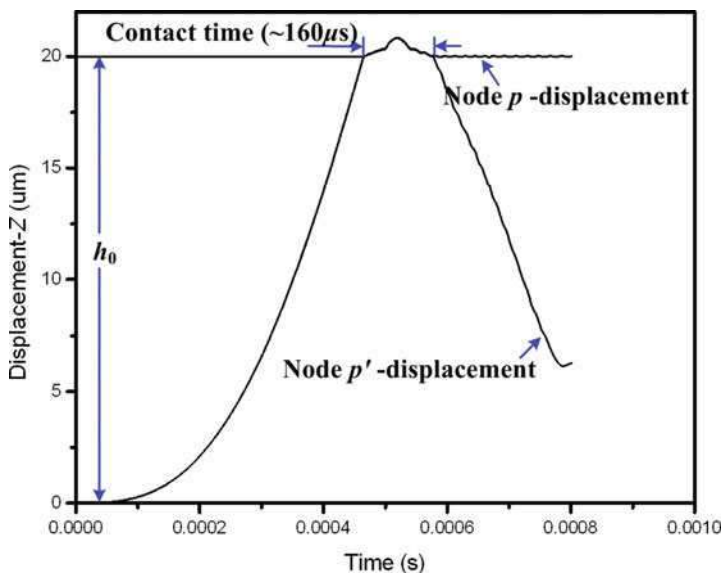
*Improvement:* In order to eliminate the bounce problem and prolong the contact time of the inertial switch, an array of cantilevers was introduced on the proof mass as a buffer between the two electrodes, while the remaining whole switch stayed the same. The stiffness (spring constant) of the cantilevers is much lower than the bridge-type fixed electrode. The structural dimensions and distributions of the cantilever components in the improved inertial microswitch are shown in Fig. 13.

FEM was used again to simulate and analyze the dynamic contact of the inertial switch with cantilevers. More than one contact pair is needed to be constructed in this model because of these cantilevers. This requires more computational time to simulate. Once again, the switch FE model was excited by applying ideal half-sine wave shock acceleration with 55 g amplitude along the negative  $z$ -axis. The dynamic response curves of the improved switch are shown in Fig. 14. It can be seen that the contact characteristic of the inertial switch with cantilevers is improved and the contact time is about 160  $\mu\text{s}$ . In addition, the bounce has been eliminated. It is indicated that the cantilever array has a good buffer effect in the contact process of the two electrodes.

The improved inertial microswitch with cantilevers was still fabricated by surfacing micromachining technology shown in Fig. 8. Only the cantilevers on proof



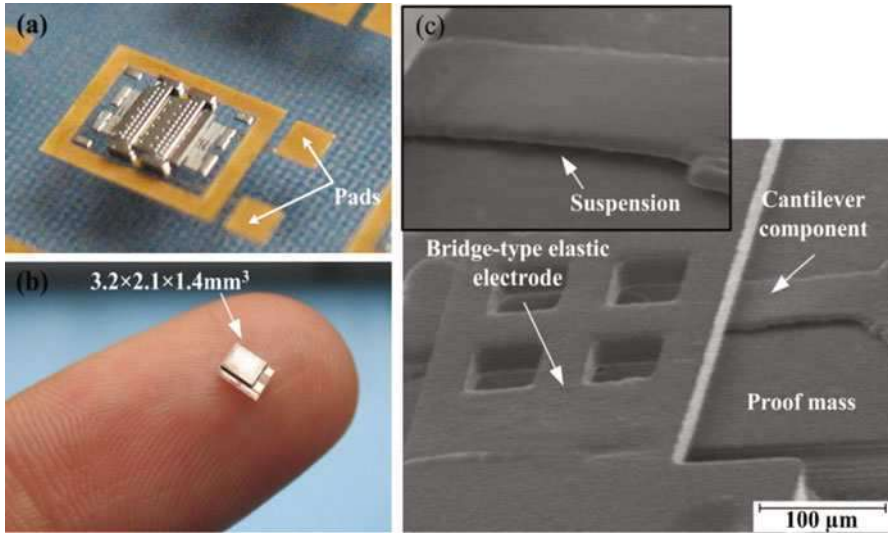
**Fig. 13** Specifications and distributions of the cantilevers in the improved inertial switch



**Fig. 14** Dynamic response displacement-time curves of nodes on two electrodes in the improved inertial switch with cantilevers buffer

mass surface were added in the process. The photograph of the fabricated and the packaged single inertial switch can be seen in Fig. 15a, b. The inertial switch was sealed using an all-metal cap that was fabricated by UV-LIGA technology. The dimensions of the packaged single device by epoxy adhesive bonding at normal





**Fig. 15** (a)–(b) Optical photographs of the fabricated and packaged single inertial switch device, (c) SEM of the improved inertial switch with cantilevers

atmosphere are  $3.2 \times 2.1 \times 1.4 \text{ mm}^3$ . SEM micrograph of the fabricated improved inertial switch with cantilevers is seen in Fig. 15c.

The inertial switch with cantilevers was tested in the same way using the drop hammer system setup. Like the switch without cantilevers, the improved inertial switch with cantilevers was tested in 55 g half-sine wave shock acceleration by setting the predetermined height  $H_0$ . The measured contact results of the switch are shown in Fig. 16. It can be seen that the bounce phenomenon did not occur here. More importantly, the contact time of the inertial switch with cantilevers increased to 240  $\mu\text{s}$ , which is over 15 times that without cantilevers. The cantilevers play an important role as a buffer for the sufficient contact between the movable proof mass electrode and the bridge-type compliant stationary electrode during the impact in the improved inertial switch.

About the abovementioned vertically driven inertial microswitch, in our previous reports (Yang et al. 2007, 2008a, b, 2009, 2010a, 2011a, b; Cai et al. 2009), the dependence of contact time, response time, and threshold level on structural parameters of the switch has been studied in details. The effect of the squeeze film damping on the dynamic response performance has been evaluated. The interconnection resistance and the package adhesive strength in the fabricated inertial microswitch have been also measured.

### With Movable Contact Point for Prolonging Contact Time

*Device design and working principle:* The inertial switch with a movable contact point is illustrated in Fig. 17, which consists of three main parts: the proof mass as the movable electrode, which is suspended by four conjoined serpentine springs

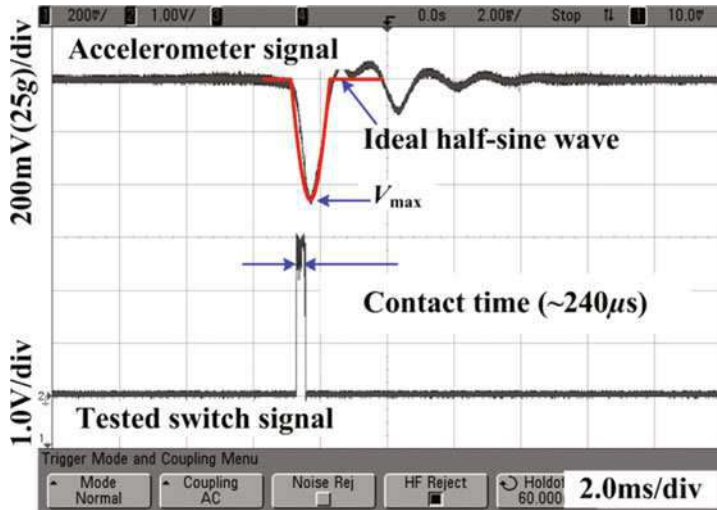


Fig. 16 Measured contact time of the improved inertial switch at the same shock acceleration

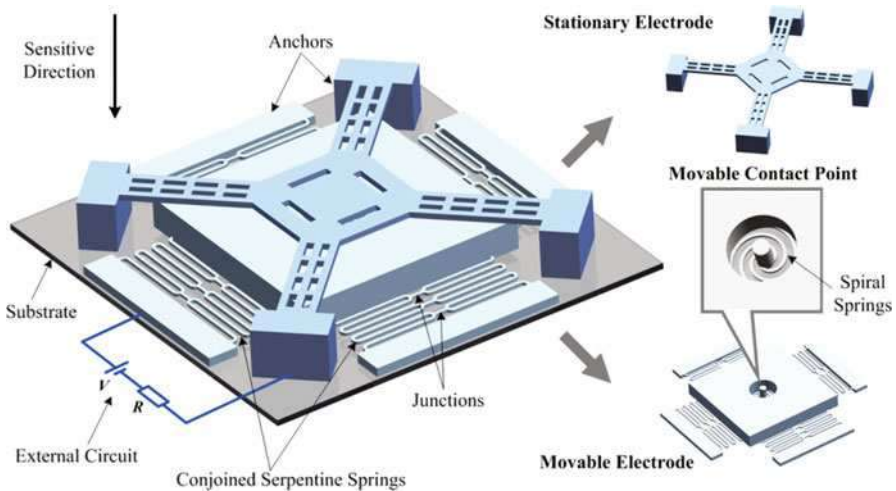
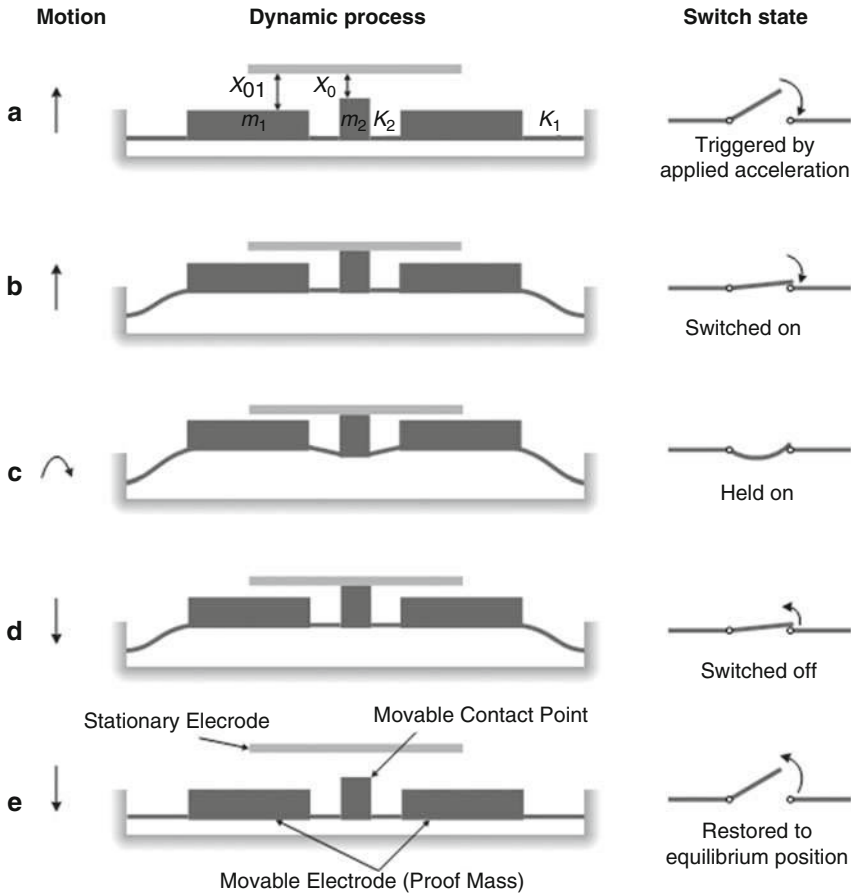


Fig. 17 Schematic of the novel inertial switch design with movable contact point

fixed on the anchors; the movable contact, a novel component which is suspended by spiral springs in the middle of the proof mass instead of fixed on it as our initial design in Fig. 4; and the crossbeam as the fixed electrode, which is a distance of  $x_0$  above the contact point and  $x_{01}$  above the proof mass, respectively. The gap between the proof mass and the substrate is  $x_{02}$ .

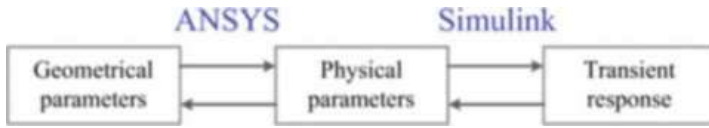
Unlike the abovementioned design in Fig. 4, it is a dual spring-mass system. Apart from the proof spring-mass system ( $m_1, k_1$ ), the contact point ( $m_2$ ) is suspended



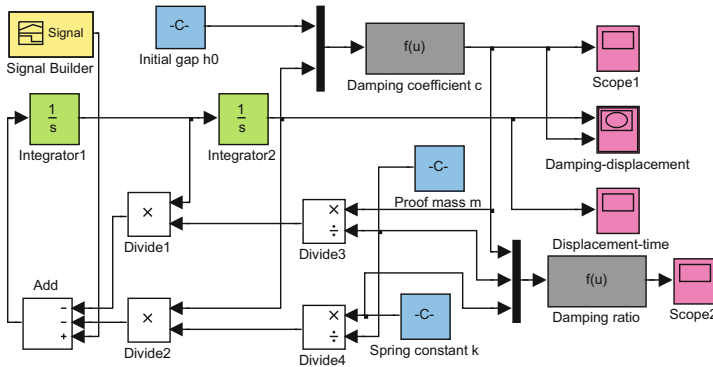


**Fig. 18** Working principle of the inertial switch with movable contact point: dynamic process and corresponding switch state

by the inner spring ( $k_2$ ) in the middle of the proof mass instead of being rigidly fixed on it. The upper stationary electrode has a distance of  $x_0$  above the contact point and  $x_{01}$  above the proof mass, respectively, where  $x_0 < x_{01}$ . The working principle is demonstrated by the dynamic process and corresponding switch state, as shown in Fig. 18: (a) The proof mass moves toward the stationary electrode as the switch is triggered by a shock acceleration. (b) When the acceleration is over the threshold, the displacement reaches  $x_0$ , so the contact point reaches the stationary electrode and the switch is turned on. (c) The proof mass keeps moving on, while the contact point is stopped by the stationary electrode; therefore, the switch-on state is held on for a longer time. (d) After extinction of the acceleration, the proof mass rebounds and pulls the contact point back, so the switch is turned off. (e) Finally, the proof mass and contact point restore to the equilibrium position after all the energy is dissipated by free vibration. In conclusion, the enhanced contact effect and long



**Fig. 19** Simulation method by ANSYS and Simulink combination

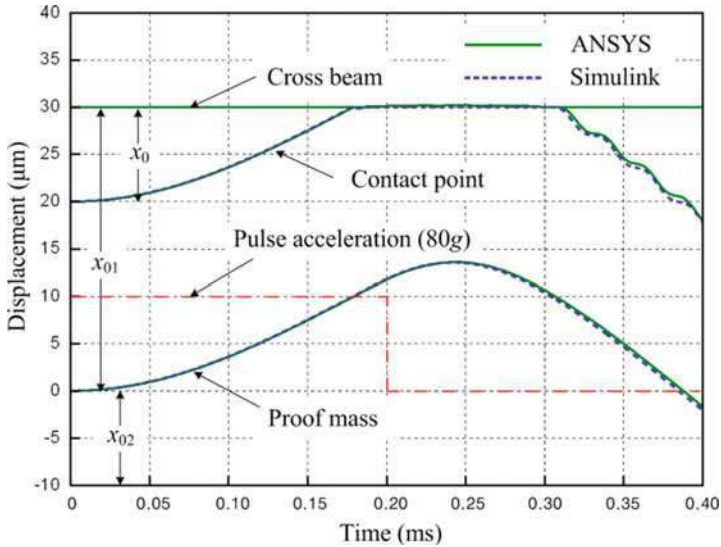


**Fig. 20** Simulink model of the inertial microswitch with movable contact point

contact time are obtained and compared with conventional designs where the proof mass with a fixed contact point rebounds intensively after impact with the stationary electrode, resulting in a transient contact time.

*Combination simulation by ANSYS and Simulink:* Simulation of the inertial microswitch is carried out using the combination of Simulink and ANSYS. ANSYS can be directly applied to transient simulation, but FEM is time consuming and computationally expensive. The dynamic simulation with Simulink is more satisfactory and provides the possibility to divide a system into subsystems without loss of accuracy (Tønnesen et al. 1997). Also, it is more convenient in Simulink to apply various accelerations as the input signal and modify the complicated damping effect. On the other hand, the use of FEM is irreplaceable to calculate the physical parameters for microstructures with complex geometry. So the physical parameters such as  $k$ ,  $m$  are extracted by ANSYS model and then applied in the Simulink model to get transient response of the device, as shown in Fig. 19. The process can be reversed to decide the geometrical parameters for desired device performance, such as the threshold acceleration, the response time, and the contact time.

The block diagram of the system was created in Simulink based on the mathematical description of the switch behavior derived in analytical analysis, as shown in Fig. 20. The output of the system describes the movement of the proof mass and the contact point as a function of the input acceleration, geometric/physical parameters, and time. When the contact point's displacement reaches  $x_0$ , the external circuit is switched on, so we get the output signal. Additional information about the damping and the velocity can also be generated by this simulation. A Simulink model without

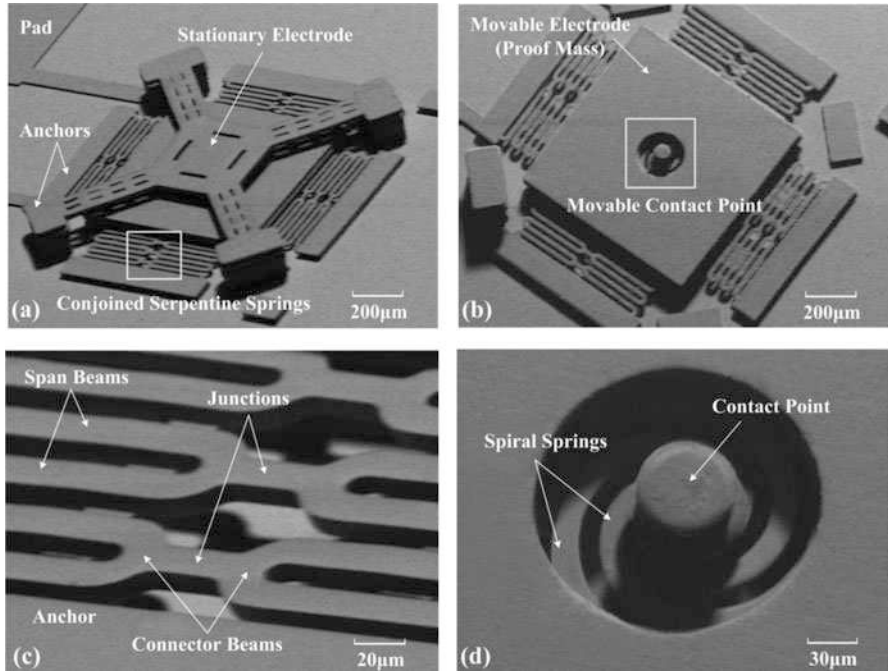


**Fig. 21** Comparison of simulation results of the inertial switch by ANSYS and Simulink

squeeze film damping was compared with ANSYS transient simulation, and the results are shown in Fig. 21, indicating that the two simulation methods were in good agreement. The only difference was that ANSYS model had taken the crossbeam's deformation into consideration. So the displacement-time curves slightly deviate from each other after the contact point rebounds. The physical parameters of the switch can be extracted by ANSYS from geometrical structure and then applied in the Simulink dynamic simulation. Based on this combination simulation, the geometrical size of the inertial microswitch can be conveniently and quickly optimized and decided for desired performance.

*Micro-fabrication and characterization:* Fabrication of the inertial microswitch with movable contact point was also carried out by low-cost and convenient multilayer electroplating technology on the insulated substrate, which allows the formation of films with thickness (hence structural height) not readily achievable using conventional surface micromachining technologies. The basic main process is like the abovementioned device fabrication as shown in Fig. 8. The SEM micrograph of the fabricated switch device is shown in Fig. 22a. A microswitch's upper stationary electrode was removed to get a clear picture of the contact point in Fig. 22b. Close-ups of the conjoined suspension serpentine springs and the contact point in the rectangles are shown in Fig. 22c, d, respectively. No severe deformation caused by residual stress was observed, and the stiction was successfully avoided during the device release.

The fabricated inertial microswitch was still tested by standard dropping hammer system like Fig. 10. When the applied acceleration was above the threshold level, the contact time of the inertial switch was prolonged by the movable contact, and it is about 56  $\mu\text{s}$  at 175 g, as shown in Fig. 23. The observed contact time was prolonged

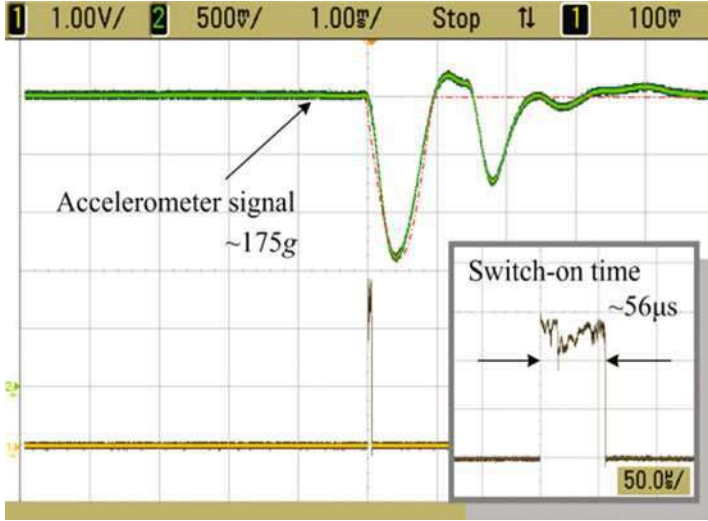


**Fig. 22** SEM micrographs of (a) the fabricated inertial microswitch and (b) with stationary electrode removed, close-ups of (c) the suspension springs and (d) the movable contact point

enough to infer that the enhanced contact effect was realized by the new designed device. A steady contact time of over  $50 \mu\text{s}$  is much longer than  $13 \mu\text{s}$  in our initial design shown in Fig. 11. Although the test contact time is still less than the improved one with cantilevers in Fig. 16, the device with movable contact point has easier fabrication process. About the detailed fabrication and simulation of the vertically driven inertial switch with movable contact point can be found in our reported works (Cai et al. 2008a, b).

## Laterally Driven Inertial Microswitch

*Design with Constraint Structures for Lower Off-Axis Sensitivity* A laterally driven inertial microswitch which can effectively lower the off-axis sensitivity and enhance the single-axial sensitivity has been designed, as shown in Fig. 24a. The device mainly consists of three parts: the double-stair shape cantilever beam as movable electrode, which attaches to the proof mass suspended by four serpentine springs, one block located at the front of proof mass as fixed electrode, and compact constraint structures consisting of constraint sleeve with holes and reverse constraint blocks. When a high enough acceleration is applied on the microswitch system along its sensing direction, the proof mass will move rapidly forward to the fixed electrode

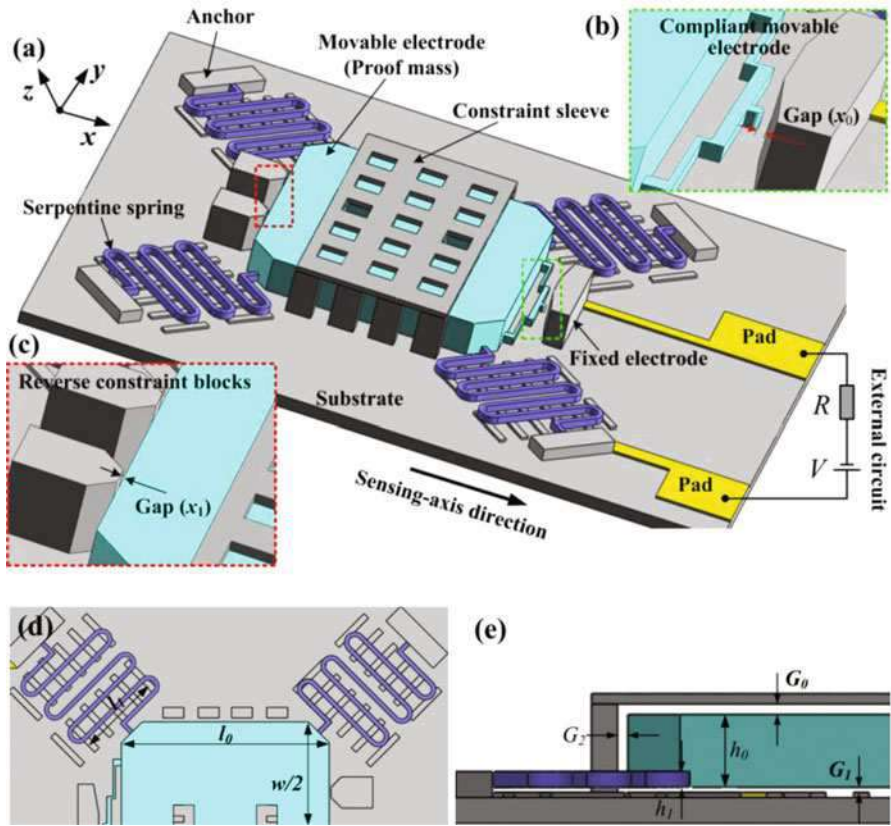


**Fig. 23** Test contact time of the fabricated inertial switch with movable contact point at 175 g

and touch it, switching the external circuit on. Subsequently, the proof mass will be drawn back quickly by serpentine springs, switching the circuit off. The double-stair shape flexible cantilever beam shown in Fig. 24b can enhance the contact between two electrodes. Introduced new constraint structures can effectively lower the off-axis sensitivity by limiting the displacement of proof mass in nonsensitive directions and, meanwhile, improve the shock-resistibility of inertial microswitch.

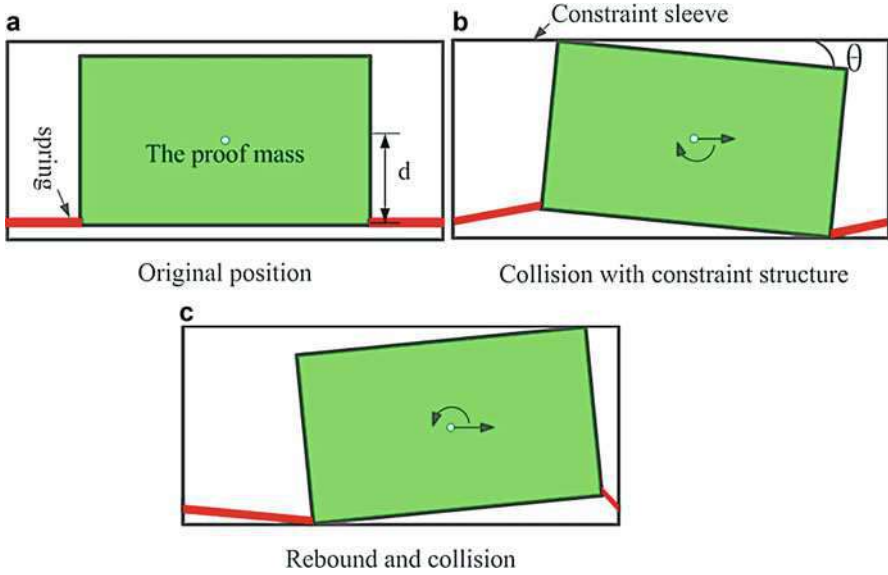
The sensitive direction of microswitch is parallel to the substrate plane, agreed with  $+x$ -direction. Comparing with the gap ( $x_0$ ) between flexible movable electrode and fixed electrode, the gap ( $x_1$ ) is smaller, as shown in Fig. 24c. When the overload acceleration is applied to the microswitch in the reverse sensitive direction, the proof mass will collide firstly with reverse constraint blocks. As long as the rebound displacement exceeding initial equilibrium position is smaller than the gap ( $x_0$ ), the double-stair shape cantilever beam would not impact to fixed electrode and avoid mistakenly triggering and improving the shock-resistibility of inertial microswitch effectively. The constraint sleeve with holes above proof mass can limit the random movement of proof mass and play a significant role to protect the device from damaging in the vertical  $z$ -direction and in the horizontal  $y$ -direction when the inertial microswitch is subjected from complicated shocks. It can effectively lower the off-axis sensitivity of inertial microswitch. Figure 24d, e shows the top view and the side view of one-half structure of inertial microswitch.

*Physical Model Analysis* In the whole motion process when the inertial switch is subjected to the applied acceleration shock, the proof mass does not move only in the sensitive direction but will also rotate in the  $xoz$  plane due to the centroid deviation



**Fig. 24** Sketch of the designed laterally driven inertial switch with constraint structures. (a) 3D view. (b) Gap between compliant movable electrode and fixed electrode. (c) Gap between proof mass and reverse constraint blocks. (d) Top view and (e) side view of one-half device

between proof mass and serpentine springs. The horizontal motion and rotation process of proof mass are simplified and described in Fig. 25a, where the whole inertial microswitch can be simplified to a spring-mass system with constraint structures. (b) When the proof mass moves toward the fixed electrode in the sensitive direction, the elastic force will increase attribute to deformation of serpentine springs. Because of the centroid deviation between proof mass and serpentine springs, the caused moment will drive the proof mass to rotate in the  $xoz$  plane until it collides with the upper and down constraint structures. (c) Then under the action of collision rebound, the proof mass will also rotate in the reverse direction. Finally, the proof mass may collide with constraint structures sometimes in moving toward the fixed electrode. Here the off-axis sensitivity ( $\epsilon$ ) of inertial microswitch could be defined as the ratio value of the maximum displacement of the proof mass in the off-axis direction and in the sensitive direction when shock acceleration with



**Fig. 25** Motion and collision process of proof mass with constraint structures in inertial switch. (a) Original position. (b) Collision with constraint structure. (c) Rebound and collision

the same amplitude is applied to the inertial microswitch along the sensitive direction, which is described as follows:

$$\varepsilon = \frac{\Delta_{\text{off}}}{\Delta_{\text{sens}}} \tag{12}$$

where  $\Delta_{\text{off}}$  and  $\Delta_{\text{sens}}$  present the maximum displacements of the proof mass in off-axis and in sensitive direction, respectively.

An analytical method is proposed to evaluate the time at that collision happens, which is significant to understand the dynamic behavior of designed inertial microswitch. When an acceleration is applied to spring-mass system in the sensing direction of inertial microswitch, the equation of horizontal motion and rotation of proof mass can be described as follows, respectively:

$$\text{Horizontal motion : } m\ddot{x} + c\dot{x} + kx = ma \tag{13}$$

$$\text{Rotation : } J_z\alpha = kx \cdot d \tag{14}$$

where  $m$  is the mass of proof mass,  $c$  is the damping coefficient,  $k$  presents the whole spring constant of four serpentine springs,  $x$  is the displacement of proof mass,  $\alpha$  is angular acceleration in the rotation process, and  $d$  is the distance between centroid of proof mass and plane of serpentine springs. The acceleration  $a$  applied to the device



in practical work is usually a half-sine wave with amplitude  $a_0$  and frequency  $\omega_0$ .  $J_z$  is the rotational inertia of system and it can be expressed as:

$$J_z = \frac{1}{12}m(l_0^2 + h_0^2) \quad (15)$$

where  $l_0$  and  $h_0$  are the length and height of proof mass, respectively. Considering that the damping coefficient is so small, it can be ignored. Solving Eq. (13), the displacement of proof mass  $x$  can be obtained:

$$x = \frac{a_0}{\omega_n^2 - \omega_0^2} \left( \sin \omega_0 t - \frac{\omega_0}{\omega_n} \sin \omega_n t \right) \quad (16)$$

where  $\omega_n$  is the inherent frequency of spring-mass inertial system and it is described as  $\omega_n = \sqrt{\frac{k}{m}}$ . Supposing that the proof mass firstly impacts to constraint structures at time  $t_c$ , the approximate expression of angular acceleration can be obtained as follows:

$$\alpha = \frac{6\omega_n^2 a_0 d \left( \sin \omega_0 t_c - \frac{\omega_0}{\omega_n} \sin \omega_n t_c \right)}{(l_0^2 + h_0^2)} \quad (17)$$

As the proof mass rotates an angle  $\theta$ , the first collision response time between proof mass and constraint structures in  $z$ -direction can be expressed as:

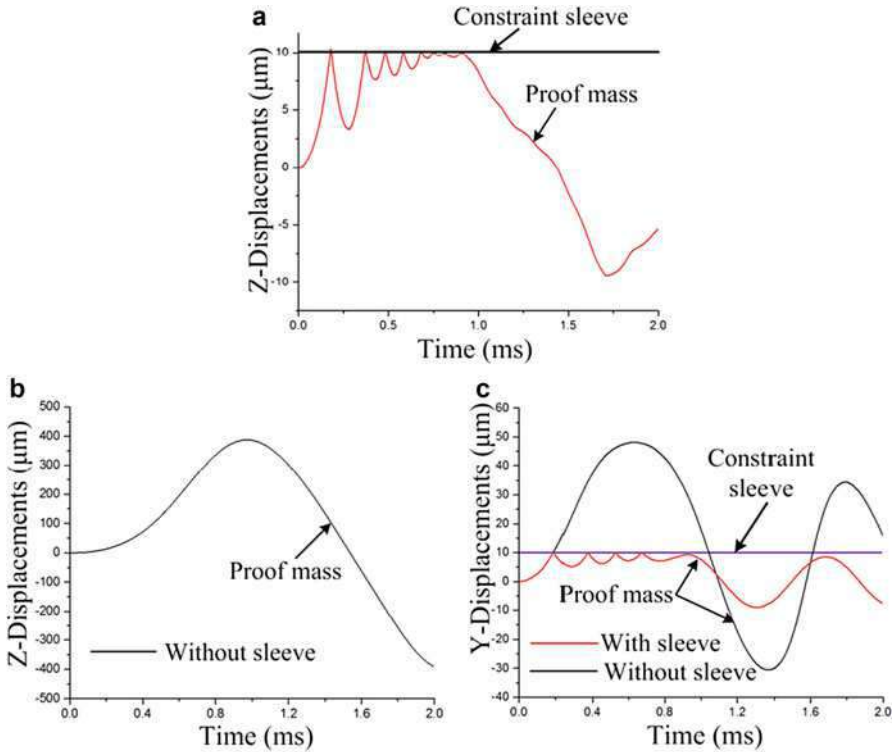
$$t_c^2 \left( \sin \omega_0 t_c - \frac{\omega_0}{\omega_n} \sin \omega_n t_c \right) = \frac{\theta(l_0^2 + h_0^2)}{3\omega_n^2 a_0 d} \quad (18)$$

where, from Fig. 25b,  $\theta$  can be expressed as follows:

$$\sin \theta = \frac{2G_0}{l_0} \quad (19)$$

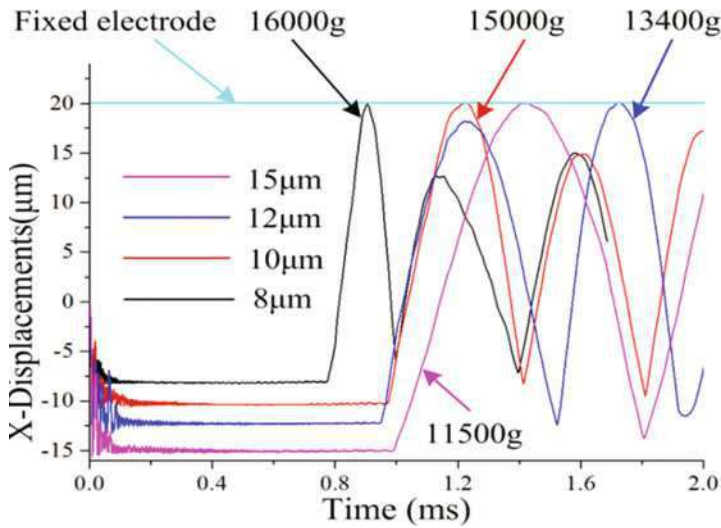
Here, the time it takes from the initial state to that the proof mass firstly impacts to constraint structures is defined as the collision response time ( $t_c$ ) in  $z$ -axis direction in the present work. The smaller the first collision response time is, the stronger the off-axis sensibility is, the lower the single-axis sensitivity is, and the proof mass would collide with constraint structures more frequently. Besides, the collision response time will be also shorter if there is a disturbance acceleration, which can drive the proof mass to move in the off-axis direction. From Eq. (18), it can be seen that the collision response time mainly depends on geometric parameters, applied shock acceleration amplitude, and inherent frequency of spring-mass inertial system. The collision response time decreases as the inherent frequency increases, and the applied acceleration amplitude increases. The collision response time increases as the length of proof mass increases.





**Fig. 26** Simulated displacement-time curves of microswitch with (a) and without (b) constraint sleeve in off-axis sensing z-direction and (c) in y-direction

*Off-Axis Sensitivity and Shock-Resistibility Simulation* In practical application, the inertial microswitch is inevitably subjected to some overload shocks from different directions because of complicated working environment. Therefore, in our novel laterally driven inertial microswitch design, constraint structures are introduced to limit overmuch displacements of proof mass in those off-axis directions and lower the off-axis sensitivity. When 170 g shock acceleration is simultaneously applied to the novel inertial microswitch in the  $x$ -,  $y$ -, and  $z$ -axis, the proof mass will move randomly and serpentine springs deform strongly. Figure 26a, b shows displacement curves of switch in off-axis sensing  $z$ -direction with and without constraint sleeve, respectively. Without constraint sleeve, the displacement of proof mass is larger than 400  $\mu\text{m}$  in  $z$ -direction. On the contrary, with constraint sleeve, the proof mass collides with constraint structures. As a result, the movement distance of proof mass is limited in the gap (10  $\mu\text{m}$ ) between proof mass and constraint sleeve. Figure 26c presents the displacement curves of switch with and without constraint sleeve in off-axis sensing  $y$ -direction. The displacement of proof mass in  $y$ -axis is 48  $\mu\text{m}$  without constraint sleeve and is

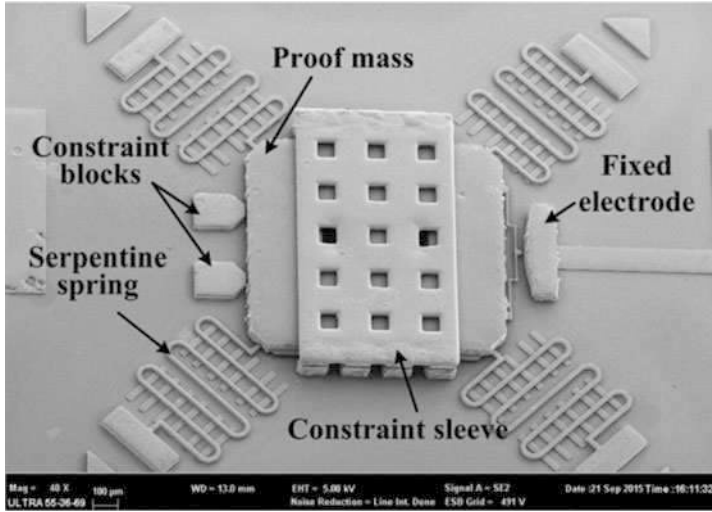


**Fig. 27** Shock-resistibility of the designed inertial microswitch with different gaps ( $x_1$ ) between proof mass and reverse constraint blocks

10  $\mu\text{m}$  with constraint sleeve. In our improved design, a laterally driven inertial switch with double-layer serpentine spring had been also proposed for realizing better single-axial sensitivity (Zhang et al. 2016a).

Generally, when the proof mass is subjected to overload acceleration in the reverse sensitive direction, the large deformation of suspension springs would cause proof mass to rebound along the sensitive direction. If the rebounded distance of flexible movable electrode is equal to or larger than the gap between fixed and movable electrode  $x_0$  (20  $\mu\text{m}$ ), it would lead to a spurious triggering in an inertial microswitch. The tolerant maximum overload acceleration in the reverse sensitive direction shows the shock-resistibility of an inertial microswitch. Good shock-resistibility indicates that the spurious triggering is not easy to occur in an inertial microswitch.

Figure 27 presents displacement-time curves of proof mass with different gaps ( $x_1$ ) between proof mass and reverse constraint blocks when reverse overload accelerations are applied to the inertial microswitch. When the gap  $x_1$  is 8, 10, 12, and 15  $\mu\text{m}$ , simulation results that show the allowed maximum overload acceleration amplitudes which did not make the microswitch be spuriously triggered are about 16,000, 15,000, 13,500, and 11,500 g, respectively. Their maximum rebounded displacements of movable electrode are close to the gap  $x_0$ , and the microswitch is just not spuriously triggered in these overload acceleration amplitudes. If the applied reverse acceleration amplitude is larger than these values, the microswitch will make the external circuit switch on. It also can be found that in the current research scale, the designed inertial microswitch can suffer a maximum of

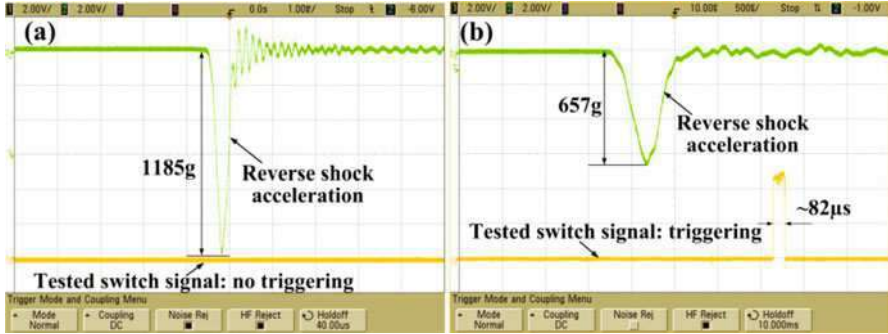


**Fig. 28** SEM of the fabricated laterally driven inertial microswitch with compact constraint structures

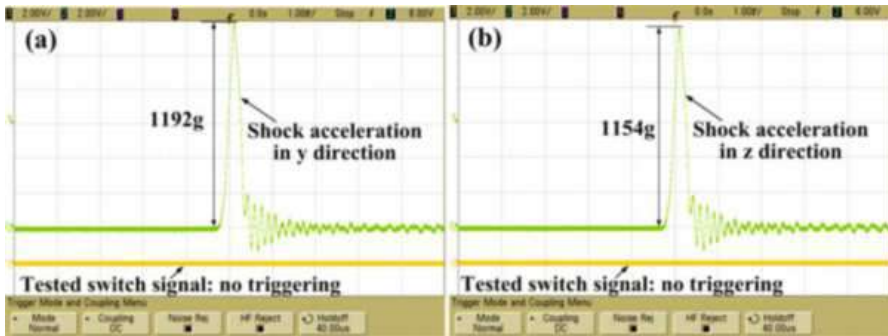
16,000 g reverse overload shock and keep no spurious triggering which occurs when the gap is set to 8  $\mu\text{m}$ .

*Micro-Fabrication and Verification* The designed inertial microswitch was still fabricated by low-temperature photoresist mold metal electroplating technology abovementioned. The nickel (Ni) metal was selected as the structure material. The detailed fabrication process steps of the device can be found in our reported work (Zhang et al. 2016b). Figure 28 shows the SEM of the fabricated laterally driven inertial microswitch with multidirectional compact constraint structures. Figure 29a shows that when the applied overload acceleration amplitude in the reverse sensitive direction ( $-x$ ) is 1185 g, the rebounded distance of movable electrode in the inertial microswitch with constraint sleeve in the sensitive direction is lower than the gap between two electrodes. So there is no spurious triggering in the fabricated inertial microswitch. However, Fig. 29b indicates that when the applied overload acceleration amplitude in the reverse sensitive direction ( $-x$ ) is 657 g, the microswitch without constraint sleeve is triggered, and the external circuit will be spuriously switched on. Therefore, the introduced constraint sleeve can effectively limit the random movement of proof mass and protect the inertial microswitch from spurious triggering.

In order to evaluate the off-axis sensitivity of the fabricated inertial microswitch in  $z$ - and  $y$ -axis, higher shock accelerations are applied to the microswitch along these two axial directions. The test results are shown in Fig. 30. It indicates that there is no spurious triggering in the inertial microswitch when the applied shock overload acceleration amplitude in  $y$ -axis direction is 1192 g. Similarly, the inertial



**Fig. 29** Test results of microswitch with (a) and without (b) constraint sleeve in the reverse sensing-axis direction: Spurious triggering is easier to happen if without constraint sleeve



**Fig. 30** Test results of microswitch when high shock accelerations are applied in (a) y-axis direction and (b) z-axis direction: No spurious triggering happens

microswitch is also not spuriously triggered when the applied shock overload acceleration amplitude in z-axis direction is 1154 g. Therefore, utilizing the introduced constraint structures, the fabricated microswitch can subject larger shock acceleration in off-axis sensing direction and keep a good shock stability. Compared with the current researches (Stauffer 2006; Tao et al. 2014; Delahunty and Pike 2014; Chen et al. 2014a, b; Wang et al. 2013a), the proposed device utilizes multidirectional compact constraint structures to realize lower off-axis sensitivity and improve shock-resistibility.

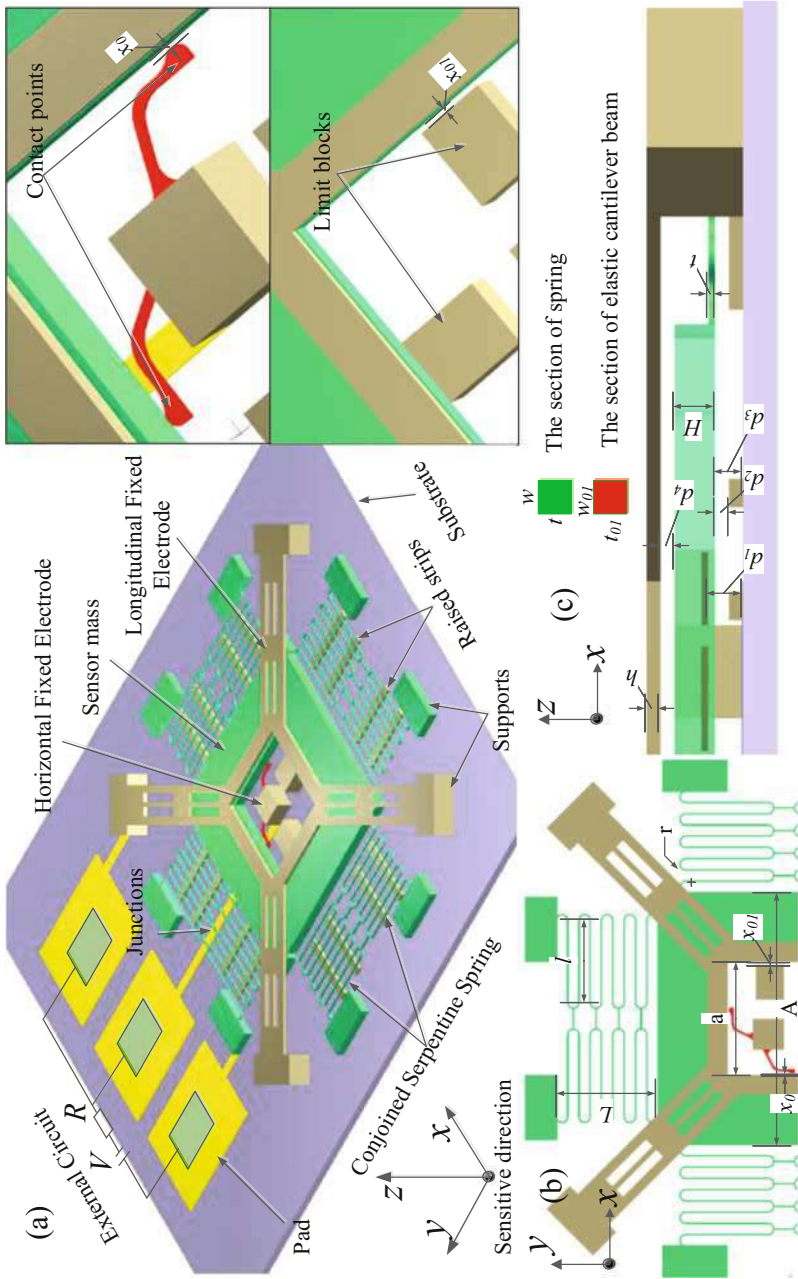
In addition, about the laterally driven inertial microswitch, we have studied the influence of the applied accelerations on the contact time and the threshold level in detail (Chen et al. 2014a) and proposed some optimized designs and dynamic simulation analysis for lowering the axial disturbances and prolonging the contact time (Chen et al. 2014c, 2015, 2016; Wang et al. 2013b; Xu et al. 2016).

## Multidirectional Sensitive Inertial Microswitch

### Triaxial Inertial Microswitch (Monolithic Integration with CMOS)

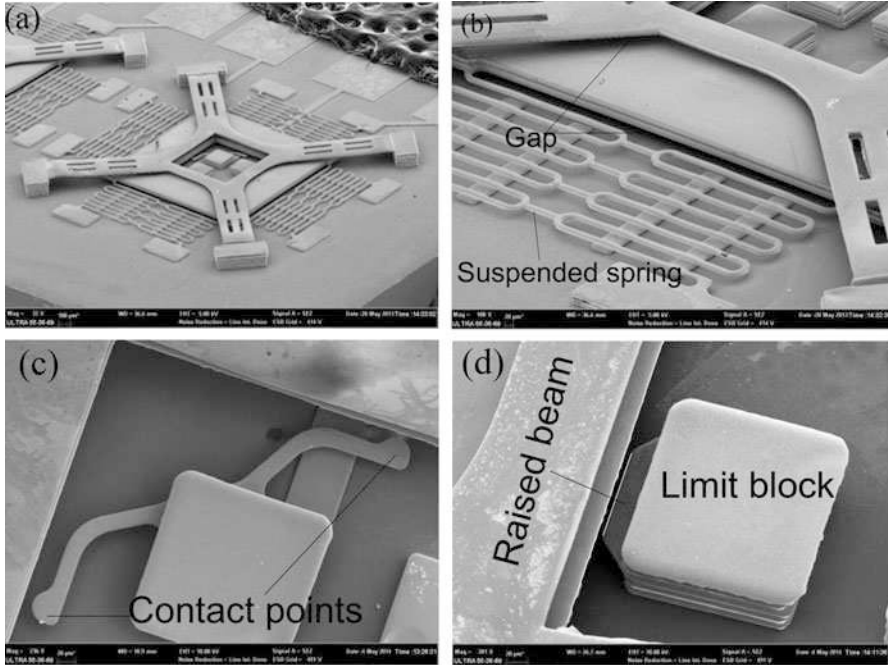
*Device design:* A triaxial inertial microswitch that can detect threshold acceleration in three directions ( $+x$ ,  $+y$ , and  $+z$ ) has been proposed and designed. Compared with the switch sensitive to symmetrical multiple directions ( $\pm x$ ,  $\pm y$ ,  $\pm z$ ), the difficulty in structure design is axial disturbance in negative directions ( $-x$ ,  $-y$ , and  $-z$ ). A novel structure that mainly includes four parts was drawn in Fig. 31a: the insulated glass substrate and the anti-stiction raised bar supports, a hollow proof mass suspended by conjoined serpentine springs as the movable electrode, two  $L$ -shaped elastic cantilever beams suspended by fixed block in the middle of proof mass as horizontal stationary electrodes, a multi-hole crossbeam above the proof mass as vertical fixed electrode, and two anchors located in the middle of proof mass as limit blocks. The movable electrode suspended by conjoined serpentine spring was designed to be a centrosymmetric structure in order to reduce the cross-axis sensitivity, increasing the axis sensitivity. The horizontal electrodes fabricated with  $L$ -shaped cantilever beam can prolong the contact time and release the inertial impaction during the contact process, which prevents the damage from overload acceleration, increasing the safe reliability of system and prolonging service life simultaneously.

In a real-world application, the switch was shocked in all direction, but many shock directions must be constrained. Two fixed blocks have been set in the inside of proof mass to restrain the contact bounce from trigger position and restrain the overload impaction from the opposite direction. The spurious triggering in reflection sensitive direction ( $-x$ ,  $-y$ ) would be avoided due to the tiny gap between proof mass and limit blocks; the gap was set as  $10\ \mu\text{m}$  in this structure based on surface micromachining technology. The details of elastic cantilever beams and fixed blocks were shown in Fig. 31a; the gap ( $x_{01}$ ) between blocks and proof mass is designed to be much narrower than the gap ( $x_0$ ) between contact point and proof mass in favor of reducing the bounce displacement. Figure 31b, c shows the top view and side view of one-half of triaxial switch, the symmetric structure of the whole device that ensures threshold accelerations in horizontal direction is consistent, and the proof mass thickness  $H$  is designed to be much larger than the spring thickness  $t$  in order to minimize the area coverage while enabling the desired sensitivity. When the acceleration (i.e., meets or surpasses set threshold) was applied on inertial switch, the proof mass moves toward the stationary electrodes in horizontal directions or vertical direction, and the  $L$ -shaped elastic electrodes and multi-hole crossbeam can realize a flexible contact to prolong the contact time and maintain system stability simultaneously. Figure 31c shows the height  $d_2$  of horizontal elastic contacts higher than the bottom of proof mass that avails to improve the effectiveness and reliability during contact process. The gap ( $d_4$ ) was designed higher enough to eliminate the cross talk between horizontal and vertical stationary electrode. Three pads were connected to the external circuit, forming a closed circuit with the switch when the system is in the conduction state.



**Fig. 31** The schematic diagram and structure parameters of triaxial inertial switch. (a) 3D sketch, (b) top view, and (c) side view of one-half of the device

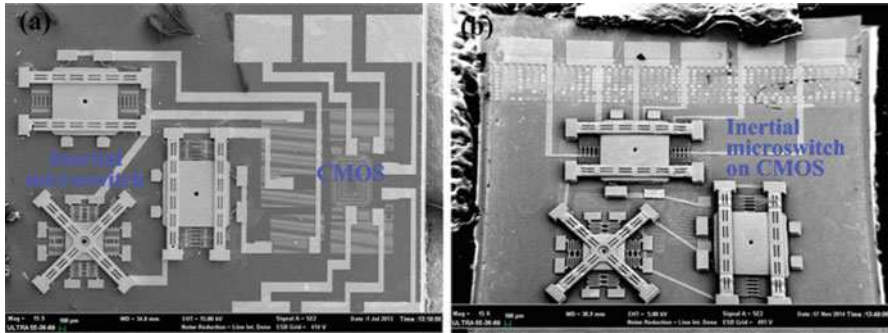




**Fig. 32** SEMs of the fabricated triaxial inertial switch. (a) The full view, (b) side view of the suspended structure, (c) horizontal fixed electrode, (d) the limit block

*Micro-fabrication and characterization:* Similarly, the designed triaxial inertial microswitch is also fabricated by low-temperature photoresist modeled metal electroplating technology abovementioned. The structure material of the device is still metal nickel (Ni). The fabrication process mainly includes sputtering, electroplating, and sacrificial layer technologies. Figure 32 shows the SEM photos of fabricated triaxial inertial microswitch. Figure 32a shows the triaxial inertial switch with complete structure. The enlargements of suspended gaps were shown in Fig. 32b; it is clear to see the gap between movable electrode and substrate and vertical fixed electrode. Figure 32c shows the elastic contact points, and Fig. 32d is limit block. The novel design of raised beam is used to simplify the technology process and which can raise the precision of manufacture. The detailed fabrication process steps and optimization simulation of the device can be found in our reported works (Chen et al. 2014d; Chen et al. 2014e).

In addition, in order to improve the integration level of the system, the monolithic integrated triaxial inertial microswitch with CMOS circuit has been proposed and fabricated on silicon wafer by abovementioned photoresist modeled metal electroplating technology. The signal extracting and preprocessing circuit is included into the CMOS, which is connected with the triaxial inertial microswitch via pads. Figure 33 shows the SEMs of the fabricated monolithic integrated triaxial inertial switch with CMOS circuit. Figure 33a illustrates the inertial microswitch located at



**Fig. 33** SEMs of the fabricated monolithic integrated triaxial inertial switch with CMOS circuit. (a) Inertial microswitch located at the side of CMOS, (b) inertial microswitch located on CMOS

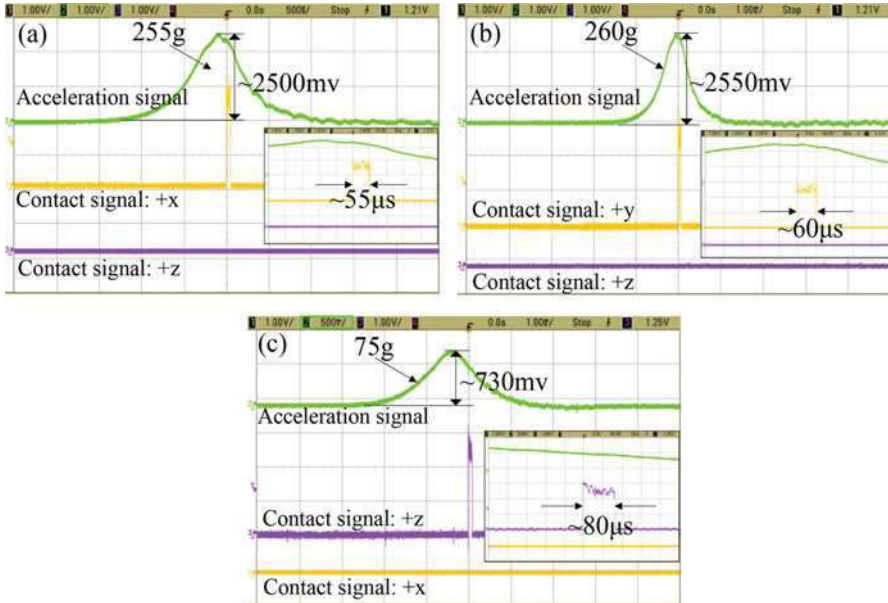
the side of CMOS, and Fig. 33b shows the inertial microswitch directly located on CMOS. During the fabrication, the effect of stress caused by multilayer electroplating on the CMOS circuit should be considered adequately.

Figure 34 shows the drop hammer experiment results of inertial switch in three sensitive directions. The acceleration and switch signals captured by oscilloscope indicate that the threshold accelerations in horizontal directions are about 255 g (Fig. 34a) and 260 g (Fig. 34b), and the contact time is about 55–60  $\mu\text{s}$ , respectively. There is minor difference of threshold acceleration between x- and y-directions due to the fabrication process are repeating steps, as a result, these cumulative errors lead to the change of gap between movable electrode and fixed electrode. Obviously, the contact time is prolonged to 60  $\mu\text{s}$  that is attributed to elastic deformation of L-shaped fixed electrodes. Figure 34c shows the threshold acceleration of triaxial inertial switch in +z-direction which is about 75 g, and contact time is about 80  $\mu\text{s}$ . The tested results shown in Fig. 34 demonstrate that the cross talk between horizontal and vertical directions is restrained. The contact in +z-direction is avoided when the prototype is shocked by acceleration in +x- and +y-directions, as shown in Fig. 34a, b. Likewise, the trigger does not appear in horizontal direction when the acceleration is applied in +z-direction, as shown in Fig. 34c.

### Omnidirectional Sensitive Inertial Microswitch

*Device Design* The omnidirectional sensitive inertial microswitch was proposed and designed in terms of non-silicon surface micromachining technology on a glass wafer, as shown in Fig. 35. It consists of two main parts: a suspended proof mass with snake springs as sensing movable electrode and a T-shaped structure on the substrate with maple leaflike top and cantilevers around the central cylinder as vertical and lateral fixed electrodes. The springs are suspended by anchors fixed on the substrate. The vertical fixed electrode contains two kinds of different materials with the same thickness. The top layer is a polymer, while the bottom one is a metal nickel. As a result, the constructed composite can realize flexible deformation and



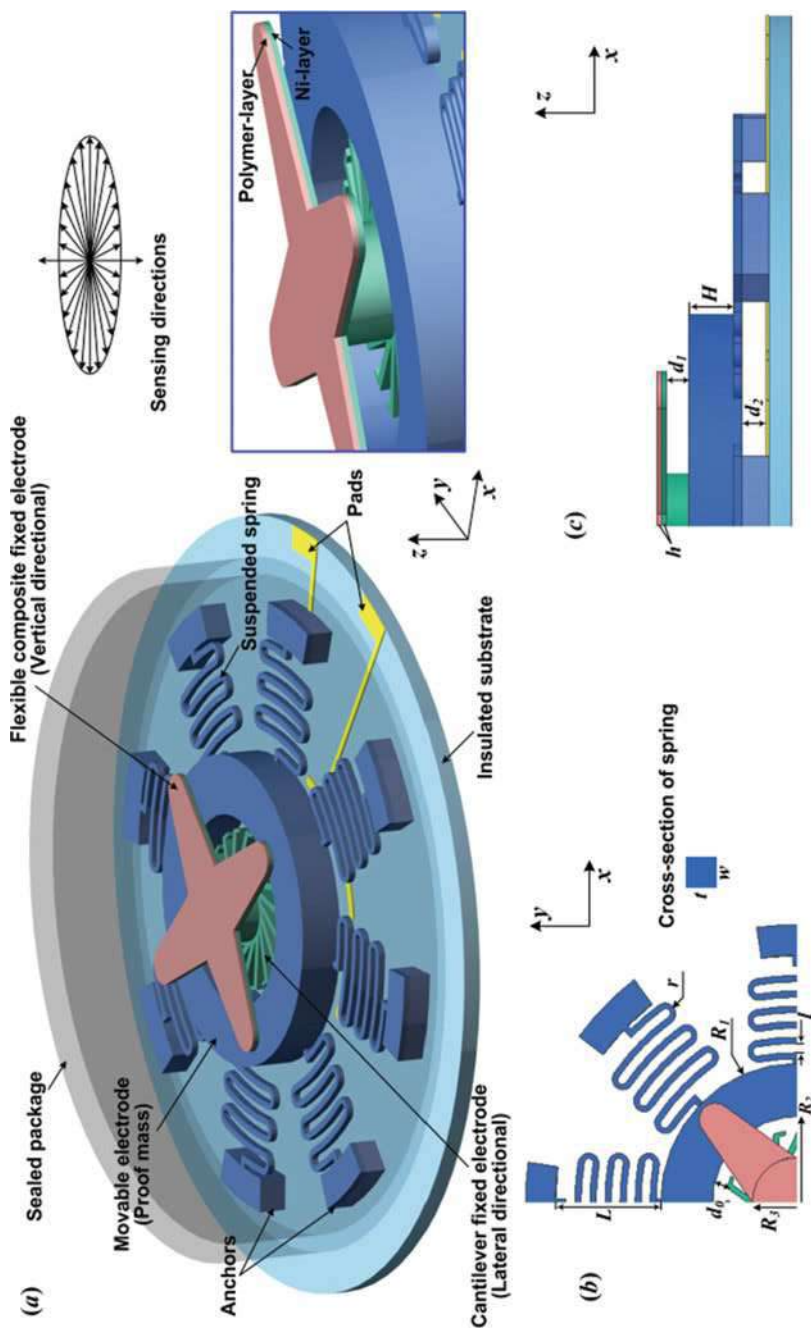


**Fig. 34** Test results of threshold accelerations and contact times in x (a)-, y (b)-, and z (c)-direction in the triaxial inertial microswitch

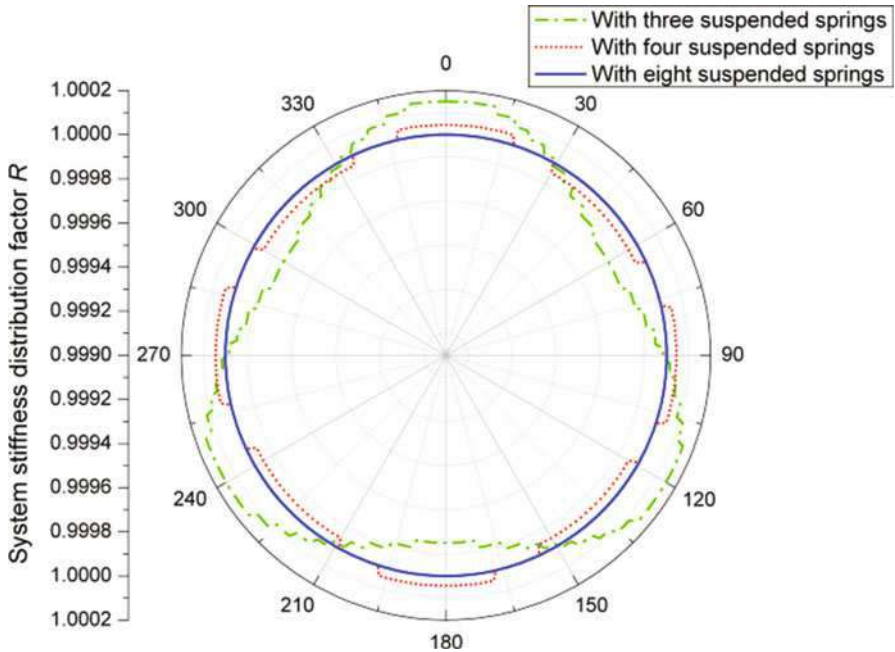
prolong the contact time under applied load. Besides, the designed microswitch can sense the applied accelerations from any radial directions in  $xoy$  plane and  $z$ -axis. Figure 35a shows the 3D sketch of the designed omnidirectional sensitive inertial microswitch, and Fig. 35b, c shows the top and side views of a quarter of the microswitch, respectively.

Based on the FE simulation, Fig. 36 shows a simulated stiffness distribution of the whole spring-mass system of the omnidirectional sensitive microswitch in  $xoy$  plane, where  $R$  is a normalization factor. It is indicated that a very uniform stiffness distribution could be obtained as long as only eight springs were used to suspend the proof mass. Of course, the stiffness will be more uniformed with increasing the number of the springs. However, too many springs would lead to excessive structural stiffness and complicated fabrication process.

*Micro-Fabrication and Characterization* Similarly, the omnidirectional sensitive inertial microswitch was fabricated on insulated glass substrate by low-cost surface multilayer metal electroplating technology. Electroplated metal nickel was still used as the main structure material. The vertical composite fixed electrode was fabricated by electroplating and electrophoresis. In our fabrication, a polyurethane-modified epoxy resin (EP-polymer, Shanghai HLS Paint Co. Ltd.) will be used as the electrophoretic polymer on the nickel surface. The ratio of electrophoretic emulsion to deionized water is 0.5. The electrophoresis conditions are pH 6.8, the temperature 30 °C, the applied voltage 50 V, and



**Fig. 35** Schematic of the proposed omnidirectional sensitive inertial microswitch. (a) 3D sketch, (b) top view, and (c) side view of a quarter of the microswitch

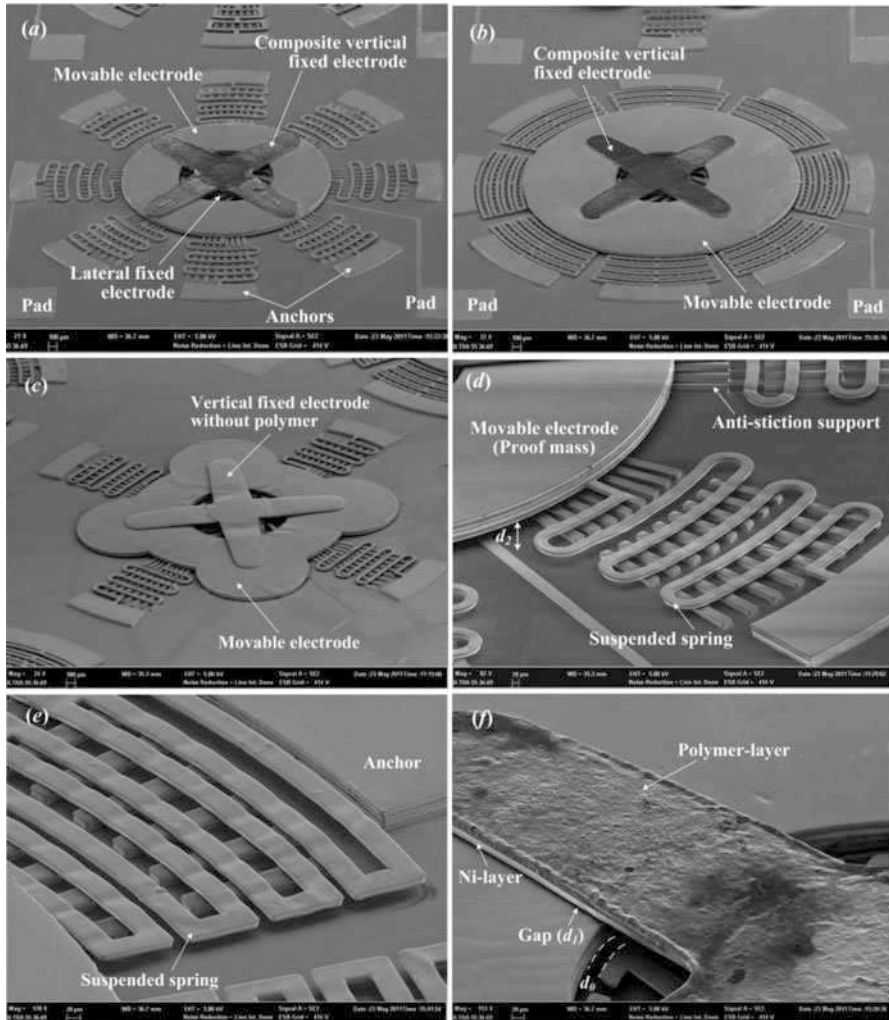


**Fig. 36** Simulated spring-mass system stiffness distributions of the omnidirectional sensitive inertial microswitch in  $xoy$  plane

duration time 30 s. The detailed fabrication process steps and optimization simulation of the device can be found in our reported works (Zhu et al. 2012; Yang et al. 2012a,b).

The SEMs of the released omnidirectional sensitive inertial microswitch prototypes are shown in Fig. 37. The microswitches with different shapes are shown in Fig. 37a–c, respectively. The close-up of Fig. 37d shows the movable proof mass has been suspended by the serpentine springs conjoined on the anchors. Figure 37e is a rectangular suspended spring used in the microswitch shown in Fig. 37b. Figure 37f shows the fabricated vertical composite fixed electrode which consists of two layers. Although the surface of the polymer layer looks not very smooth due to slight shrink during the solidification, the polymer layer and the Ni layer still have a good adhesion.

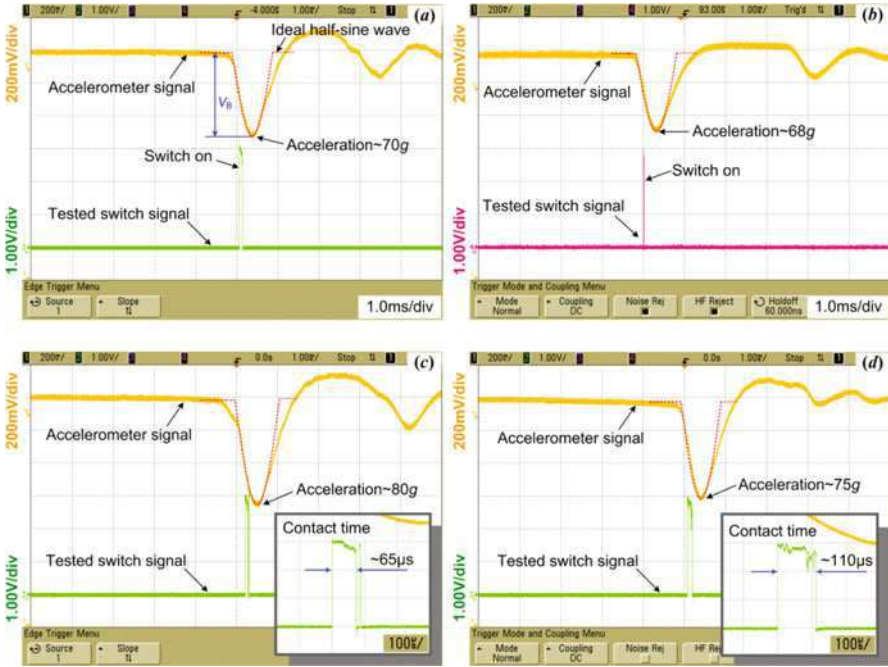
The inertial microswitch prototypes were tested by a drop hammer system. Figure 38 shows the drop hammer experiment results. Figure 38a shows a typical test threshold acceleration of the microswitch in vertical sensitive direction. Figure 38b shows a typical test threshold acceleration of the microswitch in 225°-sensitive direction in  $xoy$  plane, which is about 68 g. In addition, we tested the threshold accelerations of the switch in other sensitive directions in  $xoy$  plane, such as 72 g in 0°, 73 g in 45°, 69 g in 90°, 70 g in 135°, 68 g in 180°, 71 g in 270°, and



**Fig. 37** SEMs of (a), (b), and (c) the fabricated omnidirectional sensitive inertial microswitches with different shapes (d) and (e) released suspended serpentine springs and proof mass; (f) close-up of the vertical fixed electrode with polymer-Ni composite

67 g in 315°. It's indicated that threshold accelerations of the microswitch in lateral directions are generally uniform in  $xoy$  plane, is also 70 g, which meets our predesign.

Figure 38c, d shows the test contact times of the inertial microswitches without and with electrophoretic EP-polymer in their vertical fixed electrodes, respectively. In order to demonstrate the effect obviously, over-threshold acceleration, 80 and 75 g were applied to the tested microswitch, respectively. It's shown that the contact time of microswitch with composite fixed electrode is  $\sim 110 \mu s$ , which is longer than that



**Fig. 38** Typical test result of threshold accelerations of the omnidirectional sensitive inertial switch (a) in vertical direction and (b) in lateral 225° direction in *xoy* plane; Contact times of the switch (c) without and with (d) EP-polymer composite fixed electrode

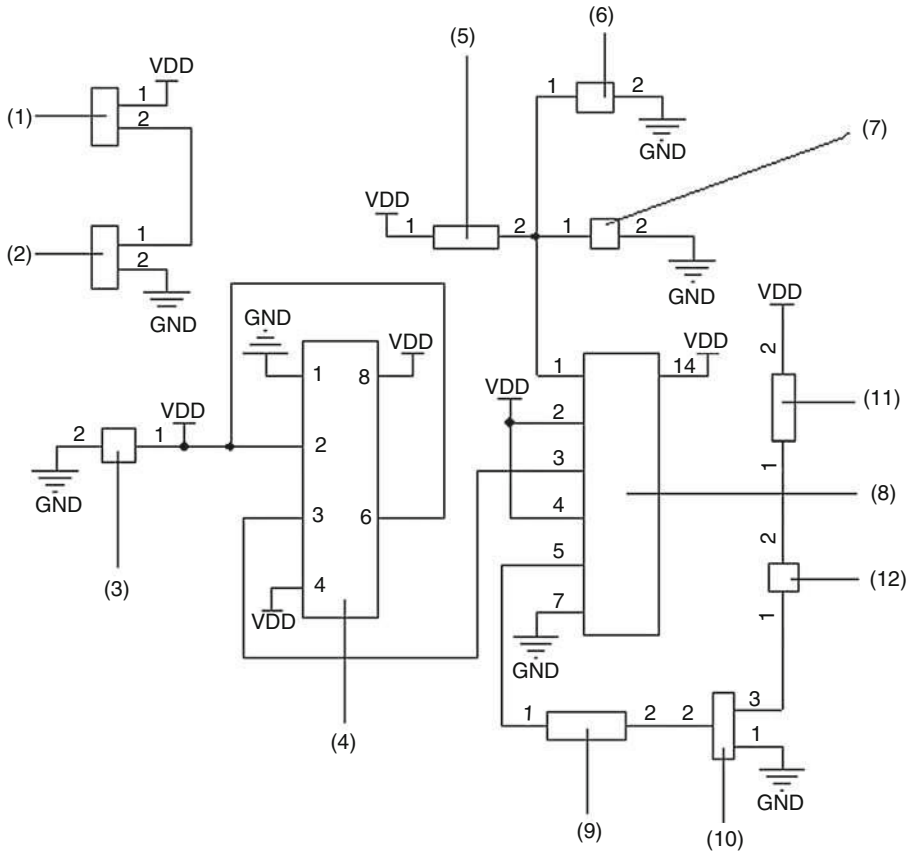
(~65 µs) of one without polymer although higher acceleration has been used to apply to the switch with all-metal fixed electrode. In addition, an undesirable bounce phenomenon even occurred in the switch without polymer, as shown in the close-up of Fig. 38c. On the contrary, the microswitch with flexible composite is very beneficial for improving contact effect and decreasing bounces.

In addition, some optimized designs and improvements about the omnidirectional sensitive inertial microswitch have been carried out in our related recent reports for prolonging the contact time and better performances (Yang et al. 2010b; Chen et al. 2014f; Wang et al. 2015a, b).

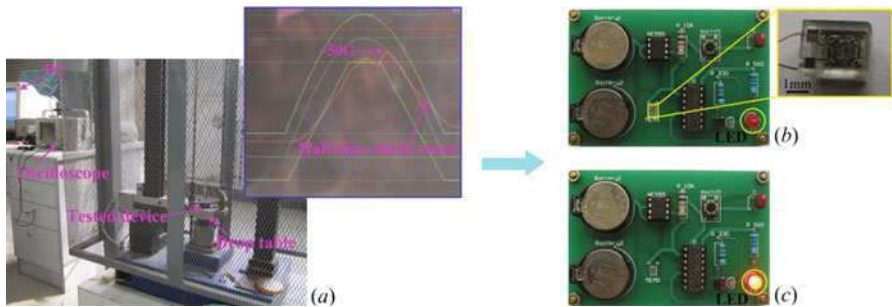
### Application Example Demonstration

As an application example, the fabricated inertial microswitch was connected with a signal extracting and processing circuit mainly consisting of a data flip-flop chip, as shown in Fig. 39. Finally, the integrated PCB with packaged MEMS inertial microswitch for vibration monitoring systems was characterized by shock equipment, as shown in Fig. 40a. Agilent-6000 oscilloscope (MSO6034A, USA) was used to collect the output signals of the impact and the tested sensor. The PC was used to





**Fig. 39** Signal extracting and processing circuit with data flip-flop chip for MEMS inertial microswitch



**Fig. 40** Test of the PCB with packaged MEMS inertial microswitch for vibration monitoring systems. (a) Testing setup; (b) before and (c) after undergoing threshold acceleration

show the applied half-sine shock wave from drop hammer system. Both the standard accelerometer and the tested PCB were fixed on the steel dropping pole, with their sensitive directions perpendicular to the ground. When the pole is dropped freely from a certain height onto the base plate, a half-sine wave shock with about 1 ms impulse width is generated and applied to the tested sensor.

Figure 40b, c shows the tested PCB with packaged MEMS inertial microswitch before and after undergoing threshold acceleration, respectively. It can be seen that the LED on the PCB just lights after a 50 g acceleration shock was applied to the sensor. It indicates that the mounted inertial microswitch device can work normally when used in the vibration monitoring application. The systematic design and simulation analysis of the device can be found in our related work (Yang et al. 2010).

---

## Conclusions

The MEMS inertial switch as a kind of passive device is essentially both a sensor and an actuator, which has functions of both threshold detection and closing action. It is widely used in industrial transportation, airbag restraint system (ABRS), and remote monitoring due to the distinctive characteristics of small size, lightweight, and ultralow power consumption and manufacturing cost. It has been paid attention by more and more researchers in MEMS field. Especially in recent years, the IoT has grown rapidly, the inertial switch is used for detecting the vibration shock and then sending corresponding overload signal in remote areas where the power supply is limited, and the power draw significantly reduces the monitoring period before the power would be changed or recharged. In this chapter, after a category and research progress summary of the MEMS inertial switch some of our developed inertial switches with different sensitive directions and configurations have been proposed and fabricated by non-silicon surface micromachining technologies mainly consisting of metal electroplating and sacrificial layer processes. They were designed and improved for realizing not only longer contact time but also quicker dynamic response. From single-axis vertically driven or laterally driven to triaxial sensitive or omnidirectional sensitive inertial microswitches, many novel designs are proposed originally and then optimized and improved step by step for better performance. There are several basic conclusions obtained as follows:

1. In the vertically driven inertial microswitch, the device with bridge-type compliant fixed electrode can prolong the contact time to some extent compared with conventional microswitches using only a rigid structure as the fixed electrode. A dynamic contact simulation well explains the bounce phenomenon in the inertial microswitch, which helps the improved switch with flexible cantilever array between two impact electrodes reach a longer contact time and eliminate the bounces. As a novel design, the inertial microswitch with movable contact point was proposed for the first time, which skillfully utilized a dual spring-mass

system configuration to enhance the contact between the movable electrode and the fixed electrode successfully.

2. In the laterally driven inertial microswitch, multidirectional compact constraint structures were introduced to lower the off-axis sensitivity and improve the shock-resistibility. The constraint sleeve and reverse stop-block structures were used to limit the overmuch displacement of proof mass in the microswitch and avoid the device damage under high shock load. The corresponding simulation and experiment have proved the proposed novel design reasonable.
3. The triaxial inertial microswitch has been proposed by using only one proof mass to sense the applied acceleration shocks from three inter-perpendicular orientations, which was fabricated by low-temperature photoresist modeled metal electroplating technology. A monolithic integrated triaxial inertial microswitch with CMOS signal extracting and preprocessing circuit has been designed and also fabricated on silicon wafer by metal electroplating technology, which can improve the integration level of the system effectively.
4. The dynamic response properties of the inertial microswitch, such as contact time, threshold level, and response time, are affected by not only structural parameters of itself but also characteristics (e.g., shock wave shape, pulse width, amplitude, etc.) of the applied shock loads.
5. The developed inertial microswitch device has been mounted on the PCB with triggering circuit, which is successfully used in a vibration monitoring module. The application demo indicates the passive inertial microswitch with ultralow power consumption is promising in various sensor network systems.

In our present research, the planarization treatment among the multilayer composite structures is very critical during the photoresist modeled metal electroplating micromachining. In addition, the low stress control is also important in the fabrication process. Some variations in the microswitch closure thresholds occur from wafer to wafer, which is primarily due to residual stress or structure dimension variations during the fabrication. The further work will focus on the optimizations of the process sequence steps and the electroplating solution compositions for obtaining better device performances. How to realize the low-cost batch manufacturing is noteworthy, which will really promote the inertial microswitch to be practically used in various fields, especially fast-growing sensor network applications. To further optimize the design and improve the system integration level of the passive inertial microswitch with CMOS signal processing circuits is also a key factor for the universal application in the future.

**Acknowledgments** The authors would like to thank the supports from the Hi-Tech Research and Development Program of China (2015AA042701), National Natural Science Foundation of China (No. 51205254, No. 61571287), Advanced Research Ministry of Education Joint Foundation (6141A02033403), SJTU SMC Young Scholar Program (16X100080044), and the National Key Laboratory of Micro/Nano Fabrication Technology Foundation (9140C790403150C79332). The authors would also like to thank Prof. H. Wang, Dr. H. Cai, Dr. W. Chen, Dr. Q. Xu, Mr. Q. Zhang, Mr. B. Zhu, Ms. H. Shen, and Ms. Y. Wang for their important works about the improvements of the device prototypes.



## References

- Cai H, Ding G, Yang Z, Su Z, Zhou J, Wang H (2008a) Design, simulation and fabrication of a novel contact-enhanced MEMS inertial switch with a movable contact point. *J Micromech Microeng* 18:115033. (10pp)
- Cai H, Yang Z, Ding G, Zhou Z (2008b) Design and simulation of a contact-enhanced MEMS inertial switch using Simulink and ANSYS. In: 1st international conference of the Chinese society of micro/nano technology, Beijing, vol 2008
- Cai H, Yang Z, Ding G, Wang H (2009) Development of a novel MEMS inertial switch with a compliant stationary electrode. *IEEE Sensors J* 9:801–808
- Chen W, Yang Z, Wang Y, Ding G, Wang H, Zhao X (2014a) Influence of applied acceleration loads on contact time and threshold in an inertial microswitch with flexible contact-enhanced structure. *Sensors Actuators A* 216:7–18
- Chen W, Wang Y, Zhang Y, Cheng P, Wang Y, Ding G, Zhao X, Yang Z (2014b) Fabrication of a novel contact-enhanced horizontal sensitive inertial micro-switch with electroplating nickel. *Microelectron Eng* 127:21–27
- Chen W, Wang Y, Wang Y, Ding G, Wang H, Zhao X, Yang Z (2014c) A laterally-driven micromachined inertial switch with compliant cantilever beam as the stationary electrode for prolonging contact time. *J Micromech Microeng* 24:065020. (10pp)
- Chen W, Wang Z, Wang Y, Yang Z, Ding G, Wang H, Zhao X (2014d) Tri-axial all-metal inertial microswitch with compliant stationary electrodes to prolong contact time. *IET Micro Nano Lett* 9:514–518
- Chen W, Yang Z, Wang Y, Ding G, Wang H, Zhao X (2014e) Simulation, fabrication and characterization of an all-metal contact-enhanced triaxial inertial microswitch with low axial disturbance. *Sensors and Actuators A* 220:194–203
- Chen W, Ding G, Wang Y, Wang H, Zhao X, Yang C, Yang Z (2014f) An all-metal passive threshold sensor for omni-directional vibration monitoring application. In: *Proceedings of IEEE sensors 2014*, Valencia, 2–5 Nov, vol 2014, pp 428–431
- Chen W, Wang Y, Wang Y, Yang Z, Ding G, Wang H, Zhao X (2015) Simulation and characterization of a laterally-driven inertial micro-switch. *AIP Adv* 041319:1–5
- Chen W, Yang Z, Wang Y, Ding G, Wang H, Zhao X (2016) Fabrication and characterization of a low-g inertial microswitch with flexible contact point and limit-block constraints. *IEEE/ASME Trans Mechatron* 21:963–972
- Choi J, Lee JI, Eun Y, Kim M, Kim J (2011) Aligned carbon nanotube arrays for degradation-resistant intimate contact in micromechanical devices. *Adv Mater* 23:2231–2236
- Currano LJ, Bauman S, Churaman W, Peckerar M, Wienke J, Kim S, Yu M, Balachandran B (2008) Latching ultra-low power MEMS shock sensors for acceleration monitoring. *Sensors and Actuators A* 147:490–497
- Currano LJ, Yu M, Balachandran B (2010) Latching in a MEMS shock sensor: modeling and experiments. *Sensors Actuators A* 159:41–50
- Currano LJ, Becker CR, Lunking D, Smith GL, Isaacson B, Thomas L (2013) Triaxial inertial switch with multiple thresholds and resistive ladder readout. *Sensors and Actuators A* 195:191–197
- Delahunty A, Pike W (2014) Metal-armoring for shock protection of MEMS. *Sensors Actuators A* 215:36–43
- Frobenius W, Zeitman S, White M, O'Sullivan D, Hamel RG (1972) Microminiature ganged threshold accelerometers compatible with integrated circuit technology. *IEEE Trans Electron Devices* 19:37–40
- Go JS, Cho YH, Kwak BM et al (1996) Snapping microswitches with adjustable acceleration threshold. *Sensors Actuators A* 54:579–583
- Jia M, Li X, Song Z, Bao M, Wang Y, Yang H (2007) Micro-cantilever shocking-acceleration switches with threshold adjusting and 'on'-state latching functions. *J Micromech Microeng* 17:567–575

- Jones EE, Begley MR, Murphy KD (2003) Adhesion of micro-cantilevers subjected to mechanical point loading: modeling and experiments. *J Mech Phys Solids* 51:1601–1622
- Lee JI, Song Y, Jung H, Choi J, Eun Y, Kim J (2011) Carbon nanotubes-integrated inertial switch for reliable detection of threshold acceleration. In: *Proceedings of Transducers'11, Beijing*, vol 2011, pp 711–714
- Lee JI, Song Y, Jung H, Choi J, Eun Y, Kim J (2012) Deformable carbon nanotube-contact pads for inertial microswitch to extend contact time. *IEEE Trans Ind Electron* 59:4914–4920
- Loke Y, McKinnon GH, Brett MJ (1991) Fabrication and characterization of silicon micromachined threshold accelerometers. *Sensors Actuators A* 29:235–240
- Ma W, Zohar Y, Wong M (2003) Design and characterization of inertia-activated electrical micro-switches fabricated and packaged using low-temperature photoresist molded metal-electroplating technology. *J Micromech Microeng* 13:892–899
- Ma W, Li G, Zohar Y, Wang M (2004) Fabrication and packaging of inertia micro-switch using low-temperature photo-resist molded metal-electroplating technology. *Sensors Actuators A* 111:63–70
- Matsunaga T, Esashi M (2002) Acceleration switch with extended holding time using squeeze film effect for side airbag systems. *Sensors Actuators A* 100:10–17
- Matsunaga T, Minami K, Esashi M (1999) Acceleration switch with extended holding time using squeeze film damping effect. In: *Proceedings of the international conference on solid-state sensors and actuators (Transducers'99)*, Sendai, vol 1999, pp 1550–1553
- Michaelis S, Timme HJ, Wycisk M et al. (1999) MEMS acceleration threshold switches fabricated with a low-cost, post-CMOS additive electroplating process. In: *Proceedings of electrochemical society spring meeting'99*, Seattle, vol 1999, p 1138
- Michaelis S, Timme HJ, Wycisk M (2000) Additive electroplating technology as a post-CMOS process for the production of MEMS acceleration-threshold switches for transportation applications. *J Micromech Microeng* 10:120–123
- Minami K, Matsunaga T, Esashi M (1999) Simple modeling and simulation of the squeeze film effect and transient response of the MEMS device. In: *Proceedings of IEEE MEMS'99*, Orland, vol 1999, pp 338–343
- Ongkodjojo A, Tay F (2006) Optimized design of a micromachined G-switch based on contactless configuration for health care applications. *J Phys Conf Ser* 34:1044–1052
- Park U, Yoo K, Kim J (2010) Development of a MEMS digital accelerometer (MDA) using a microscale liquid metal droplet in a microstructured photosensitive glass channel. *Sensors Actuators A* 159:51–57
- Petersen K, Shartel A, Raley N (1982) Micromechanical accelerometer integrated with MOS detection circuitry. *IEEE Trans Electron Devices* 29:23–27
- Rogers JW, Phinney LM (2001) Process yield for laser repair of aged, stiction-failed, MEMS devices. *J Microelectromech Syst* 2:280–285
- Selvakumar A, Yazdi N, Najafi K (1996) Low power, wide range threshold acceleration sensing system. In: *IEEE proceedings of the 9th annual international workshop on micro electro mechanical systems: an investigation of micro structures, sensors, actuators, machines, and systems*, pp 186–191
- Selvakumar A, Yazdi N, Najafi K (2001) A wide-range micromachined threshold accelerometer array and interface circuit. *J Micromech Microeng* 11:118–125
- Stauffer J (2006) Current capabilities of MEMS capacitive accelerometers in a harsh environment. *IEEE Aerosp Electron Syst Mag* 21:29–32
- Sun XQ, Zhou S, Carr WN (1997) A surface micromachined latching accelerometer. In: *Proceedings of international solid state sensors and actuators conference (Transducers'97)*, vol 2, pp 1189–1192
- Tao Y, Liu Y, Dong J (2014) Flexible stop and double-cascaded stop to improve shock reliability of MEMS accelerometer. *Microelectron Reliab* 54:1328–1337
- Tas N, Sonnenberg T, Jansen H, Legtenberg R, Elwenspoek M (1996) Stiction in surface micromachining. *J Micromech Microeng* 6:385–397

- Tønnesen T, Ludtke O, Noetzel J et al (1997) Simulation, design and fabrication of electroplated acceleration switches. *J Micromech Microeng* 7:237–239
- Wang Y, Feng Q, Wang Y, Chen W, Wang Z, Ding G, Zhao X (2013a) The design, simulation and fabrication of a novel horizontal sensitive inertial micro-switch with low g value based on MEMS micromachining technology. *J Micromech Microeng* 23:105013
- Wang Y, Chen W, Yang Z, Ding G, Wang H, Zhao X (2013b) An inertial micro-switch with compliant cantilever fixed electrode for prolonging contact time. In: *Proceedings of IEEE MEMS 2013, Taipei, 20–24 Jan, vol 2013*, pp 600–603
- Wang Y, Yang Z, Xu Q, Chen W, Ding G, Zhao X (2015a) Design, simulation and characterization of a MEMS inertia switch with flexible CNTs/cu composite array layer between electrodes for prolonging contact time. *J Micromech Microeng* 25:085012. (11pp)
- Wang Y, Yang Z, Chen W, Ding G, Wang Y, Zhang C, Zhao X (2015b) MEMS inertia switch with flexible CNTs/Cu composite array layer between electrodes for prolonging contact time. In: *18th international conference on solid-state sensors, actuators and microsystems (Transducers 2015)*, Anchorage, vol 2015, pp 1101–1104
- Whitley MR, Kranz M, Kesmodel R, Burgett S (2005) Latching shock sensors for health monitoring and quality control. In: *Progress in biomedical optics and imaging – proceedings of SPIE*, vol 5717, pp 185–194
- Whitley MR, Kranz M, Kesmodel R et al (2012) Latching shock sensors for health monitoring and quality control. In: *MEMS/MOEMS components and their applications II*, vol 2012
- Wittwer JW, Baker MS, Epp DS, Mitchell JA (2008) MEMS passive latching mechanical shock sensor. In: *Proceedings of the ASME, international design engineering technical conference & computers and information in engineering conference, Brooklyn, 3–6 Aug, vol 4*, pp 581–587
- Wycisk M, Tønnesen T, Binder J et al (1999) New sensor on-chip technology for micromechanical acceleration threshold switches. In: *Microelectronics and micro-electro-mechanical systems, Queensland, vol 1999*, pp 112–120
- Wycisk M, Tønnesen T, Binder J, Michaelis S, Timme HJ (2000) Low-cost post-CMOS integration of electroplated microstructures for inertial sensing. *Sensors Actuators A* 83:93–100
- Xu Q, Yang Z, Fu B, Li J, Wu H, Zhang Q, Sun Y, Ding G, Zhao X (2016) A surface-micromachining-based inertial micro-switch with compliant cantilever beam as movable electrode for enduring high shock and prolonging contact time. *Appl Surf Sci* 387:569–580
- Yang Z, Ding G, Chen W, Fu S, Sun X, Zhao X (2007) Design, simulation and characterization of an inertia micro-switch fabricated by non-silicon surface micromachining. *J Micromech Microeng* 17:1598–1604
- Yang Z, Ding G, Cai H, Zhao X (2008a) A MEMS inertia switch with bridge-type elastic fixed electrode for long duration contact. *IEEE Trans Electron* 55:2492–2497
- Yang Z, Ding G, Cai H, Liu R, Zhao X (2008b) Design and contact dynamics simulation of an inertia micro-switch based on non-silicon substrate. In: *Proceedings of the 3rd IEEE international conference on nano/micro engineered and molecular systems, vol 2008*, pp 19–22
- Yang Z, Ding G, Cai H, Xu X, Wang H, Zhao X (2009) Analysis and elimination of the ‘skip contact’ phenomenon in an inertial micro-switch for prolonging its contact time. *J Micromech Microeng* 19:045017. (10pp)
- Yang Z, Ding G, Zhou Z, Cai H, Zhao X (2010a) Analytical model of squeeze film damping for microelectro-mechanical systems structures with anti-stiction raised strips. *IET Micro Nano Lett* 5:258–261
- Yang Z, Ding G, Shen H, Wang H, Zhao X (2010b) Modeling and characterization of a MEMS G-sensor with anti-stiction raised strips for vibration monitoring systems. In: *proceedings of IEEE sensors 2010, Kona, 1–4 Nov*, pp 164–167
- Yang Z, Ding G, Wang H, Cai H, Zhao X (2011a) Modeling, simulation and characterization of a micromachined acceleration switch with anti-stiction raised strips on the substrate. *IEEE Trans Compon Packag Manuf Technol* 1:1198–1204
- Yang Z, Cai H, Ding G, Wang H, Zhao X (2011b) Dynamic simulation of a contact-enhanced MEMS inertia switch in simulink®. *Microsyst Technol* 17:1329–1342

- Yang Z, Zhu B, Chen W, Ding G, Wang H, Zhao X (2012a) Fabrication and characterization of a multidirectional-sensitive contact-enhanced inertial microswitch with an electrophoretic flexible composite fixed electrode. *J Micromech Microeng* 22:045006. (10pp)
- Yang Z, Zhu B, Ding G, Wang H, Wang Y, Zhao X (2012b) A multidirectional-sensitive inertial microswitch with electrophoretic polymer-metal composite fixed electrode for flexible contact. In: *Proceedings of IEEE MEMS 2012, Paris, 29 Jan–2 Feb, vol 2012*, pp 504–507
- Yoo K, Kim J (2009) A novel configurable MEMS inertial switch using microscale liquid-metal droplet. In: *IEEE international conference MEMS'09, vol 2009*, pp 793–796
- Yoo K, Park U, Kim J (2011) Development and characterization of a novel configurable MEMS inertial switch using a microscale liquid-metal droplet in a microstructured channel. *Sensors Actuators A* 166:234–240
- Younis MI, Alsalem FM, Miles R, Su Q (2007a) Characterization of the performance of capacitive switches activated by mechanical shock. *J Micromech Microeng* 17:1360–1370
- Younis MI, Alsalem FM, Jordy D (2007b) The response of clamped-clamped microbeams under mechanical shock. *Int J Non Linear Mech* 42:643–657
- Zhang Q, Yang Z, Ding G (2016a) Dynamic simulation analysis of off-axis sensitivity in a laterally-driven MEMS inertia switch. *AER-Adv Eng Res* 69:91–94
- Zhang Q, Yang Z, Xu Q, Wang Y, Ding G, Zhao X (2016b) Design and fabrication of a laterally-driven inertial micro-switch with multi-directional constraint structures for lowering off-axis sensitivity. *J Micromech Microeng* 26:055008. (15pp)
- Zhao J, Jia J, Chen G (2006) A novel MEMS parallel-beam acceleration switch. In: *Proceedings of 2nd IEEE/ASME international conference mechatron embedded syst, vol 2006*, pp 1–5
- Zhao J, Jia J, Wang H, Li W (2007) A novel threshold accelerometer with post buckling structures for airbag restraint systems. *J IEEE Sensors* 8:1102–1109
- Zhao J, Yang Y, Fan K, Hu P, Wang H (2010) A bistable threshold accelerometer with fully compliant clamped-clamped mechanism. *J IEEE Sensors* 10:1019–1024
- Zhu B, Yang Z, Chen W, Liu Q, Ding G, Zhao X (2012) Design and characterization of an inertial microswitch with electrophoretic polymer-metal composite fixed electrode for extending contact time. *IET Micro Nano Lett* 7:501–504
- Zimmermann L, Ebersohl J-P, Hung FL, Berry JP, Baillieu F, Rey P, Diem B, Renard S, Caillat P (1995) Airbag application: a microsystem including a silicon capacitive accelerometer, CMOS switched capacitor electronic and true self-test capability. *Sensors Actuators A* 46:190–195



# A Rotary Microgripper

Honglong Chang and Yongcun Hao

## Contents

Introduction .....	998
Microgripper Design .....	999
Working Principle .....	999
Design of the Grasping Structure .....	1003
Design of the Ratchet Mechanism .....	1003
Design of the Rotary Actuator .....	1005
Performance Evaluation with FEA Simulation .....	1006
Modal Analysis .....	1006
Stress Analysis .....	1006
Fabrication of the Microgripper .....	1007
Experiments Setup .....	1009
Experiments and Discussions .....	1011
Characterization of the Microgripper .....	1011
Patterning Pollen Cells Using the Gripper .....	1012
Conclusions .....	1013
References .....	1014

## Abstract

This chapter presents the design, analysis, fabrication, and testing of a micro-electromechanical systems (MEMS) rotary microgripper with self-locking function via a ratchet mechanism. The microgripper utilizes rotary actuators to solve the pull-in problem of microgrippers during large displacement manipulation and

H. Chang (✉)

MOE Key Laboratory of Micro/Nano Systems for Aerospace, Northwestern Polytechnical University, Xi'an, China

Northwestern Polytechnic University, Xi'an, China

e-mail: [changhl@nwpu.edu.cn](mailto:changhl@nwpu.edu.cn)

Y. Hao

Northwestern Polytechnic University, Xi'an, China

therefore avoids the widely used conversion systems which necessitate a high driving voltage. Furthermore, the ratchet mechanism enables long-time gripping without having to continuously apply the external excitation signal. Thus, the damage to the gripped micro-scale objects caused by the external excitation signals can be significantly reduced. The microgripper is fabricated by a silicon-on-insulator (SOI) dicing-free process to protect the delicate device structures from damage during fabrication. The microgripper has a discrete opening range and can handle micro-scale objects with a size of 20  $\mu\text{m}$ , 40  $\mu\text{m}$ , and 60  $\mu\text{m}$  based on the current design parameters. Test results showed the gripper obtained a displacement of 100  $\mu\text{m}$  with an applied voltage of 31.5 V. A pick-lock-release gripping experiment on a magnolia pollen cell is performed to form a triangle to prove the feasibility of the gripper in handling biological cells.

---

**Keywords**

Microelectromechanical systems · Rotary · Microgripper · Locking · Ratchet mechanism · Silicon-on-insulator · Dicing-free · Biological cells

---

**Introduction**

The microgripper provides an effective way to manipulate micro-scale objects and has been widely used in assembling micro objects, testing micro-devices (Cecil et al. 2007), and handling biological cells and tissues. A variety of microgrippers based on electrostatic (Piriyant et al. 2015), thermal (Kim et al. 2008; Zhang et al. 2012), piezoelectric (Wang et al. 2013, 2015; Zubir et al. 2009; Liang et al. 2017), magnetic (Ren and Gerhard 1997; Ger et al. 2013), pneumatic (Buettfisch and Buettgenbach 2001; Alogla et al. 2013), and direct mechanical (Wester et al. 2011) actuating mechanisms have been presented in recent decades. Among them, electrostatically actuated microgrippers are of considerable interest because they can be fabricated in a batch-compatible process with small device size, low power consumption, and low temperature dependence (Beyeler et al. 2006); they have a capability in precise and stable controlling over micron scale movement. However, it is challenging for electrostatic microgrippers to be used in many applications due to insufficient grasping scope which is confined by the straight comb actuators.

Many efforts have been made to achieve a larger displacement with a lower driving voltage for electrostatic microgrippers. A silicon-on-insulator (SOI) microgripper driven by straight electrostatic actuators was reported in Volland et al. (2002). This gripper utilized a mechanical conversion system comprising a group of flexible beams to convert and amplify the longitudinal motion of the comb actuators into a gripping rotational motion. With this design, the straight comb actuators moved longitudinally and achieved a large operating displacement without suffering from the pull-in problem. It achieved a displacement of 20  $\mu\text{m}$  with an applied voltage of 80 V. Similar conversion systems were used in electrostatic microgrippers in Beyeler et al. (2007), Chen et al. (2010), and Bazaz et al. (2011). The gripper in Beyeler et al. (2007) achieved a displacement of 100  $\mu\text{m}$  with a

driving voltage of 150 V, while the gripper in Chen et al. (2010) obtained a 25  $\mu\text{m}$  displacement with a driving voltage of 80 V. The usage of conversion systems effectively solved the pull-in problem of straight comb actuators during large displacement operation. However, such conversion systems necessitated a large driving force, which made the driving voltage in the order of 100 V or higher. It may be problematic to supply such a high driving voltage in practical applications. Actually, the actuating signals including electrostatic, thermal, piezoelectric, and magnetic may affect or even damage the micro objects such as biological tissues or biological cells in the long-time manipulating experiments (Choi et al. 2013; Lai and Singh 2004; Valic et al. 2003).

To achieve a larger displacement, rotary comb actuators are proposed to be employed in microgrippers (Chang et al. 2012). The rotary comb actuator can generate a larger displacement than a straight one because it can efficiently suppress the lateral pull-in phenomenon. Furthermore, the displacement of a rotary comb actuator can be amplified to the gripper tip by a long gripper arm. Therefore, a rotary comb actuator can directly generate a large angular displacement without the inefficient mechanical conversion system. With this design approach, the proposed rotary comb-actuated microgripper can achieve a larger displacement consuming a lower driving voltage than the existing microgrippers.

The microgripper with self-locking mechanism can alleviate the effect on gripped objects generated by the actuating signals (Yuan et al. 2015). In this chapter, the ratchet mechanism is also introduced into the microgripper to achieve the self-locking function. The ratchet mechanism is a mechanical device that allows the linear or rotary motion in only one direction while preventing the motion in the opposite direction. The ratchet mechanism has been successfully used in MEMS actuators (Barnes et al. 2000; Pham et al. 2006) to lock the actuators with preventing the reverse movement. However, the ratchet pawl in Barnes et al. (2000) and Pham et al. (2006) is a passive mechanical structure which means that the ratchet pawl cannot be separated from the ratchet tooth without external actuators. In this chapter, the control method on the ratchet pawl by a lateral electrostatic actuator is also introduced. Then the microgripper only requires the excitation signal in the moment of picking and releasing the micro-objects, while in the long-time locking state it is kept by the mechanical structure.

---

## Microgripper Design

### Working Principle

Schematic of the presented rotary microgripper is depicted in Fig. 1. The microgripper consists of a pair of jaws with an initial gap of 100  $\mu\text{m}$ . The jaws connect to the curved shuttles via the long arms that transfer and amplify the output displacement generated by the rotary actuator. The shuttles are integrated with the ratchet teeth which work with ratchet pawls to form the ratchet mechanism. The ratchet mechanism, as depicted in Fig. 2, mainly consists of the ratchet beam, the lateral actuator, the ratchet pawl, the ratchet tooth, and the stop cusp. The ratchet

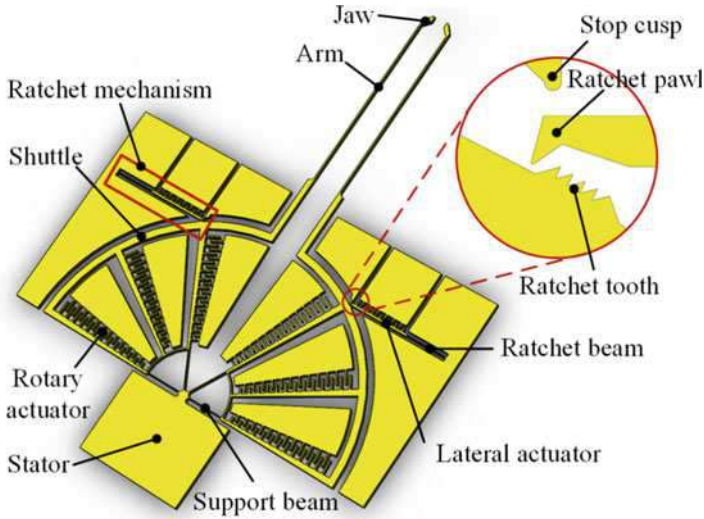


Fig. 1 Schematic of the proposed microgripper with ratchet mechanism

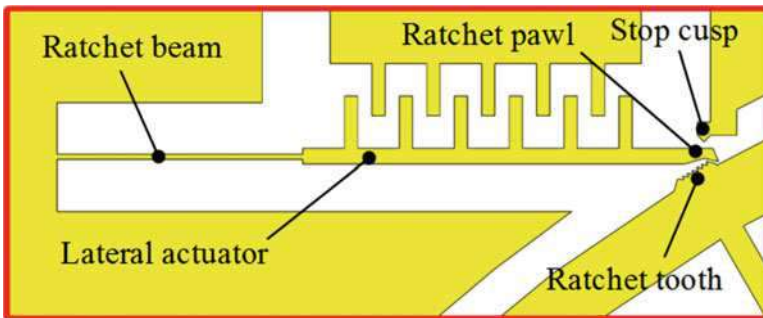
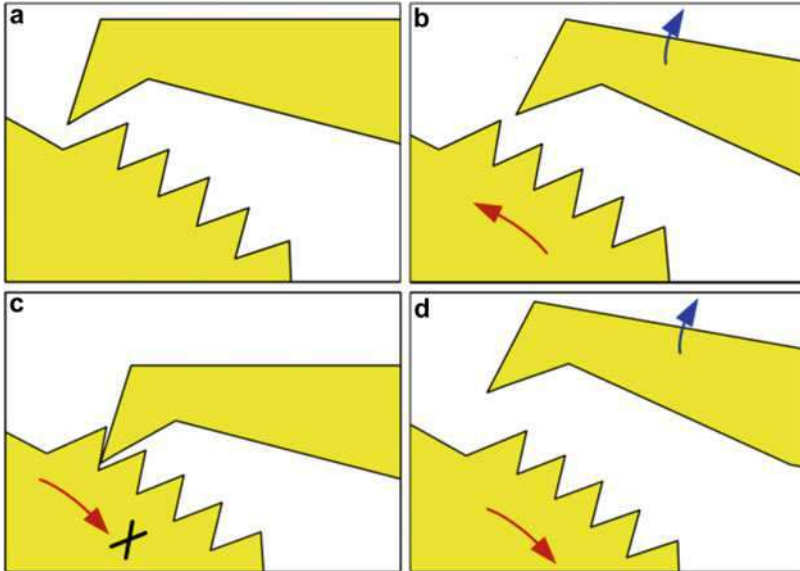


Fig. 2 Schematic view of the ratchet mechanism

mechanism restricts the opening of the arms and dominates the release of the micro-scale object by lateral actuator. In addition, a stop cusp is designed to prevent the pull-in problem of the lateral actuator.

The working process of the gripper includes three stages: picking, locking, and releasing. The corresponding states of the ratchet mechanism are schematically shown in Fig. 3. Figure 3a show the initial position of the ratchet teeth and ratchet pawl. Figure 3b shows the picking process in which the pawl is lifted and the shuttle rotates in anticlockwise to pick up the micro-object. In Fig. 3c, the ratchet pawl engages with the ratchet tooth and prevents the microgripper from rotating in the reverse direction. Thus, the micro-object is locked by the mechanical structure. In Fig. 3d, the ratchet pawl and ratchet tooth are separated, then the gripper returns to the initial position and the gripped objects is released.





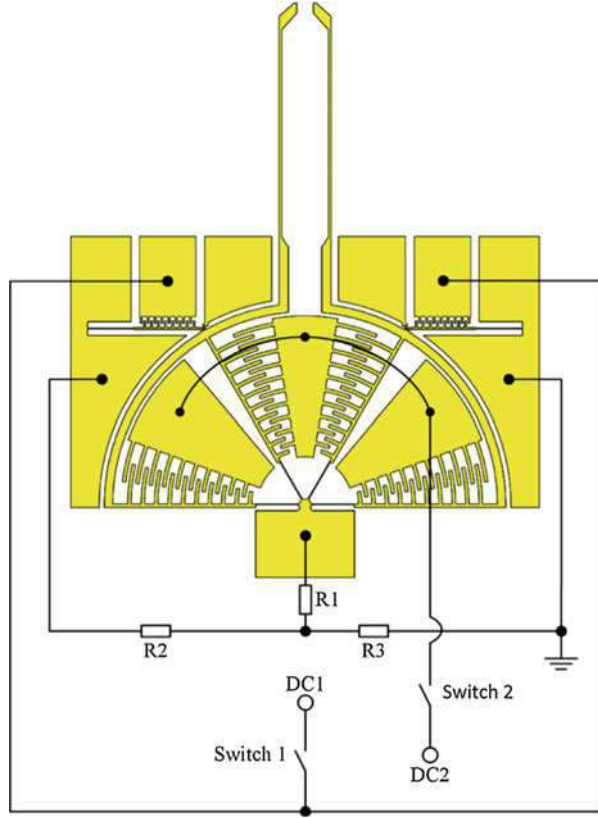
**Fig. 3** Working process of the ratchet mechanism. **a** Initial position. **b** Picking. **c** Locking. **d** Releasing

The schematic circuit diagram is shown in Fig. 4. The switch 1 controls the rotary actuator circuit and the switch 2 controls the ratchet mechanism circuit. The actuator circuit and ratchet mechanism circuit is equivalent to a RC circuit. The resistors (R1, R2, and R3) are used to reduce the oscillation frequency of the RC circuit to eliminate the complex vibration mode of microgripper and improve the dynamic characteristics of the system.

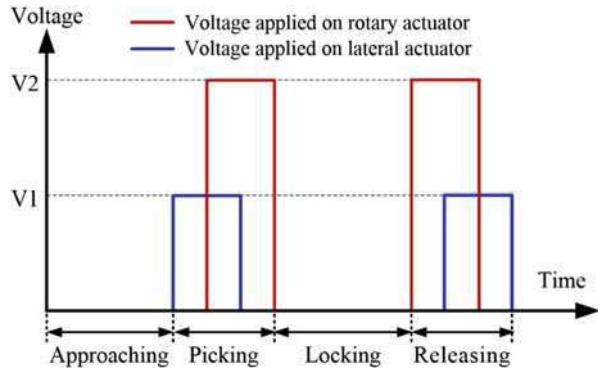
To avoid the wearing between the ratchet pawl and ratchet tooth, a logic control strategy as described in Fig. 5 is proposed to manipulate the gripper. In picking stage, the lateral actuator is applied with the voltage V1 to lift the ratchet pawl and the rotary actuator is driven to pick up the micro-object with voltage V2. The excitation signal of the lateral actuator and the rotary actuator is switched off successively and the microgripper is locked. In releasing stage, the rotary actuator is applied with the voltage V2 to eliminate the pressure on ratchet pawl from the ratchet tooth. Then, the pawl is lifted up by the lateral actuator with voltage V1. The driving voltage of the rotary actuator and the lateral actuator is powered off in order and the microgripper is released. There is no contact between the ratchet pawl and the ratchet tooth in the picking step and releasing step. So the logic control strategy protects the ratchet mechanism from wearing by eliminating the friction between ratchet pawl and ratchet tooth.

As shown in Fig. 5, during the working process of the proposed microgripper, the excitation signal is applied only in the moment of picking and releasing step, and the excitation signal is set to zero in the locking step because the gripper is locked by the ratchet mechanical structure. However, as shown in Fig. 6, for the gripper without self-locking function, the excitation signal will last during the whole working

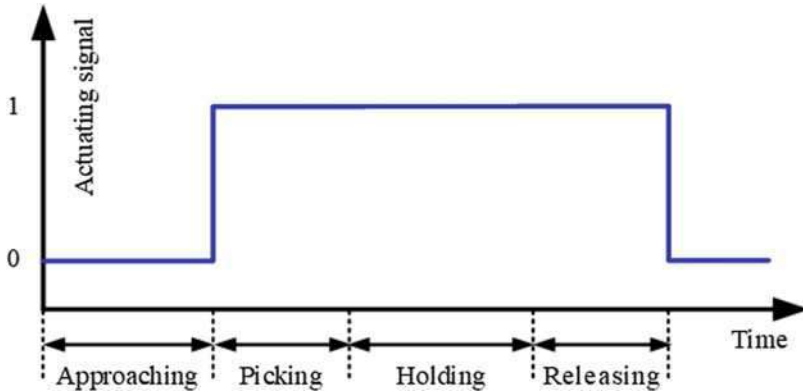
**Fig. 4** Schematic circuit diagram of microgripper



**Fig. 5** Logic control strategy of the proposed gripper



process from the picking step to releasing step. For experiments that the objects will be hold for long time, the proposed microgripper can significantly reduce the effect and damage on the gripped objects by shortening the exposure time of electrical, thermal, or magnetic fields generated by excitation signals.



**Fig. 6** Typical logic control strategy of the gripper without self-locking function

### Design of the Grasping Structure

The grasping structure including the jaw, arm, shuttle, and support beam is the main part of the microgripper. The geometries of the grasping structure determine the grasping scope of the microgripper. The displacement of jaw can be amplified by the arm. The geometrical relationship of the gripper is shown in Fig. 7 and the amplification  $\alpha$  can be described as:

$$\alpha = \frac{D2}{D1} = \frac{L1 + L2}{L1} \tag{1}$$

where, D1 is the displacement of the shuttle, L1 is the radius of the shuttle, D2 is the displacement of the jaw, and L2 is the length of the jaw.

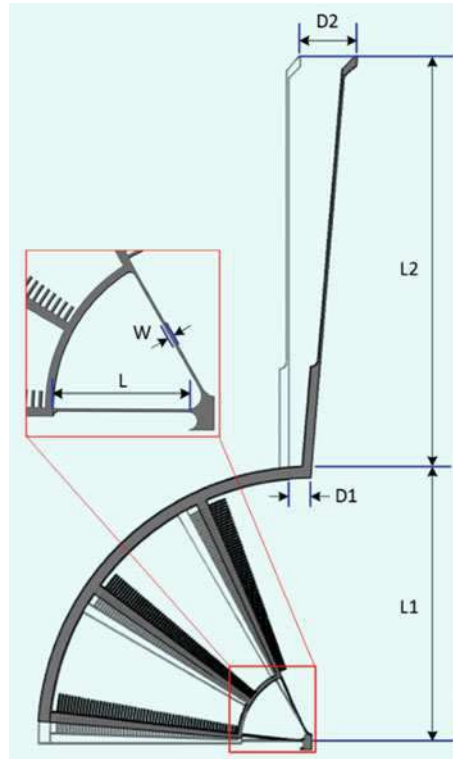
In this design, L1 is 1,200  $\mu\text{m}$  and L2 is 1,800  $\mu\text{m}$  which means an amplification of 2.5. Since the maximum displacement of the shuttle is 20  $\mu\text{m}$ , the displacement of the jaw can reach 50  $\mu\text{m}$ .

The stiffness of support beam will affect the displacement of the jaws. To drive the gripper at a lower voltage, the support beam with a smaller width and larger length is expected. However, the width of the beam is at least 5  $\mu\text{m}$  and the length-width ratio should be less than 50 which is confined by our process capability. In this chapter, the width (W) and length (L) of the support beam is set as 5  $\mu\text{m}$  and 250  $\mu\text{m}$ , respectively, in order to realize the available minimum stiffness of support beam.

### Design of the Ratchet Mechanism

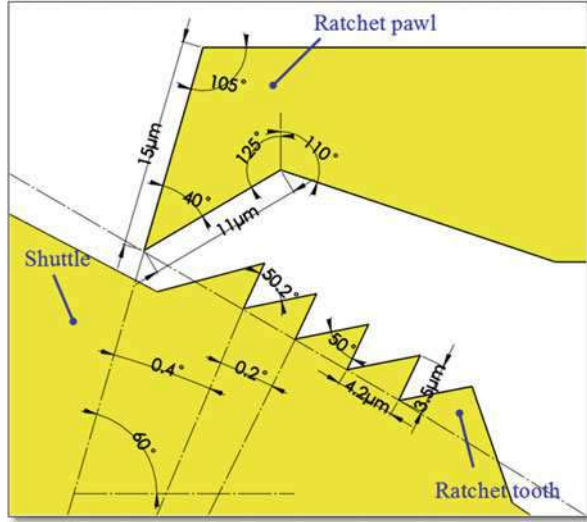
The working principle of the gripper makes it to only grip micro-objects with specific sizes. Actually, the size that can be gripped is greatly dependent on the

**Fig. 7** The geometrical relationship of the grasping structure



dimension of the ratchet pawl and ratchet teeth, which is usually defined by the MEMS fabrication process. As shown in Fig. 8, the ratchet teeth are designed to be  $3.5\ \mu\text{m}$  tall and  $4.2\ \mu\text{m}$  wide with a teeth angle of  $50^\circ$ . The tip of the ratchet pawl is designed on the root circle of the ratchet teeth which coincides with the rotation center of the microgripper, and the angle of the pawl is designed  $10.2^\circ$  smaller than the tooth root angle of  $50.2^\circ$  for easily engaging and separating from each other. The radial line which connects the pawl tip and the gripper rotation center has an angle of  $60^\circ$  relative to the horizontal line. The angle between radial line and the first ratchet tooth is  $0.4^\circ$ , while the angle for the adjacent tooth is  $0.2^\circ$ . Considering that the jaws are closed with a  $50\ \mu\text{m}$  movement which corresponds to the shuttle rotation angle of  $1^\circ$ , the first engagement will occur when the jaw moves  $20\ \mu\text{m}$ , and the subsequent engagement will occur for a  $10\ \mu\text{m}$  movement of the jaws. Thus, the engagement and separation occur when the jaw reaches a displacement of  $20\ \mu\text{m}$ ,  $30\ \mu\text{m}$ ,  $40\ \mu\text{m}$ , and  $50\ \mu\text{m}$ . So, the microgripper can grip a target with a size of  $20\ \mu\text{m}$ ,  $40\ \mu\text{m}$ , and  $60\ \mu\text{m}$  as the initial gap of the jaws is  $100\ \mu\text{m}$ . But during the operation, the elastic deformation of the slender arms enables the microgripper grip objects that are slightly larger than  $20\ \mu\text{m}$ ,  $40\ \mu\text{m}$ , and  $60\ \mu\text{m}$ .

**Fig. 8** Geometry of the ratchet pawl and tooth



### Design of the Rotary Actuator

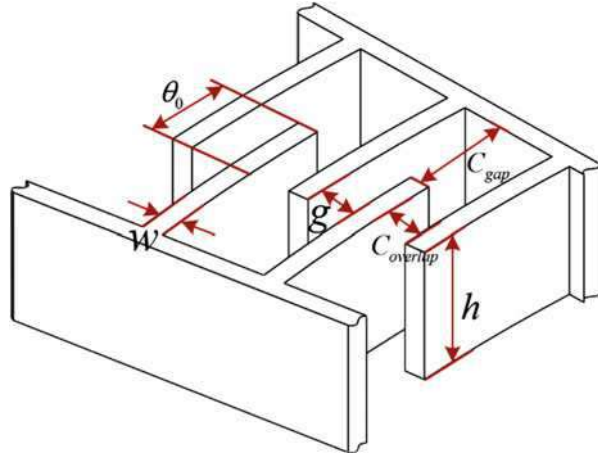
The fabricated gripper is an electrostatic driving device using rotary actuators to grip the micro-scale object. The rotary actuator contains two concentric interdigitated curved fingers – one set of fingers are fixed onto the anchor while the other set are suspended by support beams. As depicted in Fig. 9, the total capacitance  $C_{tot}$  between moveable comb finger and fixed comb finger is mainly composed of two parts, the overlap capacitance  $C_{overlap}$  and the gap capacitance  $C_{gap}$ . In most practical cases, the gap capacitance  $C_{gap}$  is sufficiently small to be neglected. The total capacitance  $C_{tot}$  of the rotary electrostatic actuator (Tu et al. 2005) can be calculated from (Eq. 2),

$$C_{tot} = \epsilon_0(\theta + \theta_0)h \left\{ \sum_{i=1}^{n-1} [\ln(a_i)]^{-1} + \sum_{i=0}^{n-1} [\ln(b_i)]^{-1} \right\} \quad (2)$$

where  $a_i = \frac{R_0+2i(w+g)}{R_0+2i(w+g)-g}$ ,  $b_i = \frac{R_0+(2i+1)(w+g)}{R_0+2i(w+g)+w}$ ,  $\epsilon_0$  is the free-space permittivity with a value of  $8.85 \times 10^{-12} \text{Fm}^{-1}$ ,  $\theta$  is the rotated angle,  $\theta_0$  is the initial engagement angle,  $h$  is the thickness of comb fingers,  $R_0$  is the radius of the first finger,  $w$  is the finger width,  $g$  is the gap between fingers.  $i$  refers to the  $i$ th comb finger, and  $n$  is the number of fingers. The key structure parameter of the rotary actuator is listed in Table 1.

An electrostatic attraction force is created when a voltage  $V$  is applied to the rotary actuator. The generated torque  $\tau$  is proportional to the square of the excitation voltage and the differential change of capacitance as expressed in (3),

**Fig. 9** Solid model of the rotary actuator



**Table 1** Key structure parameter of the rotary actuator

Parameter	$R_0$	$w$	$g$	$h$	$n$
Value	335 $\mu\text{m}$	4 $\mu\text{m}$	3 $\mu\text{m}$	60 $\mu\text{m}$	56

$$\tau = \frac{1}{2} \left( \frac{\partial C_{tot}}{\partial \theta} \right) V^2 = \frac{1}{2} \varepsilon_0 h \left\{ \sum_{i=1}^{n-1} [\ln(a_i)]^{-1} + \sum_{i=0}^{n-1} [\ln(b_i)]^{-1} \right\} V^2 \quad (3)$$

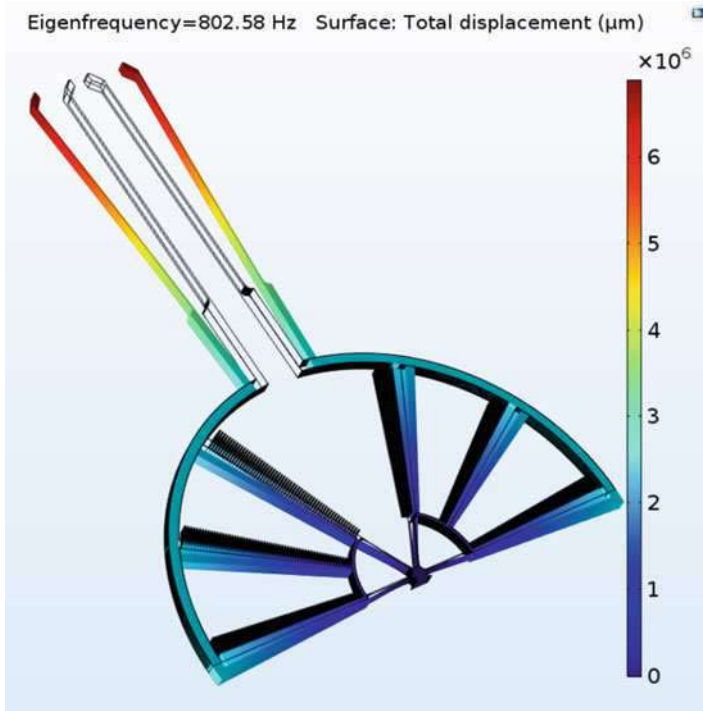
## Performance Evaluation with FEA Simulation

### Modal Analysis

A modal analysis was performed using COMSOL Multiphysics to predict the dynamic performance and the natural frequency of the proposed microgripper. The proposed microgripper rotates in the gripping direction on its first natural frequency as shown in Fig. 10. Table 2 lists the first three natural frequencies of the microgripper. The driving frequency should be lower than the gripper’s first modal frequency to avoid the structure resonance in manipulations.

### Stress Analysis

During grasping, the deformation of the spring beam and elastic beam will produce a considerable stress. The stress concentration occurs usually near the anchor which may lead to structural failure. So, it is necessary to evaluate the stress distribution contours of microgripper. Figure 11 shows the physical conditions of the gripper for



**Fig. 10** First order natural frequency of the grasping structure

**Table 2** Resonant frequency of the grasping structure

Order	1	2	3
Frequency(Hz)	803	3,160	3,309

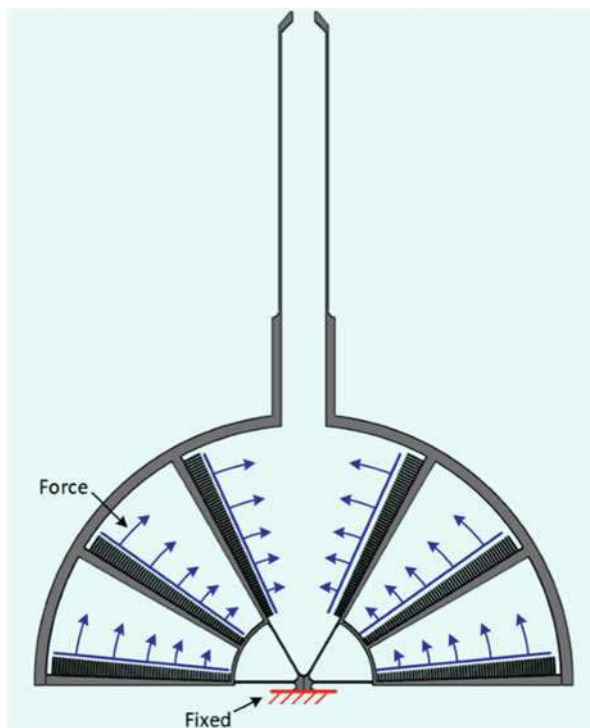
the stress analysis. The gripper is fixed at the root of the support beams and the equivalent electrostatic force is applied on the curved fingers.

Figure 12 show the Von Mises stress contour plot of grasping structure. The maximum stress occurs near the anchor with a value of about 33 MPa. The maximum stress is much less than yield strength of single crystal silicon which is 7 GPa (Zhang et al. 2012).

## Fabrication of the Microgripper

The microgripper can be fabricated using a relative simple single-mask process (Hao et al. 2015). However, there are three problems which should be taken into consideration. First, the remaining handle layer under the jaws of the gripper brings troubles in handling the micro-objects. Second, the under etching in the single-

**Fig. 11** Physical conditions of the gripper for the stress analysis



mask selective release process is hard to be avoided, which leads to reduction in the thickness of the structure. Third, the dicing step of the fabrication process probably damages the delicate device structures of the gripper due to the impact of water flow used for cooling the blade and wafer. To solve the three problems, a dicing free process based on wet releasing technology is used to fabricate the microgripper. In the fabrication process, wet etching is used to release the moveable structure and detach the devices from the wafer grid. The detailed fabrication process flow is illustrated in Fig. 13.

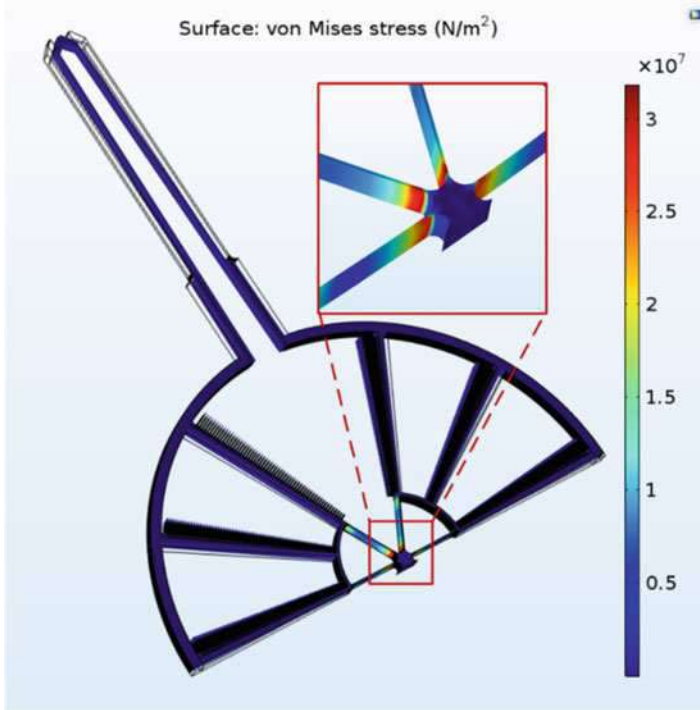
*Step (a).* A silicon-on-insulator wafer with a device layer of  $60\ \mu\text{m}$ , a handle layer of  $400\ \mu\text{m}$ , and a buried  $\text{SiO}_2$  layer of  $4\ \mu\text{m}$  is used for the process.

*Step (b).* Patterned with aluminum on the handle layer and deep reactive ion etching (DRIE) etching until the buried  $\text{SiO}_2$  layer to form the backside trenches and cavity.

*Step (c).* HF wet etch to remove the buried  $\text{SiO}_2$ .

*Step (d).* DRIE through etching the device layer to form the device features and front trenches on the top side. A support wafer with grooves is used in this step to balance the pressure of the two sides of the wafer. Otherwise, the device layer tends to crack inside the process chamber of DRIE due to the difference in pressure on the two sides of the wafer.





**Fig. 12** Von Mises stress contour of the grasping structure

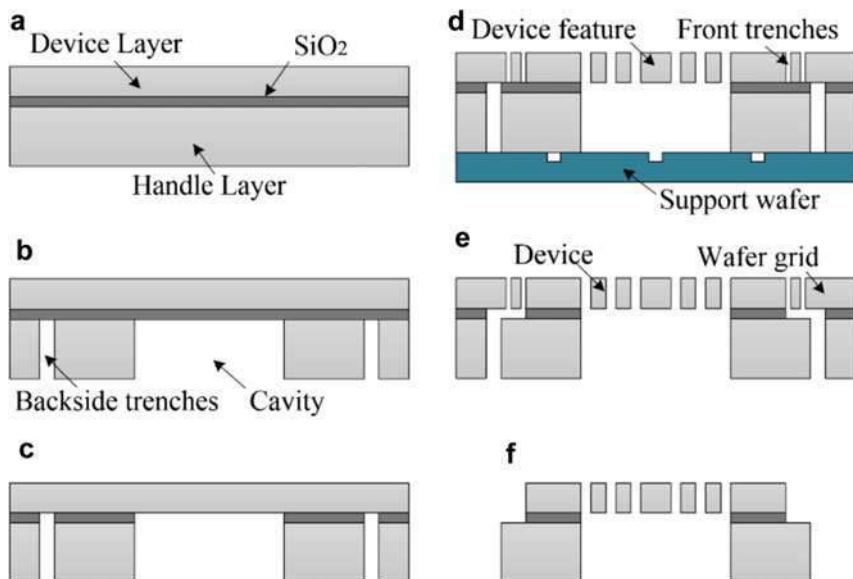
*Step (e).* Detach the support wafer and remove the  $\text{SiO}_2$  under the front trenches using HF wet etch.

*Step (f).* Unload the fabricated microgripper from the wafer grid and the fabrication is completed.

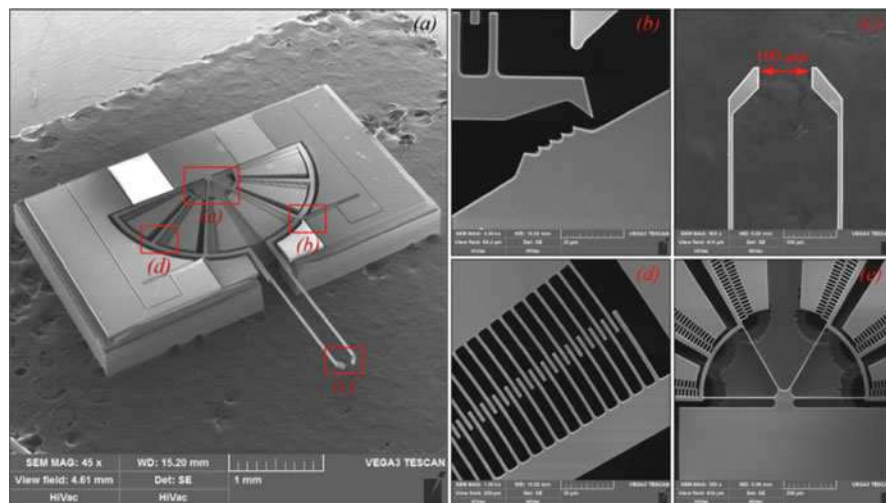
Figure 14 is the SEM images of the fabricated microgripper. Figure 14a shows the global view of the microgripper with a general size of  $4,200 \times 3,800 \mu\text{m}$ . Figure 14b is the ratchet mechanism and shows the morphology of the ratchet tooth and ratchet pawl. Figure 14c shows the jaws with a gap of  $100 \mu\text{m}$ . Figure 14d is the curved fingers with a width of  $4 \mu\text{m}$  and a gap of  $3 \mu\text{m}$ . Figure 14e is the support beams with a width of  $5 \mu\text{m}$  and length of  $250 \mu\text{m}$ .

## Experiments Setup

To operate the microgripper, a control and observing system is set up. As shown in Fig. 15, the system consists of a printed circuit board (PCB) as the carrier of the microgripper, a position platform with four degrees of freedom (4-DOF) which is heaving, swaying, surging, and pitching to adjust the position of microgripper,

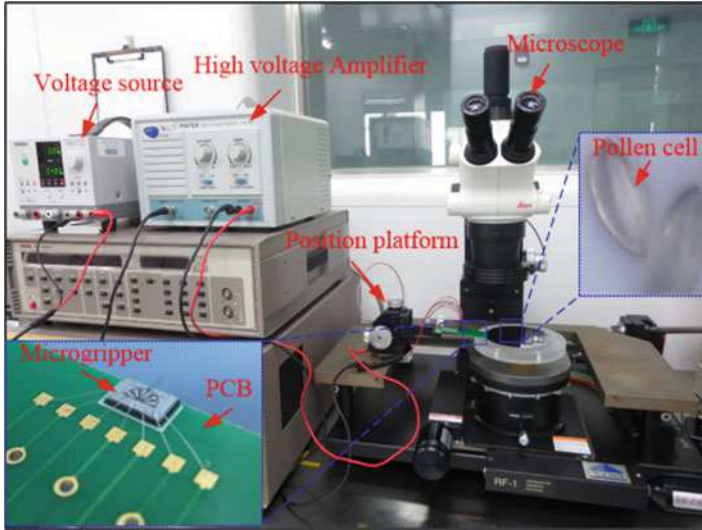


**Fig. 13** Fabrication process of the microgripper



**Fig. 14** SEM of the manufactured microgripper. (a) The global view. (b) The ratchet mechanism. (c) The jaws. (d) The curved fingers. (e) The support beams

a voltage source and a high voltage amplifier to supply and increase the driving signal, and a microscope for observing. All experiments are performed at room temperature with a relative humidity of 46%. And the magnolia pollen cells with an equatorial diameter around 21  $\mu\text{m}$  are chosen to be gripped by the microgripper.



**Fig. 15** Apparatus setup and arrangement

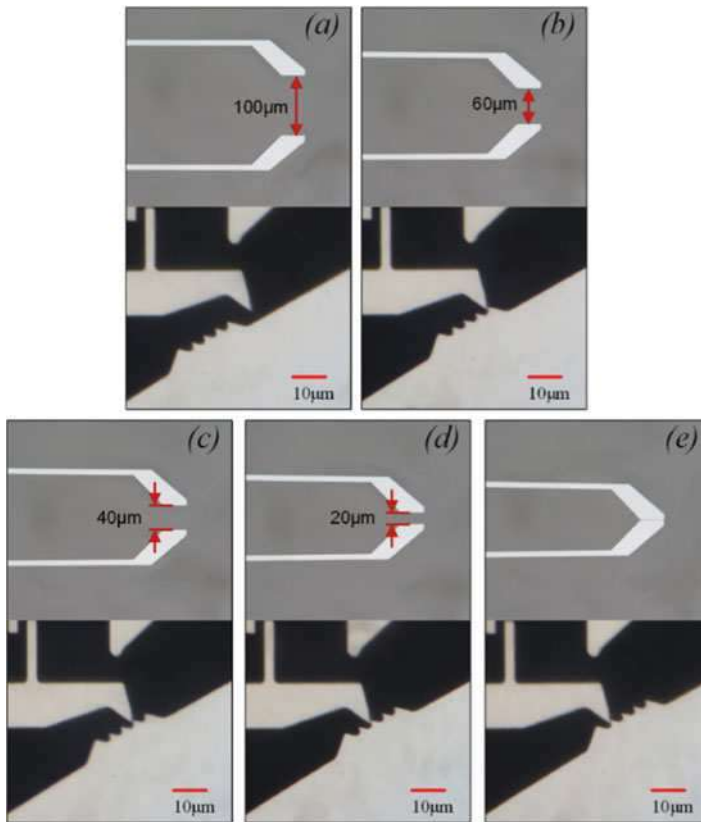
## Experiments and Discussions

### Characterization of the Microgripper

Using the logic control strategy described in Fig. 5, the gripper is actuated without gripping any object to confirm the operating scope and the corresponding excitation voltage.

Figure 16 shows the state of the jaws and the ratchet mechanism with different excitation voltage. In Fig. 16a, the gap distance between jaws is 100  $\mu\text{m}$  and the ratchet mechanism is in the initial state. In Fig. 16b, the gap reduces to 60  $\mu\text{m}$  with an actuating voltage of 20.5 V, and the ratchet pawl is engaged with the first ratchet tooth. The gap becomes 40  $\mu\text{m}$  with an actuating voltage of 24.5 V, and the ratchet pawl is engaged with the second ratchet tooth (Fig. 16c). In Fig. 16d, the gap becomes 20  $\mu\text{m}$  with an actuating voltage of 28 V, and the ratchet pawl is engaged with the third ratchet tooth. And the jaw is closed with a movement of 50  $\mu\text{m}$  when a voltage of 31.5 V is applied, and the ratchet pawl is engaged with the fourth ratchet tooth as depicted in Fig. 16e. The voltage of lifting the ratchet pawl in picking stage is always 16 V, but the voltage to separate the ratchet pawl from ratchet tooth in releasing stage, as summarized in Table 3, is variable. It is likely caused by the adhesion between ratchet pawl and ratchet tooth which is generated for their contact in locking stage.

Table 4 compares the operating displacement of the microgripper prototype with recent published results. With a driving voltage of 31.5 V, the arm tips relative displacement reach to 100  $\mu\text{m}$ . The results demonstrate the proposed structure can



**Fig. 16** The state of jaws (*upper*) and ratchet mechanism (*lower*) with different excitation voltage

effectively suppress the pull-in problem in large displacement manipulation. Meanwhile, we can see the proposed microgripper shows a good performance in operating displacement considering it moved  $100\ \mu\text{m}$  with an applied electrostatic force of  $31.5\ \text{V}$  which is much smaller than the previously reported gripper. The driving voltage can be further reduced through the optimization of the supporting beams and actuator dimensions.

### **Patterning Pollen Cells Using the Gripper**

The performance of the gripper is assessed by patterning the magnolia pollen cells, which is about  $21\ \mu\text{m}$  in equatorial diameter. First, the gripper slowly approaches to the cell, as illustrated in Fig. 17a. Then the ratchet pawl is lifted with a voltage of  $16\ \text{V}$  and the cell was gripped with a voltage of  $29.5\ \text{V}$  applied on the rotary actuator (Fig. 17b). The excitation signal of the lateral actuator and the rotary actuator is

**Table 3** Voltage of separating the ratchet pawl from ratchet tooth in releasing stage

Order of tooth	Voltage
First	16 V
Second	16 V
Third	16.5 V
Fourth	17.5 V

**Table 4** The operating displacement of the microgripper when actuated

Arm tip displacement ( $\mu\text{m}$ )	Driving voltage (V)	Reference
20	80	Volland et al. (2002)
100	150	Beyeler et al. (2007)
25	80	Chen et al. (2010)
17	50	Bazaz et al. (2011)
100	31.5	This work

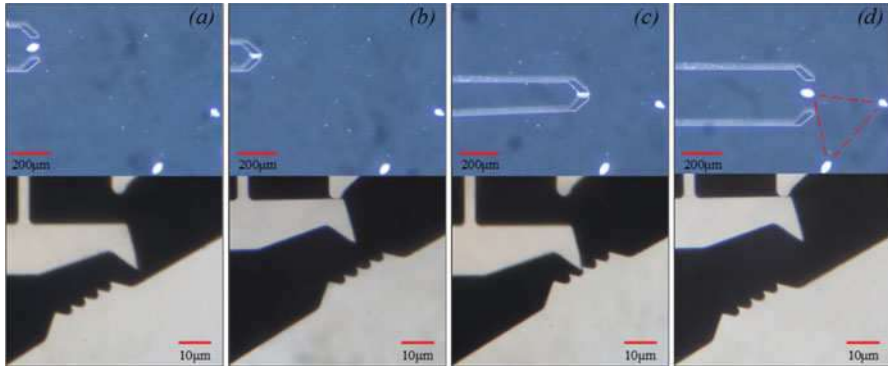
powered off in order, thus the ratchet pawl is engaged with the third ratchet tooth simultaneously and the pollen cell is locked. In Fig. 17c, the pollen cell is moved to the destination and left in air for 24 h to suffer the impact of the surroundings. Finally, the rotary actuator is applied with a voltage of 29.5 V and the ratchet pawl is separated from the ratchet tooth when the lateral actuator is applied with a voltage of 17 V. Powering off the excitation signal of the rotary actuator and lateral actuator successively, the pollen cell is released and forms a triangle with two other cells (Fig. 17d). The logic control strategy in the patterning experiment is shown in Fig. 18. The experiment demonstrates the capability of the gripper in handling micro-objects and the long-time locking function.

The equatorial diameter of the gripped magnolia pollen cell is about 1  $\mu\text{m}$  larger than 20  $\mu\text{m}$  which is the gap distance between two jaws in Fig. 16d, so the rotary actuators must overcome the deformation of slender arm to lock and release the pollen cell. Thus, the voltage applied on the rotary actuator is larger than the corresponding value in Table 3. Hence, the slender arm can tolerate the size deviation to some extent and makes the microgripper more practicable and flexible.

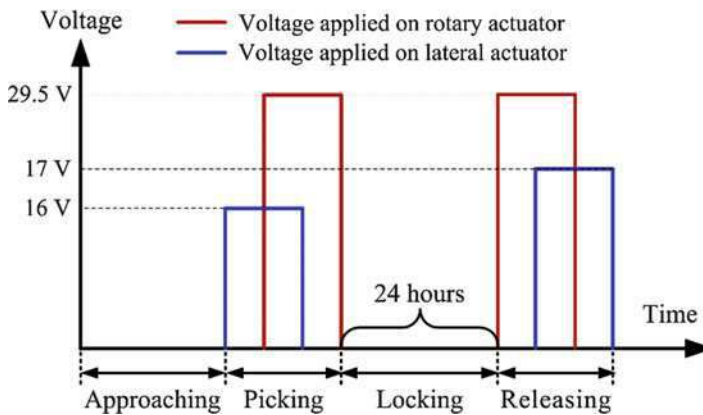
---

## Conclusions

A rotary microgripper with locking function via a ratchet mechanism is presented in this chapter. The feasibility of the gripper is demonstrated by patterning the magnolia pollen cells. Currently, the microgripper only can handle micro-scale objects with a specific size of 20  $\mu\text{m}$ , 40  $\mu\text{m}$ , and 60  $\mu\text{m}$ , but the range can be expanded to several microns or hundreds of microns according to practical requirements by adjusting the position of the ratchet tooth and the space of the jaws.



**Fig. 17** Experiment of patterning the magnolia pollen cells. (a) Approaching. (b) Picking. (c) Locking. (d) Releasing



**Fig. 18** Logic control strategy in the patterning experiment

## References

- Alogla A, Scanlan P, Shu WM, Reuben RL (2013) A scalable syringe-actuated microgripper for biological manipulation. *Sensors Actuators A Phys* 202:135–139
- Barnes S, Miller S, Rodgers M, Bitsie F (2000) Torsional ratcheting actuating system. In: 2000 international conference on modeling and simulation of microsystems – MSM 2000. Computational Publications, San Diego, pp 273–276
- Bazaz SA, Khan F, Shakoor RI (2011) Design, simulation and testing of electrostatic SOI MUMPs based microgripper integrated with capacitive contact sensor. *Sensors Actuators A Phys* 167:44–53
- Beyeler F, Bell DJ, Nelson BJ et al (2006) Design of a micro-gripper and an ultrasonic manipulator for handling micron sized objects. In: 2006 IEEE/RSJ international conference on intelligent robots and systems. IEEE, Beijing, pp 772–777

- Beyeler F, Neild A, Oberti S et al (2007) Monolithically fabricated microgripper with integrated force sensor for manipulating microobjects and biological cells aligned in an ultrasonic field. *J Microelectromech Syst* 16:7–15
- Buetefisch S, Buettgenbach S (2001) A new pneumatically actuated miniature gripper for micro assembly. In: Nelson BJ, Breguet J-M (eds) *Proceedings of SPIE 4568, microrobotics and microassembly III*, 32 (October 8, 2001). SPIE, Newton, pp 32–39
- Cecil J, Powell D, Vasquez D (2007) Assembly and manipulation of micro devices—a state of the art survey. *Robot Comput Integr Manuf* 23:580–588
- Chang H, Zhao H, Xie J et al (2012) Design and fabrication of a rotary comb-actuated microgripper with high driving efficiency. In: 2012 I.E. 25th international conference on micro electro mechanical systems (MEMS). IEEE, Paris, pp 1145–1148
- Chen T, Sun L, Chen L et al (2010) A hybrid-type electrostatically driven microgripper with an integrated vacuum tool. *Sensors Actuators A Phys* 158:320–327
- Choi J, Morrissey M, Bischof JC (2013) Thermal processing of biological tissue at high temperatures: impact of protein denaturation and water loss on the thermal properties of human and porcine liver in the range 25–80 °C. *J Heat Transf* 135:61302
- Ger T, Huang H, Chen W, Lai M (2013) Magnetically-controllable zigzag structures as cell microgripper. *Lab Chip* 13:2364–2369
- Hao YC, Yuan WZ, Zhang HM, Chang HL (2015) A microgripper with a ratchet self-locking mechanism. In: 2015 28th IEEE international conference on micro electro mechanical systems (MEMS). IEEE, Estoril, pp 1106–1109
- Kim K, Liu X, Zhang Y, Sun Y (2008) Nanonewton force-controlled manipulation of biological cells using a monolithic MEMS microgripper with two-axis force feedback. *J Micromech Microeng* 18:55013
- Lai H, Singh NP (2004) Magnetic field-induced DNA strand breaks in brain cells of the rat. *Environ Health Perspect* 112:687–694
- Liang C, Wang F, Tian Y et al (2017) Development of a high speed and precision wire clamp with both position and force regulations. *Robot Comput Integr Manuf* 44:208–217
- Pham PH, Dao DV, Amaya S et al (2006) Straight movement of micro containers based on ratchet mechanisms and electrostatic comb-drive actuators. *J Micromech Microeng* 16:2532–2538
- Piriyanont B, Fowler AG, Moheimani SOR (2015) Force-controlled MEMS rotary microgripper. *J Microelectromech Syst* 24:1–1
- Ren H, Gerhard E (1997) Design and fabrication of a current-pulse-excited bistable magnetic microactuator. *Sensors Actuators A Phys* 58:259–264
- Tu C, Fanchiang K, Liu C (2005) A 1xN rotary vertical micro mirror for optical switching applications. In: El-Fatraty A (ed) *SPIE 5719, MOEMS and miniaturized systems V*, 14 (April 11, 2005). SPIE, San Jose, pp 14–22
- Valic B, Golzio M, Pavlin M et al (2003) Effect of electric field induced transmembrane potential on spheroidal cells: theory and experiment. *Eur Biophys J: EBJ* 32:519–528
- Volland BE, Heerlein H, Rangelow IW (2002) Electrostatically driven microgripper. *Microelectron Eng* 61–62:1015–1023
- Wang DH, Yang Q, Dong HM (2013) A monolithic compliant piezoelectric-driven microgripper: design, modeling, and testing. *IEEE/ASME Trans Mechatron* 18:138–147
- Wang F, Liang C, Tian Y et al (2015) Design of a piezoelectric-actuated microgripper with a three-stage flexure-based amplification. *IEEE/ASME Trans Mechatron* 20:2205–2213
- Wester BA, Rajaraman S, Ross JD et al (2011) Development and characterization of a packaged mechanically actuated microtweezer system. *Sensors Actuators A Phys* 167:502–511
- Yuan G, Yuan W, Hao Y et al (2015) A microgripper with a post-assembly self-locking mechanism. *Sensors* 15:20140–20151
- Zhang R, Chu J, Wang H, Chen Z (2012) A multipurpose electrothermal microgripper for biological micro-manipulation. *Microsyst Technol* 19:89–97
- Zubir MNM, Shirinzadeh B, Tian Y (2009) A new design of piezoelectric driven compliant-based microgripper for micromanipulation. *Mech Mach Theory* 44:2248–2264



# Traveling-Wave Micropumps

Guohua Liu and Wei Zhang

## Contents

Introduction .....	1018
Principle .....	1019
Structure and Fabrication .....	1023
Piezoelectric Travelling-Wave Micropump .....	1023
Magnetic Travelling-Wave Micropump .....	1026
Performance .....	1028
Piezoelectric Travelling-Wave Micropump with Different Channels .....	1029
Magnetic Travelling-Wave Micropump with Different Parameters .....	1031
Conclusion .....	1034
References .....	1035

## Abstract

In this chapter, a valveless piezoelectric and magnetic micropump driven by travelling wave is presented. The micropump, fabricated with polydimethylsiloxane (PDMS) and polymethylmethacrylate (PMMA), consists primarily of a saw-tooth microchannel, substrates, and two kinds of integrated actuator arrays (piezoelectric bimorph arrays and NdFeB permanent magnetic arrays). The travelling wave beneath the top wall of the elastic microchannel can be induced by the actuator arrays, and the liquid particles are then transported along with the travelling wave in the microchannel. The micropumps are designed and fabricated with different microchannels (the saw-tooth and the straight microchannels). Appropriate geometry of the saw-toothed micro-

---

G. Liu (✉) · W. Zhang

Key Laboratory of Photo-electronic Thin Film Devices and Technology of Tianjin, College of Electronic Information and Optical Engineering, Nankai University, Tianjin, China  
e-mail: [liugh@nankai.edu.cn](mailto:liugh@nankai.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,  
[https://doi.org/10.1007/978-981-10-5945-2\\_29](https://doi.org/10.1007/978-981-10-5945-2_29)

1017



channel was also studied for optimizing the performance of the micropump. Experimental characterization of the micropump has been performed in terms of the frequency response of the flow rate and back pressure. The results demonstrate that this micropump is capable of generating a stable flow rate in microfluidic systems.

---

**Keywords**

Micropumps · Travelling wave · Piezoelectric actuator arrays · Magnetic actuator arrays · Microfluidics · MEMS · PDMS microchannel

---

## Introduction

Micropumps play an essential role in transporting precise volumes of sample fluid through the various components of microelectromechanical systems (MEMS), particularly in the fields of medical implants, drug delivery systems, lab-on-a-chip (LOC) systems, and micro-total analysis systems ( $\mu$ TAS) (Wang et al. 2013; Iverson and Garimella 2008). Different driving modes are used to build the micropump such as piezoelectric, electrostatic, pneumatic, and thermodynamic. Generally, micropumps can be divided into two different types: mechanical and nonmechanical micropumps (Jang and Lee 2000; Zengerle 1995). Nonmechanical pumps usually employ the properties of the working liquid to generate the flow (Darabi et al. 2001). Therefore, the available liquids are limited and are often incompatible with biochemical protocols. On the other hand, mechanical micropumps can transport a variety of liquids and be less sensitive to liquid properties. They usually use moving parts, such as check valves, oscillating diaphragms, turbines, and gears, to deliver a constant fluid volume, which could provide a more versatile pumping solution (Muralt 2005). One type of mechanical micropump is the valveless micropump which has several advantages over other micropumps, such as no moving parts, extended working range, ease of fabrication, and cost effectiveness.

The most typical valveless micropump is the valveless nozzle/diffuser micropump which makes fluid flow by implementing diffuser and nozzle elements to function as passive check valves (Olsson et al. 1996, 1997, 2000). In nozzle/diffuser valveless micropumps, the flow field is generated typically by piezoelectric actuators which are equipped with a pressure chamber for generating a pressure difference between the inlet and outlet. Though piezoelectric micropumps can transport liquid at a high flow rate and with excellent controllability, further miniaturization is difficult because it reduces the generated pressure and because of the difficulty in fabricating a micropressure chamber using conventional MEMS processes. To solve these problems, the valveless piezoelectric travelling-wave micropump has been proposed (Ogawa et al. 2009; Zhang et al. 2011), which is advantageous for miniaturization because of its simple structure. The vibrating wall is a unimorph structure of piezoelectric thin films directly deposited on the ceiling of the microchannel. Additionally, there is no need to prepare a pressure chamber with this configuration. The principle of the piezoelectric travelling-wave micropump is that

the piezoelectric actuator array induces a travelling wave propagated outside the top wall of the microchannel, which makes the fluid particles move forward continuously. Though piezoelectric actuators usually produce high actuation forces and fast mechanical responses, they need high input voltages. Also, most piezoelectric actuators are realized by mounting standard metal-piezo-composites as membrane actuators on top of a plastic microchannel. This requires a reliable mounting and sealing technology, with the danger of wearing at the matching points of metal and polymer. Overall, the two main demerits associated with current piezoelectric travelling micropumps are (1) fabrication complexity and (2) power consumption.

In an attempt to simplify the fabrication process and to address the above demerits, a novel valveless magnetic travelling-wave micropump is presented, whose principle is that a travelling wave beneath the top wall of the elastic microchannel is generated by the interactive forces between two integrated magnetic arrays. As compared to piezoelectric types, it also has the merits of fast response time and high actuation forces but can be manufactured more easily and actuated with a relatively low voltage. In this pump, a modified nozzle/diffuser channel structure called a saw-toothed microchannel is adopted, based on the flow-resistance coefficient equations. Since previous numerical and experimental studies have demonstrated that the geometric design of the nozzle/diffuser elements can significantly affect the performance of the valveless micropump (Ye et al. 2014), systematic characterization of the microchannel has been performed. In addition, a fully printed circuit board (PCB) fabrication technology is presented in our work and the substrate used in this technology is polymethylmethacrylate (PMMA), which allows for easy fabrication and integration with other microfluidic components.

---

## Principle

The travelling-wave micropump is similar to the peristaltic micropump in structure, but they have different working principle and different properties yet. The peristaltic micropump pumps the fluid by alternately extruding and releasing the elastic channel of the pump. As the extrusion position on the elastic channel moves alternately, negative pressure is formed in the tube, and the liquid flows along with the extrusion position movement. The flow rate of the peristaltic micropump depends on the driving frequency and force of the extrusion, and the relationship between the flow rate and driving frequency shows a linear relationship. However, the peristaltic micropump produces an unsteady pulse flow at the outlet during working. On the other hand, travelling-wave micropump depends on the travelling wave propagating along a certain direction in the elastic channel. Because of the viscosity of the fluid, the fluid in the channel moves slightly along the travelling wave propagating direction after each period and obtains a continuous directional flow rate at the outlet of the micropumps. The travelling wave in the elastic channel is excited by several standing waves at a certain vibration frequency, thus the flow rate of travelling-wave micropump is not linear with the driving frequency, which reaches the maximum at a certain frequency, and declines at both sides of the maximum. In the travelling-wave

micropumps, the fluid transport is driven by a high frequency and low amplitude travelling wave, so the flow rate of the travelling-wave micropump is relatively stable. The travelling wave in the elastic channel is generated by several standing waves, and the relationship of travelling wave and standing wave is introduced follows.

Standing waves and travelling waves have many different characteristics in physical properties, but there is a close relationship between these two waves. A travelling wave is the wave of the spatial distribution of a physical quantity that propagates forward in space and time. With difference from the travelling wave, the standing wave does not propagate forward like a travelling wave, in which each point on the axis of the wave has an associated constant amplitude. In the process of vibration, the locations at which the amplitudes is always minimum are called the wave nodes, and the locations where the amplitude can reach maximum are called the antinodes, there is always 1/4 wavelength length difference between the adjacent wave nodes and antinodes. With certain conditions, the travelling wave and standing wave can be transformed into each other. For example, the superposition of two travelling waves, which have the same amplitude, frequency, direction of vibration, and the opposite propagation direction at the same time, could obtain a travelling wave with the same frequency of two standing waves.

Standing waves do not transfer the energy in space; it means the average energy density of the standing wave is equal to zero in one period. The standing wave can be expressed as:

$$\xi(x, t) = A \cos \frac{2\pi}{\lambda} x \cdot \cos \omega t \quad (1)$$

Different from the standing wave, the travelling wave propagates continuously in the medium. Each point of the travelling wave vibrates at the equilibrium position, its vibration and energy spread along the direction of wave propagation, and the average energy density is not zero in one period. The travelling wave can be expressed as:

$$\xi(x, t) = A \cos \left( \omega t - \frac{2\pi}{\lambda} x \right) \quad (2)$$

Two standing waves which met some certain conditions can generate a travelling wave. Assume two standing waves (named as  $\xi_1$ ,  $\xi_2$ ) have the same amplitude  $A$ , wave number  $k$ , angular frequency  $\omega$ , and the same vibration direction, but  $\pi/2$  difference in time and space respectively, then these two standing waves can be expressed as:

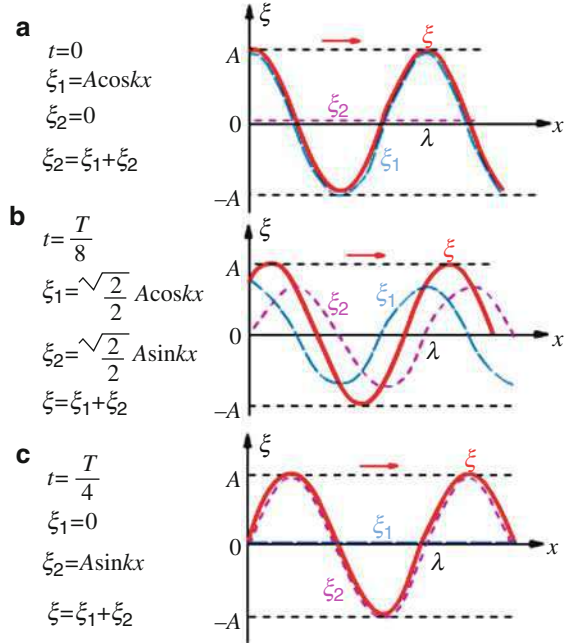
$$\begin{aligned} \xi_1(x, t) &= A \cos kx \cdot \cos \omega t = A \sin(kx + \pi/2) \sin(\omega t + \pi/2) \xi_2(x, t) \\ &= A \sin kx \cdot \sin \omega t \end{aligned} \quad (3)$$

Combining these two standing waves, a travelling wave  $\xi$  can be obtained:

$$\begin{aligned} \xi(x, t) &= \xi_1(x, t) + \xi_2(x, t) \\ &= A \cos kx \cdot \cos \omega t + A \sin kx \cdot \sin \omega t \\ &= A \cos(\omega t - kx) \end{aligned} \quad (4)$$

Thus, two standing waves with phase difference of  $\pi/2$  could generate a travelling wave propagating along the positive direction of axis  $x$ , and the amplitude  $A$ , wave

**Fig. 1** Waveform diagram of the superposition process of two standing waves



number  $k$ , and angular frequency  $\omega$  of the travelling wave are same to the original standing waves. The superposition process of two standing waves at different time can be directly expressed by the waveform curve in Fig. 1.

According to the above expressions, when the following conditions (Eq. 5) are met, two wave standing wave will generate a travelling wave propagating along the positive direction of axis  $x$ .

$$(\varphi_2 + a_2) - (\varphi_1 + a_1) = \pm(2m + 1)\pi \quad (m = 0, 1, 2, \dots) \tag{5}$$

Similarly, four square standing waves with the same amplitude, frequency, direction of vibration, and phase difference of  $\pi/2$  can be expressed as:

$$\begin{aligned} \xi_A(t) &= \frac{4h}{\pi} \left( \sin \frac{2\pi}{\lambda} x \sin \omega t + \frac{1}{3} \sin \frac{2\pi}{\lambda} 3x \sin 3\omega t + \dots \right) \\ \xi_B(t) &= \frac{4h}{\pi} \left( \cos \frac{2\pi}{\lambda} x \cos \omega t + \frac{1}{3} \cos \frac{2\pi}{\lambda} 3x \cos 3\omega t + \dots \right) \\ \xi_C(t) &= \frac{4h}{\pi} \left( \sin \frac{2\pi}{\lambda} x \sin \omega t + \frac{1}{3} \sin \frac{2\pi}{\lambda} 3x \sin 3\omega t + \dots \right) \\ \xi_D(t) &= \frac{4h}{\pi} \left( \cos \frac{2\pi}{\lambda} x \cos \omega t + \frac{1}{3} \cos \frac{2\pi}{\lambda} 3x \cos 3\omega t + \dots \right) \end{aligned} \tag{6}$$

The above four square standing waves can be combined into a square travelling waves.

$$\begin{aligned}
\xi(t) &= \xi_A(t) + \xi_B(t) + \xi_C(t) + \xi_D(t) \\
&= 2\frac{4h}{\pi} \left[ \left( \sin \frac{2\pi}{\lambda}x \sin \omega t + \cos \frac{2\pi}{\lambda}x \cos \omega t \right) + \right. \\
&\quad \left. \frac{1}{3} \left( \sin \frac{2\pi}{\lambda}3x \sin 3\omega t + \cos \frac{2\pi}{\lambda}3x \cos 3\omega t \right) + \dots \right] \\
&= \frac{8h}{\pi} \left[ \cos \left( \frac{2\pi}{\lambda}x - \omega t \right) + \frac{1}{3} \cos 3 \left( \frac{2\pi}{\lambda}x - \omega t \right) + \dots \right] \\
&= \frac{8h}{\pi} \left[ \sin \left( \frac{\pi}{2} + \omega t - \frac{2\pi}{\lambda}x \right) + \frac{1}{3} \sin 3 \left( \frac{\pi}{2} + \omega t - \frac{2\pi}{\lambda}x \right) + \dots \right] \quad (7)
\end{aligned}$$

Thus, the four square standing waves which have difference of  $\pi/2$  in time and space can also generate a square travelling wave; the travelling-wave micropumps are also based on this principle, generating the travelling wave in the elastic microchannels, and realizing the transport of the fluid.

The configuration parameters of the micropump are determined by the calculations and simulations below. The actuator array composes of four actuators, which have  $1/4$  wavelength spacing between adjacent actuators. The voltage signals applied to the actuator array have the same amplitude, frequency, and different phase differed  $\pi/2$  between the driving signals of adjacent actuator. The microchannel is made from PDMS (Sylgard 184, Dow Corning) which comprises the base and the curing agent with volume ratio of 10:1. In this condition, the Young's modulus of PDMS is  $E \approx 750$  kPa, and the Poisson's ratio is  $\nu \approx 0.45$  (Schwarz et al. 2002), and the shear modulus  $G$  is gained from the following equation:

$$G = E/2(1 + \nu) = 258.62\text{kPa} \quad (8)$$

Then, the velocity of transverse wave in the elastic channel is:

$$v = \sqrt{G/\rho} \quad (9)$$

The density of PDMS is about  $1.03$  g/cm<sup>3</sup>. As a result of calculating, the travelling wave velocity in the microchannel is  $v = 15.8$  m/s, and the frequency of driving signals caused travelling wave propagating along positive  $x$  direction is about 1,886 Hz, as the spacing of piezoelectric actuators is 2.1 mm.

The average flow velocity is expressed by the following equation (Ogawa et al. 2009):

$$\bar{u} = \frac{1}{2}a^2\Psi(z) \quad (10)$$

Where  $a$  is the amplitude of the travelling wave,  $\Psi(z)$  is a function of the microchannel's height. So the average flow velocity is proportional to the intensity of the oscillation which is defined as the square of the amplitude of the travelling wave, assumed the microchannel's height  $z$  is a constant.

The travelling wave in the microchannel is a forced resonance caused by the vibrations of the actuator array. In general, the response curve of the resonance is not exactly symmetric about the resonant frequency. But towards the low damping forced resonance, the intensity of oscillation with a driving frequency  $\omega$  is approximated by the Lorentzian function, which is symmetric about the resonant frequency  $\Omega$ . And the theoretical expression of the average flow velocity of the micropump could be expressed as follows.  $\Gamma$  is a parameter dependent on the damping of the oscillation, which specifies the full-width at half-maximum (FWHM). The parameter  $\Gamma$  is also a measure of the sharpness of the resonance, the smaller  $\Gamma$  value responds to the sharper resonant peak.

$$\bar{u} \propto a^2 = I(\omega) \propto \frac{2}{\pi} \cdot \frac{\Gamma}{4(\omega - \Omega)^2 + \Gamma^2} \quad (11)$$

When the vibration transferred to the PDMS, the amplitude gets smaller because of the absorption of elastic PDMS. Figure 2 shows the response of the maximum displacement of the microchannel's inside wall to the vibration amplitude on the microchannel's outside wall, which resulted from ANSYS. It is a conclusion that the curve is almost linear and the transfer efficiency of the amplitude is more than 93%.

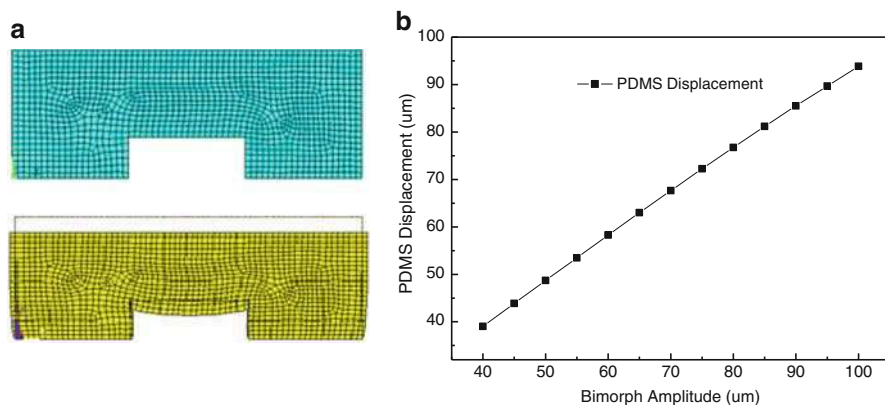
The principle of the travelling-wave micropump is that the alternating voltage is transformed to the mechanical vibration by the actuator array, which induces a travelling wave on the elastic microchannel. The actuator array is attached to the surface of the microchannel, and the vibration of each actuator with different phase is transmitted to the corresponding position of the microchannel. The fluid particles beneath the top wall of the microchannel moves up-and-down followed the vibration of the top wall, and also moves forth and back due to the periodic variation of the local pressure in the microchannel, which caused by the travelling wave propagating along. Such two motions make the trajectory of fluid particles as an elliptic form. After a period of the travelling wave, the fluid particles move forward along the microchannel slightly. And the net flow at the outlet is achieved by repeating such motion.

---

## Structure and Fabrication

### Piezoelectric Travelling-Wave Micropump

The piezoelectric travelling-wave micropump mainly includes three parts: the micropump substrate, microchannel, and piezoelectric actuator array, shown in Fig. 3. The micropump substrate made from PMMA is used to support the microchannel and the piezoelectric actuator array, which consists of inlet, outlet, and actuator array slot (not illustrated in Fig. 3). The actuator array is attached on the microchannel and also connected to the voltage signals of different phase, which induces the travelling wave in the top wall of the microchannel.



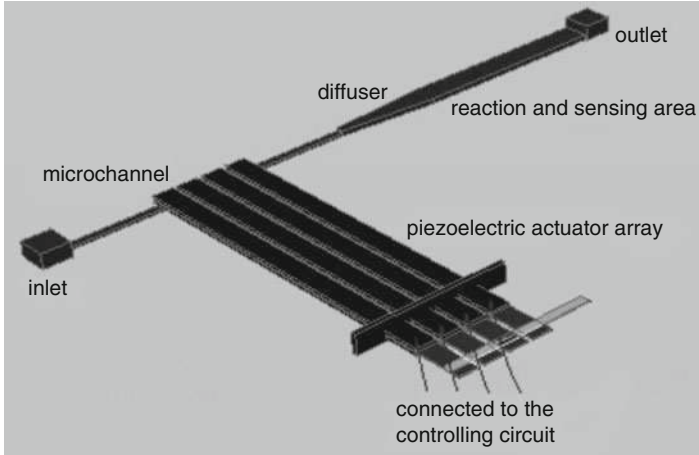
**Fig. 2** Response of the maximum displacement of the microchannel to the vibration amplitude. (a) Simulation diagram of the microchannel with displacement of 0 and 50  $\mu\text{m}$ ; (b) response curve of the maximum displacement of the microchannel's inside wall

The microchannel made by PDMS is an important part of travelling-wave micropump, as its elasticity and configuration affects the performance of the micropump directly. PDMS is a polymer material which exhibits very well optical transparency and biological inertia. It is also convenient to fabricating an integrated micropump by binding with silicon, glass, polystyrene, etc. The different microchannel structures of pump area are designed as the saw-tooth figuration and straight figuration, and the bumps on the top of the microchannel transferred the deformation caused by external pressures to the microchannel inside effectively. Figure 4 shows the configuration of saw-tooth channel. Because of the larger flank area and less pressure loss, the saw-tooth channel should exhibit less fluidic impedance than the straight channel and the traditional diffuse/nozzle structure.

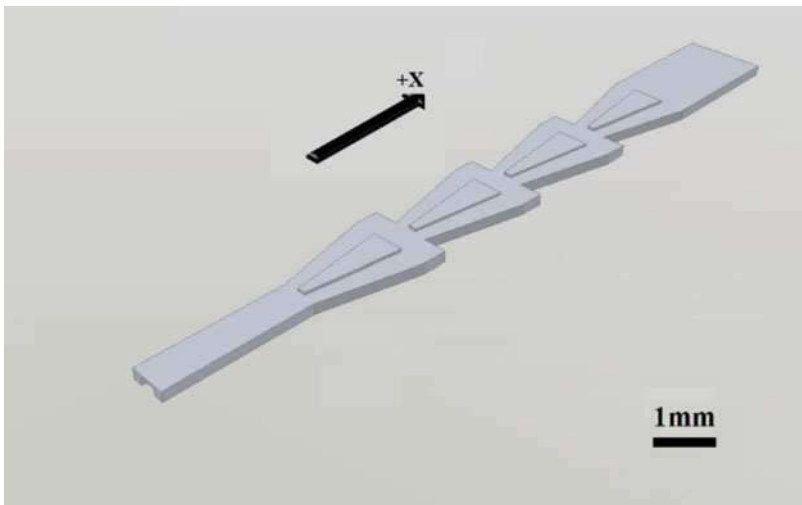
The saw-tooth microchannel is a cascaded configuration of four diffuse elements; the parameter of each diffuse element is that the width of inlet is 200  $\mu\text{m}$ , the width of outlet is 700  $\mu\text{m}$ , the length of diffuse element is 2.1 mm, the height of microchannel is 100  $\mu\text{m}$ , the thickness of side wall and top wall of the microchannel is 300 and 200  $\mu\text{m}$ , and the height of the trapeziform bumps on the top wall is 100  $\mu\text{m}$ .

The molds of microchannel are made from PMMA which has good chemical stability and mechanical performance. Firstly, the PDMS base and curing agent with volume ratio of 10:1 are mixed sufficiently. And the mixture is injected to the molds after degassing in the relative vacuum of  $-1 \text{ kg/cm}^2$  pressure. The molds are put into the incubator of 65  $^{\circ}\text{C}$  about 30 min, made the PDMS precuring. The microchannel which are shown in Fig. 4 are finished after stripping from the molds subsequently and put into a clean container and treated in UV of low pressure mercury lamp for 3 h.

Afterwards, the glass was treated to clean organic residues by piranha solution (98%  $\text{H}_2\text{SO}_4$ :30%  $\text{H}_2\text{O}_2 = 7:3$ ), rinsed extensively in distilled water for several times and dried thoroughly. Then the PDMS microchannel after UV treatment is put



**Fig. 3** Diagram of the travelling-wave micropump



**Fig. 4** Configuration of the saw-tooth microchannel

on the glass substrate and baked at 65 °C for 12 h to bonding the microchannel on the glass, while the additional pressure always lay on the microchannel. And the bonded PDMS microchannel fabricated by this method has enough watertight for the requirement of the travelling-wave micropump.

Finally, the piezoelectric actuator array is fixed on the pump area of microchannel, while the top of the cantilever is contacted on the bumps surface of microchannel. Four piezoelectric bimorphs of actuator array are parallel distributed



**Fig. 5** Photograph of the piezoelectric travelling-wave micropump

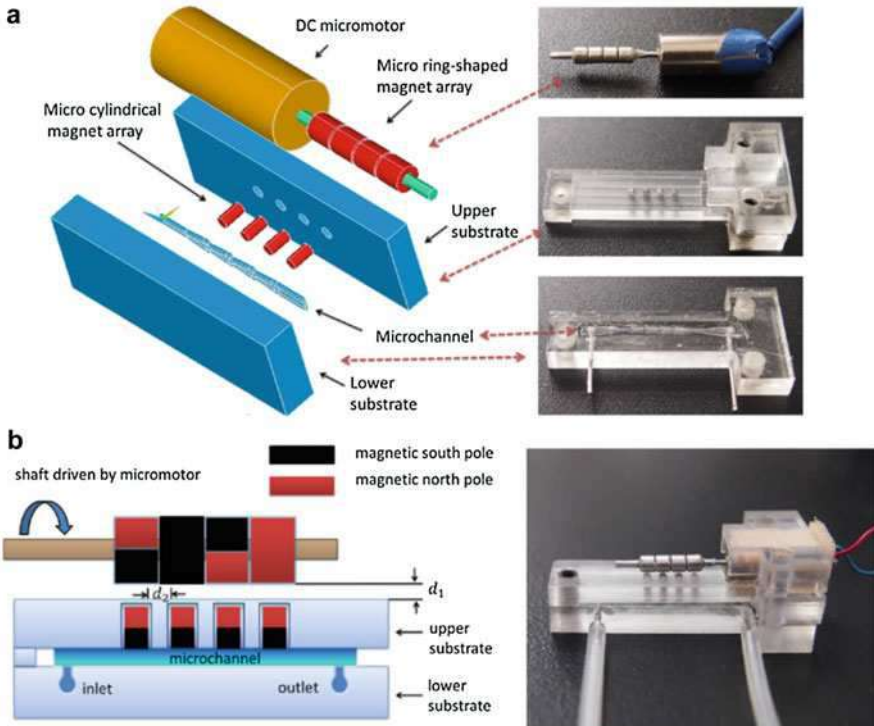


with spacing of 2.1 mm. The photograph of the fabricated piezoelectric travelling-wave micropump is shown in Fig. 5.

### Magnetic Travelling-Wave Micropump

The magnetic travelling-wave micropump employs two actuated micromagnet arrays and a saw-toothed microchannel to provide unidirectional fluid flow. Figure 6a shows an expanded diagram of the structure of the magnetic micropump. From bottom to top, it comprises three layers, namely a lower PMMA substrate containing an outlet and an inlet, a saw-toothed microchannel, and an upper PMMA substrate containing four round cavities. Specifically, the lower substrate was designed to combine with the microchannel, and the four round cavities in the upper substrate were molded and integrated with the cylindrical micromagnet array. A DC electric minimotor (6 mm in diameter and 14 mm in length) was mounted on the top side of the upper substrate. The gap between the ring-shaped magnets and the upper substrate is  $d_1 = 0.8$  mm, and the interval between adjacent cylindrical magnets is  $d_2 = 1$  mm.

Figure 6b presents the cross-sectional view of the magnetic micropump. The saw-toothed microchannel was bonded onto the lower substrate and covered by the upper substrate. For easy assembly, four cylindrical permanent micromagnets (each 2 mm in height and 1 mm in diameter, with  $B = 500$  Gs at 1 mm above its surface) with the same magnetic pole orientation were embedded in the round



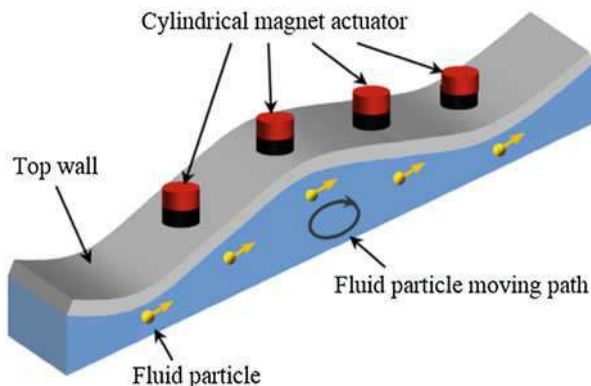
**Fig. 6** Schematic of the magnetic travelling-wave micropump. (a) 3D components schematic; (b) cross-sectional view

cavities of the upper substrate side by side and in contact with the ceiling wall the elastic microchannel. Another four ring-shaped permanent micromagnets (each 2 mm in height, 1 mm in inner diameter, and 2.3 mm in outer diameter, with  $B = 800$  Gs at 1 mm above its surface) were mounted on the rotation shaft of the motor and aligned to the cylindrical magnets one by one. The magnetic pole orientations of the ring-shaped magnets were arranged to successively rotate through  $\pi/2$  rad. The gap between the ring-shaped magnets and the upper substrate is  $d_1 = 0.8$  mm, and the interval between adjacent cylindrical magnets is  $d_2 = 1$  mm. In this configuration, the travelling wave was induced beneath the ceiling wall by the rotating magnetic field created by these two magnetic arrays.

In Fig. 6, the cylindrical magnet actuators were arranged with the same magnetic pole orientation, and the ring-shaped magnets were arranged with the different magnetic pole orientation of rotating  $\pi/2$  rad in turn. During one rotation cycle of the motor, the four cylindrical magnet actuators generate four standing waves with the same amplitude but with  $\pi/2$  rad phase differences.

Therefore, the pumping mechanism is based on the successive actuation of the two micromagnet arrays, and the travelling wave is responsible for the fluid displacement

**Fig. 7** Working mechanism of magnetic travelling-wave micropump



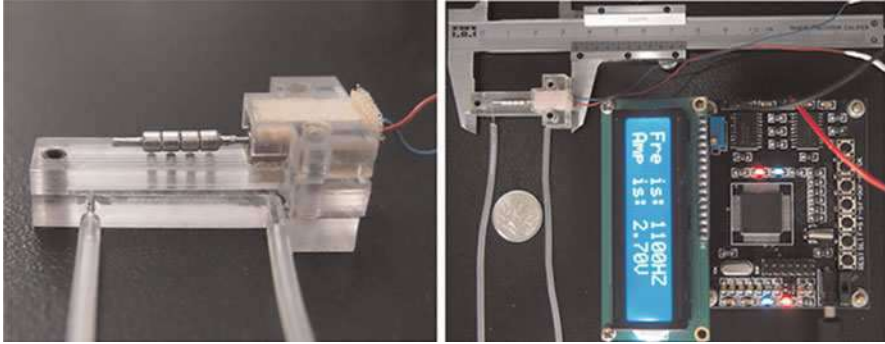
in the microchannel. Figure 7 shows the fluid particles beneath the top wall of the microchannel moving forward with the travelling wave due to its viscosity and the periodic variation of the local pressure. The flow rate of the liquid particles can be controlled by changing the frequency of the generated travelling wave.

The microchannel was fabricated by polydimethylsiloxane (PDMS) molding technology. First, the PDMS base and curing agent were mixed sufficiently with a volume ratio of 11:1. The mixture was injected into the PMMA molds after degassing in a relative vacuum of  $-1 \text{ kg cm}^{-2}$  pressure and cured at  $50 \text{ }^\circ\text{C}$  for about 50 min. Afterward, the microchannel was finished after stripping from the mold and treated by UV from a low-pressure mercury lamp for 3 h in a clean container. Then, the PDMS prepolymer was coated on the lower substrate and cured at  $45 \text{ }^\circ\text{C}$  for about 30 min. Subsequently, the microchannel was bonding to lower substrate at  $45 \text{ }^\circ\text{C}$  for 2 h. The bonded PDMS microchannel fabricated by this method is watertight enough for the requirements of the travelling-wave micropump. Finally, the cylindrical micromagnet array was embedded in the upper substrate, while the minimotor was mounted on the top side of the upper substrate. The upper and lower substrates were fixed by three bolts as shown in Fig. 6b.

The photograph of the micropump and drive circuitry is illustrated in Fig. 8. The motor was driven by programmable pulse-width-modulation (PWM) square waveforms ( $V_{p-p} = 3.7 \text{ V}$ ), which were generated by an MSP430 microcontroller circuitry. Travelling waves with various frequencies ranging from 1 to 1.6 kHz were applied to the tested device, corresponding to motor rotation speeds from 15,000 to 24,000 rpm (approximately).

## Performance

To optimizing the performance of the travelling-wave micropump, the piezoelectric travelling-wave micropump and magnetic travelling-wave micropump should be tested. The optimization of the travelling-wave micropump includes the



**Fig. 8** Photograph of the micropump and drive circuitry

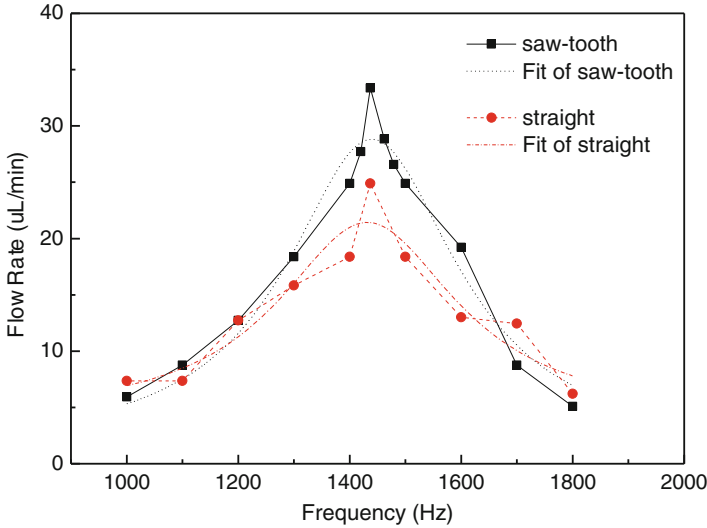
configuration, height of the microchannel, and the diffusion angle of the diffuser element. First, the piezoelectric travelling-wave micropumps with different microchannel are tested to compare the performance of different configuration of microchannels. And then, the performance of magnetic travelling-wave micropumps with different height and diffusion angle of saw-tooth microchannels are also tested, optimizing the parameters of microchannels.

### Piezoelectric Travelling-Wave Micropump with Different Channels

The frequency property of two micropumps with different microchannel configuration is tested in the same conditions firstly. And then, at a certain frequency, the relationship of flow rate to back pressure and the voltage amplitude property are also tested. The back pressure is a pressure at the outlet caused by the output liquid of the micropump, which relates to the tightness, export capability of the pump, and the configuration of the channel. In the experiments, the adjustment of the back pressure is achieved by the height change of the liquid level at the micropump's outlet.

Figure 9 shows the dependency of average flow rate of micropumps with different microchannels on frequency when the voltage amplitude is fixed at 26 V. The average flow rate of micropump with the saw-tooth microchannel reaches the maximum of 33.36  $\mu\text{L}/\text{min}$  when the frequency of driving signal is 1,437 Hz, and that with the straight microchannel is only 24.88  $\mu\text{L}/\text{min}$  at the same frequency. At both sides of the maximum, the average flow rate is declined rapidly. The corresponding frequency of maximal flow rate in experiment (1437 Hz) is different to that in theory (1886 Hz), which is mainly caused by the error of the calculated elastic modulus of PDMS, but the calculated result is still a valuable reference.

The function fitting of the experimental data according to Eq. 4 is also shown in Fig. 9. The resonant central frequency  $\Omega$  and the FWHM  $\Gamma$  of the fitting curve of the saw-tooth microchannel are 1439.7 and 374.4, respectively, and these two parameters of the straight microchannel are 1431.5 and 394.8, respectively. It is concluded



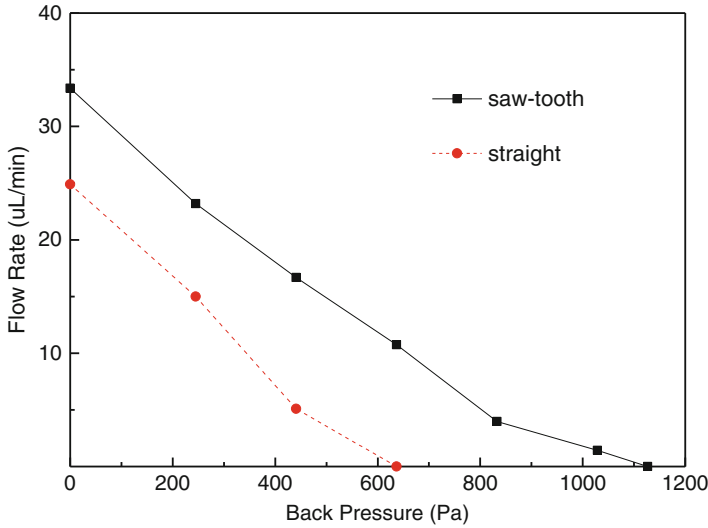
**Fig. 9** Comparison between saw-tooth microchannel and straight microchannel

that the resonant frequencies of these two microchannels are basically the same, and the FWHM of the saw-tooth microchannel is smaller than that of the straight microchannel, indicating the better frequency selectivity.

From Fig. 9, it can be concluded that the saw-tooth microchannel can produce larger flow rate, and the frequency of maximum flow rate is the same for the two kinds of microchannels as the elastic modulus of the microchannel is unchanged. The same frequency of maximum flow rate infers that the frequency property is only dependent on the material property of the microchannels and independent on the structure.

Besides, the maximum back pressure of saw-tooth microchannel reaches 1.13 kPa, which is much larger than that of the straight microchannel (0.64 kPa). The curve of the flow rate to the back pressure is shown in Fig. 10. The curve shows the inverse proportion between average flow rate and back pressure of the micropump at driving frequency of 1,437 Hz. With the same flow rate, the micropump with saw-tooth microchannel could achieve larger back pressure than that with straight microchannel, because of the larger fluidic impedance along the reverse direction.

From Eq. 10, the average flow rate is proportional to the square of the voltage amplitude of driving signal. The voltage amplitude is picked as 20 V, 22 V, 24 V, 26 V, 28 V, respectively, and the average flow rate of the micropump is detected at frequency of 1,437 Hz. The fitting curve which is similar to the quadratic function and in accordance with the theoretical equation exactly is shown in Fig. 11.



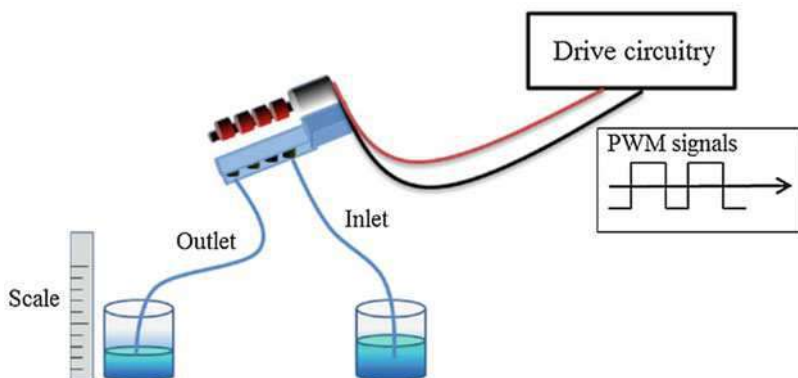
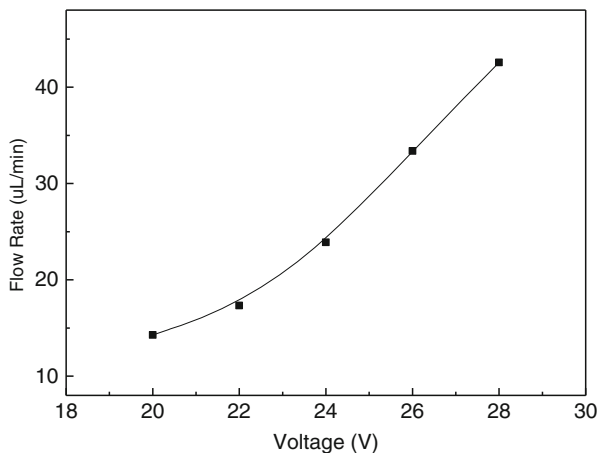
**Fig. 10** Curve of flow rate to back pressure

## Magnetic Travelling-Wave Micropump with Different Parameters

During testing, deionized water was used as a sample liquid for characterization of the micropump. Figure 12 shows the schematic of the micropump experimental setup and the averaged flow rate was tested (calculated by measuring the moving speed of the water front in the outlet tubing). In back pressure-related micropump tests, a pressure difference was obtained by placing the inlet and outlet tubes at a different height  $H$  and could be calculated as  $\Delta\rho = \rho gH$ , where  $\rho$  is the density of water ( $10^3 \text{ kg m}^{-3}$ ) and  $g$  is gravitational acceleration.

The valveless magnetic travelling-wave micropump's characteristics by varying the driving frequency and microchannel geometry is studied systematically. Transient flow measurements on the different micropumps show that they all have flow-directing properties with the diverging-wall direction as the positive direction. First, the diffusion angle of the diffuser element is fixed at  $6.5^\circ$  and then studied the magnetic micropump characteristics by increasing heights of the microchannel to find the optimal height, which is  $350 \mu\text{m}$ . The frequency response of the average micropump flow rates at varied heights of the microchannel is shown in Fig. 13a. It can be seen that the tested average pumping flow rates show a consistent trend with the driving frequency. Specifically, the flow rate initially increases with the frequency until it reaches a peak value at a certain frequency. The average flow rates decline steadily on both sides of the maximum, in correspondence with the principle of the travelling-wave micropump, which implies that only a certain frequency of the driving signal, depending on the mechanical properties of the PDMS materials and

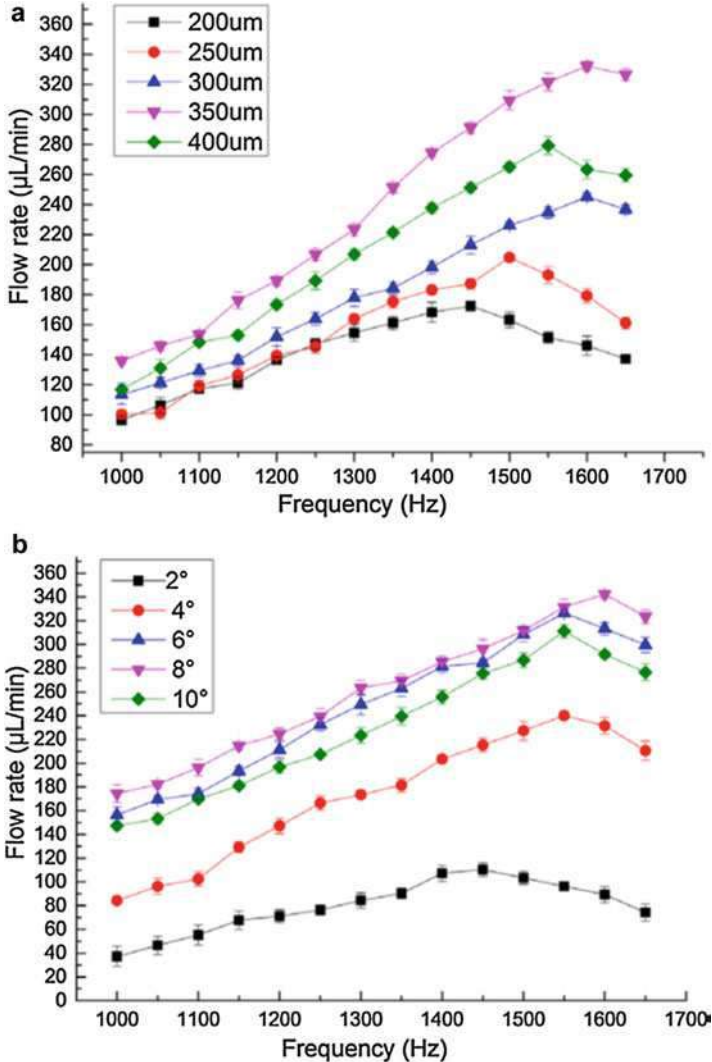
**Fig. 11** Voltage amplitude property of travelling-wave micropump



**Fig. 12** Schematic of micropump experimental setup

the coupling between the channel and the flow, can efficiently induce a travelling wave in the microchannel.

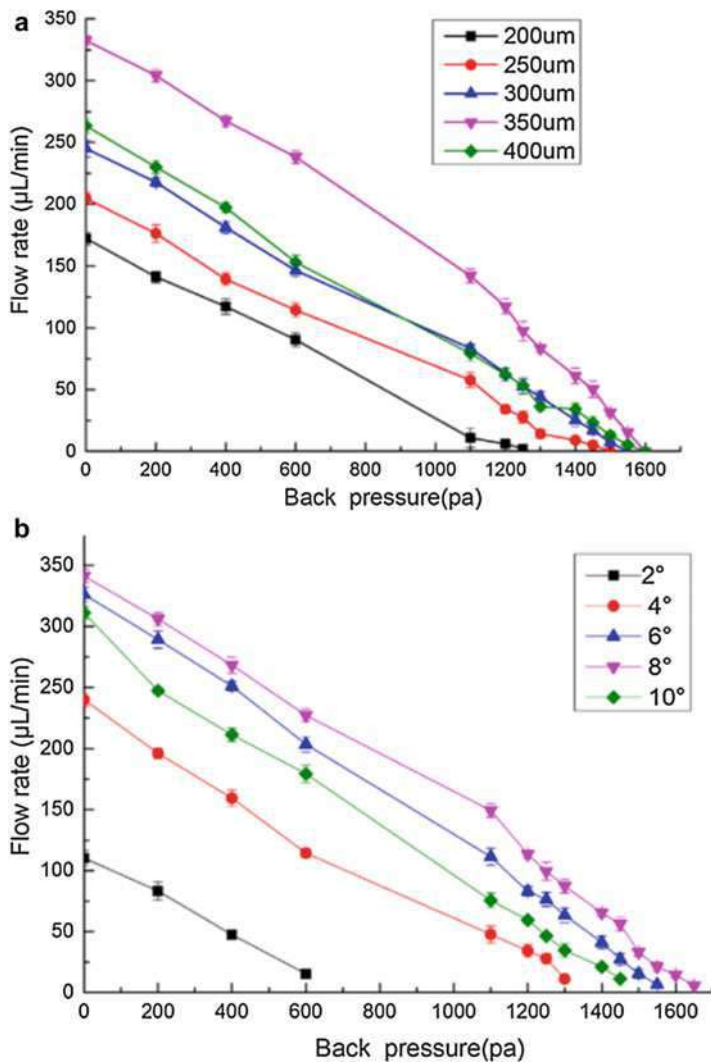
Secondly, the height of the microchannel is fixed at  $350\ \mu\text{m}$ , and then the magnetic travelling-wave micropump characteristics is tested by increasing the diffusion angle of the diffuser element to find the optimal angle. Figure 13b clearly showed that average pumping flow rate increases with increasing diffusion angle of the diffuser element until it reaches an optimal angle of approximately  $8^\circ$ , which is in good agreement with previous studies. The behavior is proposed to be due to the flow separation as a result of the diffuser/nozzle elements, and further flow visualization investigation is required.



**Fig. 13** Curve of average flow rates. (a) Flow rates at varied heights of the microchannel. (b) Flow rates at varied diffusion angles of the diffuser element

Figure 14a and b shows the relationship between the pumping flow rates and the back pressures at different driving frequencies. It can be observed that the pumping flow rates have an almost linear relationship with the back pressures. Accordingly, the magnetic micropump provides the maximum flow rate and back pressure ( $342.4 \mu\text{L min}^{-1}$  and  $1.67 \text{ kPa}$ ) with the height of the microchannel ( $350 \mu\text{m}$ ) and the diffusion angle of the diffuser element ( $8^\circ$ ).





**Fig. 14** Curve of flow rates against back pressures. (a) Relationship between flow rates and back pressures at varied heights of the microchannel. (b) Relationship between flow rates and back pressures at varied diffusion angles of the diffuser element

## Conclusion

A piezoelectric travelling-wave micropump combined with the saw-tooth microchannel is presented, and compared the performance of this micropump with straight microchannel. The inducing mechanism of the travelling wave is discussed and simulated. The performance of the micropump such as average flow rate and back

pressure exhibits improvement in contrast to the straight microchannel micropump. And the micropump is applied in low voltage, which is benefit to the miniaturization and the integration.

An efficient valveless magnetic travelling-wave micropump is also presented here, which is actuated by the interactive forces between two integrated magnetic arrays. This results in a simple pump structure and permits easy system integration. With flow-rate regulation realized by the rotation speed of the motor, the device exhibits excellent performance, demonstrating its ability to drive fluid continuously and efficiently by inducing a travelling wave in the microchannel. This micropump was fabricated using simple molding technology and bonding techniques. In particular, the micropump device under varying microchannel parameters is characterized systematically, including the height of the microchannel and the diffusion angle. Experimental results demonstrate that this micropump is able to efficiently generate a maximum average flow rate of  $342.4 \mu\text{L min}^{-1}$  and reliably operate against high back pressures up to 1.67 kPa. In addition, the performance of a magnetic micropump can even be further improved by optimizing the assembling technology. It is reasonable to expect that a magnetic micropump with this simple structure design and reliable operation can be readily integrated with various actuation methods and potentially applied in various lab-on-a-chip systems, such as drug delivery systems.

---

## References

- Darabi J, Ohadi MM, DeVoe D (2001) An electrohydrodynamic polarization micropump for electronic cooling. *J Microelectromech Syst* 10:98–106
- Iverson BD, Garimella SV (2008) Recent advances in microscale pumping technologies: a review and evaluation. *Microfluid Nanofluid* 5:145–174
- Jang J, Lee SS (2000) Theoretical and experimental study of MHD (magnetohydrodynamic) micropump. *Sensors Actuators* 80:84–89
- Murali P (2005) Piezoelectric micromachined ultrasonic transducers based on PZT thin films. *IEEE Trans Ultrason Ferroelectr Freq Control* 52:2276–2288
- Ogawa J, Kanno I, Kotera H et al (2009) Development of liquid pumping devices using vibrating microchannel walls. *Sensors Actuators* 152:211–218
- Olsson A, Stemme G, Stemme E (1996) Diffuser-element design investigation for valveless pumps. *Sensors Actuators* 57:137–143
- Olsson A, Enoksson P, Stemme G et al (1997) Micromachined flat-walled valveless diffuser pumps. *J Microelectromech Syst* 6:161–166
- Olsson A, Stemme G, Stemme E (2000) Numerical and experimental studies of flat-walled diffuser elements for valveless micropumps. *Sensors Actuators* 84:165–175
- Schwarz US, Balaban NQ, Rivelino D et al (2002) Calculation of forces at focal adhesions from elastic substrate data: the effect of localized force and the need for regularization. *Biophys J* 83:1380–1394
- Wang C, Zhang W, Liu RP et al (2013) Time-division self-referencing multichannel spectral SPR sensor without mechanical scanning. *Electron Lett* 49:493–495
- Ye WX, Zhang W, Wang C et al (2014) Travelling wave magnetic valveless micropump driven by rotating integrated magnetic arrays. *Micro Nano Lett* 9:232–234
- Zengerle R (1995) A bidirectional silicon micropump. *Sensors Actuators* 50:81–86
- Zhang W, Wang C, Yue Z et al (2011) Travelling-wave piezoelectric micropump with low resistance microchannel. *Electron Lett* 47:1065–1066

---

**Part VI**  
**RF MEMS**



# RF MEMS Switch

Li-Feng Wang, Qing-An Huang, and Lei Han

## Contents

Overview .....	1040
Basic Principles of RF MEMS Switches .....	1041
Basic RF Principles .....	1041
Basic Mechanical Principles .....	1049
Vertical Actuating RF MEMS Switches .....	1051
Shunt Capacitive RF MEMS Switch on GaAs Substrate .....	1051
Shunt Contact RF MEMS Switch on LCP Substrate .....	1056
Lateral Actuating RF MEMS Switches Based on SOG Process .....	1059
Fabrication .....	1060
Design and Analysis of SOG Process Based CPW Lines .....	1061
Design A: Push-Pull Type Switch .....	1064
Design B: Three-State Switch by Rhombic Beam Deformation .....	1065
Mechanical Design .....	1067
Measurement Results and Discussion .....	1068
Conclusions and Further Reading .....	1074
References .....	1075

## Abstract

RF MEMS switch is a new type of RF component developed by MEMS technology. Like the macro switches and relays, RF MEMS switches use a mechanical way to control the signal on and off. The difference is that, RF MEMS switches have very small volume and are used to process RF or microwave

---

L.-F. Wang (✉) · L. Han  
Southeast University, Nanjing, Jiangsu, China  
e-mail: [wanglifeng@seu.edu.cn](mailto:wanglifeng@seu.edu.cn); [hanlei@seu.edu.cn](mailto:hanlei@seu.edu.cn)

Q.-A. Huang  
Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, Jiangsu, China  
e-mail: [hqa@seu.edu.cn](mailto:hqa@seu.edu.cn)

signals. Compared with the traditional solid-state semiconductor RF switches, RF MEMS switches have the advantages of low insertion loss, high isolation, low power consumption, and high linearity. Hence, RF MEMS switches have a wide range of application prospects in many fields, such as radar, satellite, base station, and portable radio communication equipment.

In this chapter, the configuration principles of RF MEMS switches is first introduced, including four basic EM models and two basic movement styles. According to the two basic movement styles, the vertical movement and the lateral movement, several RF MEMS switches with different features are shown. Two different vertical actuating membrane bridge RF MEMS switches are obtained based on two different substrates. The first one has the advantage of compatibility with GaAs MMIC process, and the second one has the feature of flexibility. Furthermore, another two lateral actuating RF MEMS switches are demonstrated based on the SOG process. One is the push-pull type switch controlled by only one actuation signal, and the other is the three-state switch actuated by rhombic structures. By designing lateral actuating structures, some performances of the RF MEMS switches are improved.

---

**Keywords**

RF MEMS switch · Vertical actuating · GaAs substrate · LCP substrate · Shunt capacitive switch · Shunt contact switch · Lateral actuating · SOG process · Push-pull type switch · Three-state switch · Deep-off state

---

**Overview**

RF MEMS (micro-electro-mechanical system) switch is a new type of RF component developed by MEMS technology. Like the macro switches and relays, RF MEMS switches use a mechanical way to control the signal on and off. The difference is that, RF MEMS switches have very small volume and are used to process RF or microwave signals. Compared with the traditional solid-state semiconductor radio frequency switches, RF MEMS switches have the advantages of low insertion loss, high isolation, low power consumption, and high linearity. Hence, RF MEMS switches have a wide range of application prospects in many fields, such as radar, satellite, base station, and portable radio communication equipment.

K. E. Petersen, founder of the MEMS field, made the first micromechanical switch by micromachining technology in 1978 (Petersen 1978). Although Petersen's switch could only deal with low frequency signals, he provides a good idea for the use of MEMS technology to fabricate RF switches. Goldsmith of TI successfully developed a kind of membrane bridge RF MEMS switch using micromachining technology (Goldsmith et al. 1995), for the first time. The contact resistance of this switch is only 1.5–2.5  $\Omega$ , and the off state parasitic capacitance is only 35 fF. Goldsmith also used his MEMS switch to produce a 4-bit delay phase shifter. After 2000, more and more companies and research institutions saw the potential of MEMS RF switches, and joined to the research of RF MEMS switch. Those

companies and research institutions include TI, Michigan University, Berkeley UC University, MIT University, ADI, Raytheon, Omron, Motorola, Samsung, etc. Because numerous research institutions joined in, a large number of different structure, principle, and driving mode RF MEMS switches were reported (Goldsmith et al. 1996; Pacheco et al. 1998; Chan et al. 2003; Schauwecker et al. 2004; Adam et al. 2009; Sedaghat-Pisheh and Rebeiz 2010; Verger et al. 2010; Persano et al. 2010).

In this chapter, we first introduce the configuration principle of different kinds of RF MEMS switches, including RF principles and mechanical principles. Then, based on two basic movement styles, several MEMS RF switches with different characteristics are demonstrated, respectively. In the end, some conclusions and further reading are given.

---

## Basic Principles of RF MEMS Switches

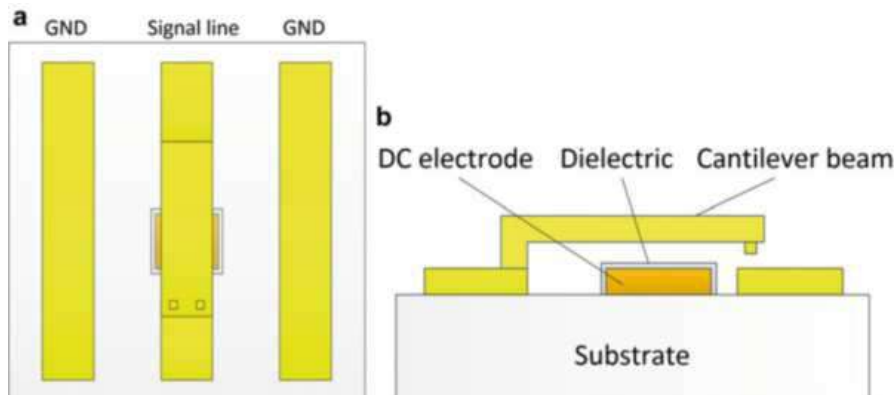
Unlike with traditional solid state semiconductor RF switches, RF MEMS switches consist of both the RF signal processing part and the mechanical actuating part. In this section, the basic RF principle and the basic mechanical principle of RF MEMS switches will be introduced, respectively.

### Basic RF Principles

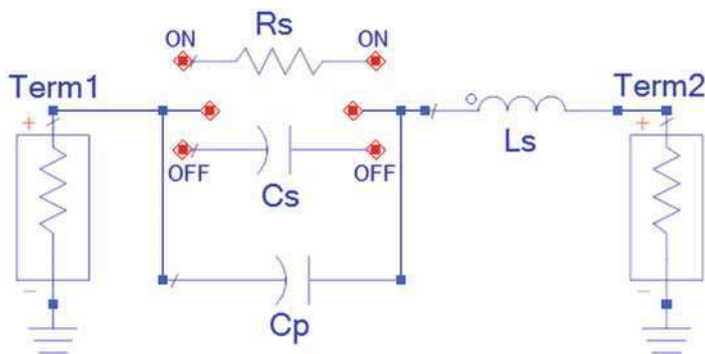
After adding a micromovable beam to the RF transmission line, the on and off state of the RF transmission line can be controlled by the movable beam. There are mainly two ways of controlling the on and off state of the transmission line, one is the capacitance variation generated by the movable beam, and the other is the resistance variation (or called the contact condition variation) generated by the movable beam. Also, the way that the movable beam is added to the RF transmission line is divided into two types: one is the series type, and the other is the parallel type. Therefore, we got the basic four types of RF principles of RF MEMS switches: the series contact type, the shunt capacitive type, the series capacitive type, and the shunt contact type. Below, the four basic RF principles of RF MEMS switches will be explained, respectively.

#### The Series Contact Type

The schematic of a series contact type RF MEMS switch is shown in Fig. 1a, which shows that a cantilever is serially inserted between the RF input and output CPW line. Figure 1b shows the section of the series contact type RF MEMS switch. When no voltage is applied between the DC bias electrode and the cantilever, the input and output signal line is disconnected. Due to the disconnection, the input RF/MW signal is reflected, and the switch is at its off-state. When a large enough voltage is applied between the bias electrode and the cantilever, the cantilever bends and finally touches the transmission line. The connection between the cantilever and the signal line makes the input RF/MW signal successfully transmit to the output port, and the switch is at its on-state.



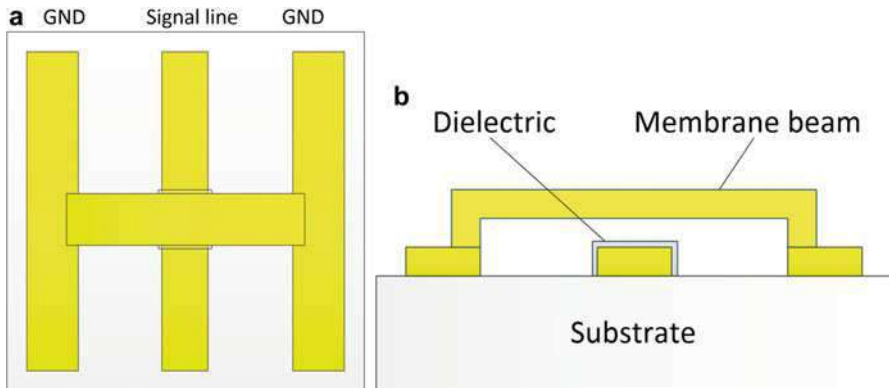
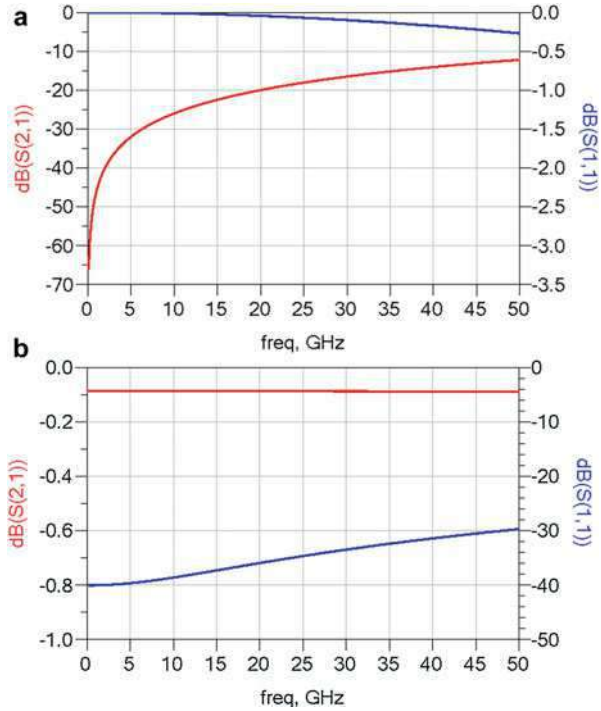
**Fig. 1** The schematic of a series contact type RF MEMS switch. (a) Top view, (b) section view



**Fig. 2** The lumped circuit model of the series contact switch

Figure 2 is the lumped circuit model of the series contact switch. The off-state isolation of the series contact switch is decided by the parasitic capacitance between the cantilever beam and the output signal line. This parasitic capacitance consists of the series parasitic capacitance  $C_s$  between contact area and signal line and the parallel parasitic capacitance  $C_p$  between input and output signal lines.  $L_s$  is the inductance of the cantilever beam, and  $R_s$  is the contact resistance when the beam contacts the signal line. Figure 3 is a typical isolation and insertion loss curve of a series contact switch, which are simulated by its lumped circuit model in ADS. We can see that the isolation of the series contact switch decreases with the increase of the signal frequency. The switch has excellent isolation performance at low frequency. However, when the signal frequency is higher than 20 GHz, the isolation of the switch will be less than 20 dB. So the series contact switch is more suitable for middle and low frequency band.

**Fig. 3** (a) Isolation and (b) insertion loss, curve of a series contact switch simulated by lumped circuit model, where  $C_s = 5\text{fF}$ ,  $C_p = 3\text{fF}$ ,  $R_s = 1\ \Omega$



**Fig. 4** The schematic of a shunt capacitive type RF MEMS switch. (a) Top view, (b) section view

**The Shunt Capacitive Type**

Figure 4 is the schematic of a shunt capacitive type RF MEMS switch. A membrane bridge connects the two ground lines of the CPW line. Figure 4b is the section view of the shunt capacitive switch. Beneath the membrane bridge, there is a dielectric layer on the signal line. When no voltage is applied between the signal line and the membrane bridge, the capacitance between the membrane bridge and the signal line



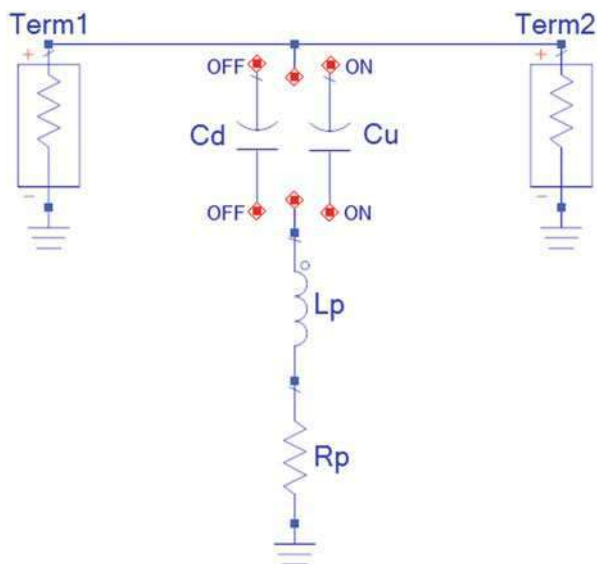
is very small. The input RF/MW signal can successfully transmit to the output port, and the switch is at its on-state. When an adequate voltage is applied between the signal line and the membrane bridge, the membrane bridge is pulled down by the electrostatic force and finally contacts with the dielectric layer. A relatively large capacitance is now formed between the signal line and the ground line, the input RF/MW signal at specific frequency band can be reflected, and the switch is at its off-state.

The lumped circuit model of the shunt capacitive switch is shown in Fig. 5. The off-state isolation of the shunt capacitive type switch is determined by the off-state capacitance  $C_d$ , the inductance  $L_p$ , and resistance  $R_p$  of the membrane. Figure 6 is the simulation result of a typical isolation and insertion loss curve of a shunt capacitive switch. It can be seen that the shunt capacitance switch has very poor isolation at low frequency band especially near DC, so it is not suitable for low frequency band applications. At a certain frequency the shunt capacitance switch shows a very high isolation. The frequency of the maximum isolation can be adjusted through the off-state capacitance  $C_d$ . Therefore, shunt capacitive switch is more suitable for medium and high frequency situations.

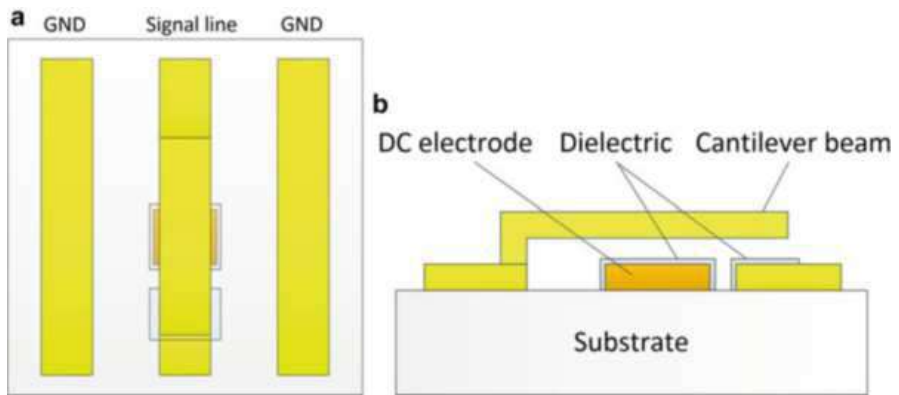
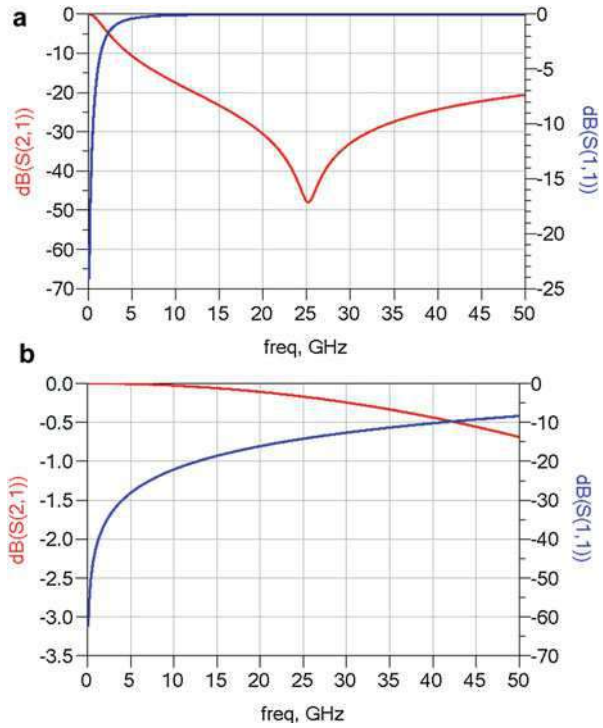
### The Series Capacitive Type

The schematic of the series capacitive type RF MEMS switch is shown in Fig. 7. Like the series contact type switch, a cantilever is serially inserted between the RF input and output transmission line in the series capacitive switch. So the off-state of the series capacitive switch is similar with the series contact switch. The difference is that, the contact area of the series capacitive switch is covered with a dielectric layer.

**Fig. 5** The lumped circuit model of the shunt capacitive switch

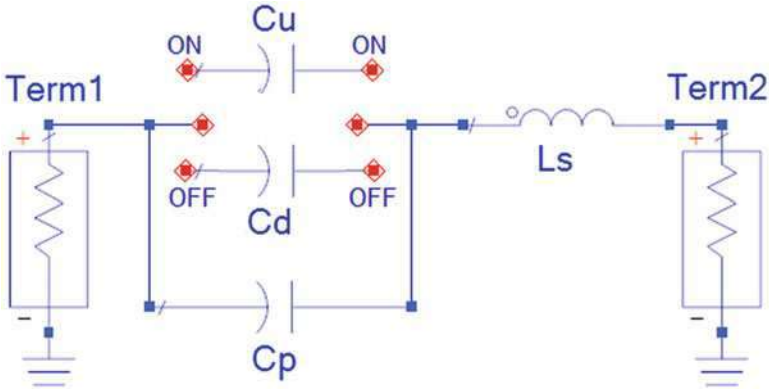


**Fig. 6** (a) Isolation and (b) insertion loss, curve of a shunt capacitive switch simulated by lumped circuit model, where  $C_d = 4$  pF,  $R_p = 0.1 \Omega$ ,  $L_p = 10$  pH,  $C_u = 50$  fF



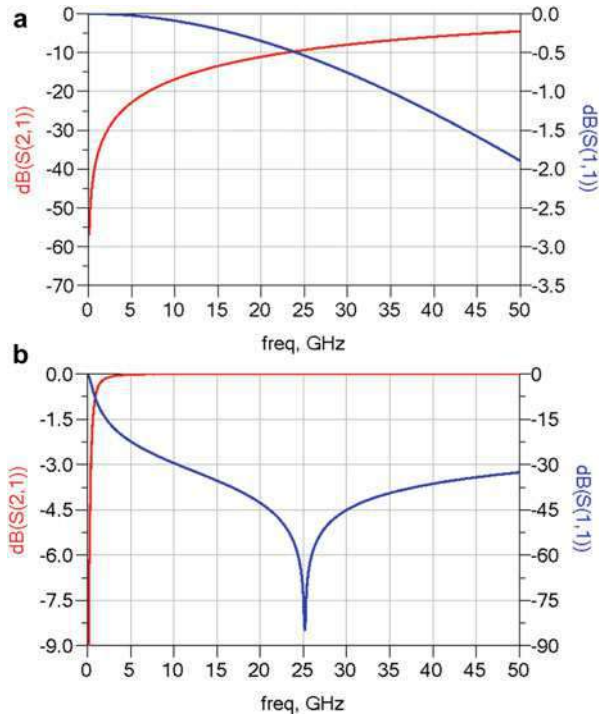
**Fig. 7** The schematic of a series capacitive type RF MEMS switch. (a) Top view, (b) section view

Hence, when the cantilever is pulled down and contacts the dielectric layer, a large capacitance is formed between the cantilever and the signal line. The input RF/MW signal is coupled to the output port through this relatively large capacitor, and the switch is at its on-state.

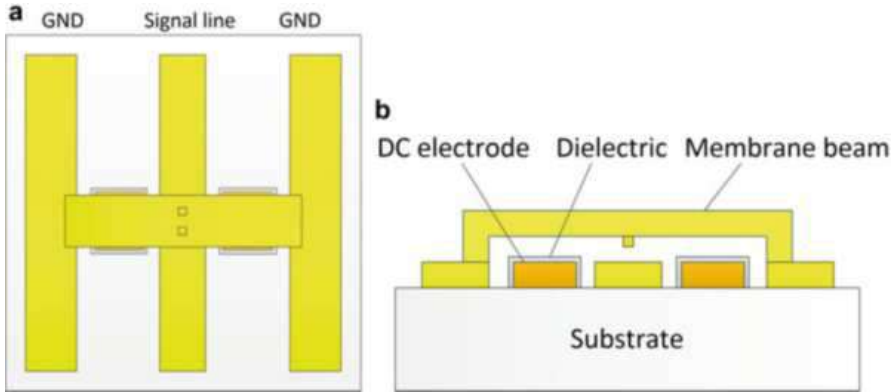


**Fig. 8** The lumped circuit model of the series capacitive switch

**Fig. 9** (a) Isolation and (b) insertion loss, curve of a series capacitive switch simulated by lumped circuit model, where  $C_u = 20\text{fF}$ ,  $C_p = 3\text{fF}$ ,  $L_s = 10\text{pH}$ ,  $C_d = 4\text{pF}$



The lumped circuit model of the series capacitive switch is shown in Fig. 8. Figure 9 shows the isolation and insertion loss curve of a series capacitive switch simulated by the lumped circuit model. The RF characteristic of the off-state of the series capacitive switch is similar with that of the series contact switch. However, the



**Fig. 10** The schematic of a shunt contact type RF MEMS switch. (a) Top view, (b) section view

much large parasitic capacitance  $C_u$  in the series capacitive switch leads to a much large insertion loss. At the on-state, the RF/MW signal is transmitted by a capacitive coupling between the cantilever beam and the signal line. So the on-state insertion loss is very high in the low frequency range, especially near DC. Therefore, the series capacitive switch has large insertion loss at low frequency, at the same time, has poor isolation the high frequency band.

### The Shunt Contact Type Switch

Figure 10 is the schematic of the shunt contact type switch. As the shunt capacitive switch, a membrane bridge connects the two ground lines of the CPW line in the shunt contact switch. So the on-state of the series capacitive switch is very similar with the series contact switch. When the membrane bridge is pulled down, the membrane will contact the signal line and leads a short circuit between them. Due to this short circuit, the input RF/MW signal can be reflected, and the switch is at its off-state.

Figure 11 is the lumped circuit model of the shunt contact switch. The off-state isolation of the shunt contact switch is determined by the resistance and inductance value of the short circuit after the membrane bridge contacts the signal line. In which,  $R_s$  is the contact resistance of the membrane and the signal line, and  $L_p$  is the inductance of the membrane bridge. A typical off-state isolation curve is shown in Fig. 12. The isolation of the shunt contact switch decreases with the increase of the signal frequency. However, even at low frequencies, the isolation of the shunt contact switch is not very good.

The RF characterizations of most RF MEMS switches could be described using the four models above. By changing the parameters of the components in the lumped circuit models, RF MEMS switches with different RF performance requirements could be designed.

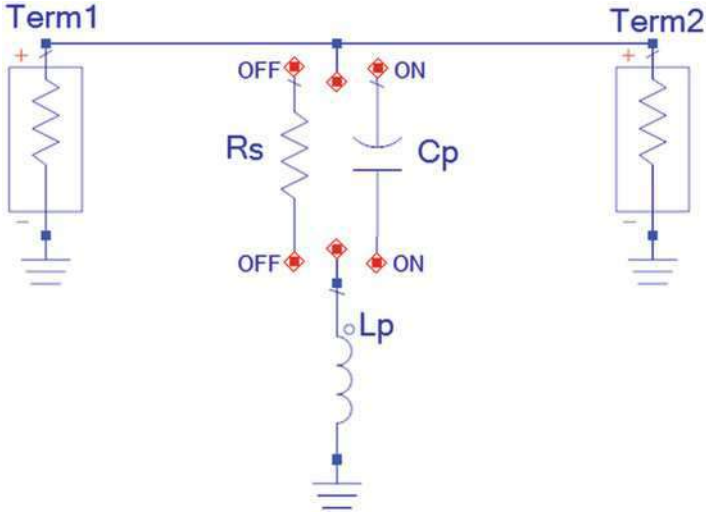
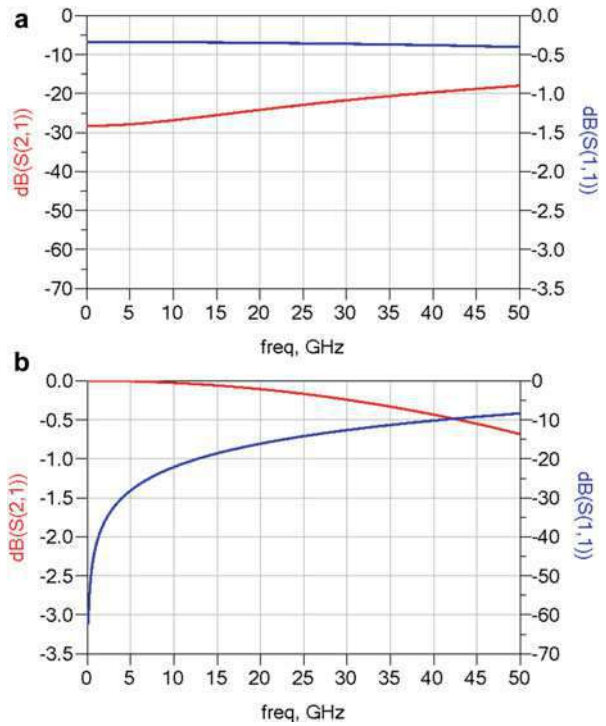


Fig. 11 The lumped circuit model of the shunt contact switch

Fig. 12 (a) Isolation and (b) insertion loss, curve of a shunt contact switch simulated by lumped circuit model, where  $R_s = 1 \Omega$ ,  $L_p = 10 \text{ pH}$ ,  $C_p = 50 \text{ fF}$



**Table 1** The performance comparing of different actuation mechanisms

Actuation mechanism	Switching time	Chip size	Power consumption
Electrostatic	Fast	Small	Low
Thermal	Slow	Large	High
Electromagnetic	Medium	Medium	High
Piezoelectric	Medium	Medium	Low

## Basic Mechanical Principles

Different with traditional IC circuits, MEMS devices usually include movable parts. The actuation mechanism and the movement type are the two most important factors of the movable structures.

### Actuation Mechanism

The actuation mechanisms of driving micromechanical structures are many and varied, the most common of which are electrostatic, thermal, electromagnetic, or piezoelectric (Cho et al. 2005; Lee et al. 2005; Daneshmand et al. 2009; Solazzi et al. 2010). The performance comparing of different actuation mechanisms is listed in Table 1. Compared with other actuation types, electrostatic actuation has relatively short switching time, small chip size, and very low power consumption. Due to these advantages, electrostatic actuation becomes the most popular technique used in RF MEMS switches. As a result, the RF MEMS switches discussed in this chapter will be based on electrostatic driven.

The most important parameter of the electrostatic driven structure is the actuation voltage. It is determined by two factors: one is the force generated by the driving structure, and the other is elastic coefficient of the movable structure.

Figure 13 shows a schematic of a parallel plate electrostatic actuator. The electrostatic driving force generated by two parallel plates is given by,

$$F_e = -\frac{\epsilon_0 A V^2}{2g^2} \quad (1)$$

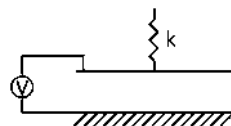
where  $V$  is the voltage applied between the two electrodes,  $A$  is the area of the electrode,  $\epsilon_0$  is the permittivity of free space, and  $g$  is the initial gap distance.

When the displacement of the upper electrode reaches one third of the initial gap distance, the force equilibrium is broken, and the upper electrode collapses. The DC voltage at which the upper electrode start to collapses is defined as the threshold voltage  $V_{pi}$  of the actuator (Yao et al. 1999).

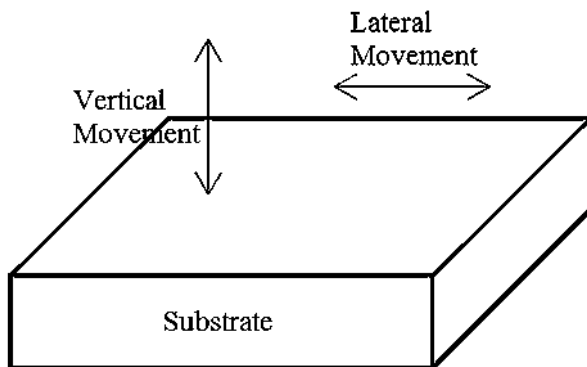
$$V_{pi} = \sqrt{\frac{8kg^3}{27\epsilon_0 A}} \quad (2)$$

However, the elastic coefficient  $k$  is determined by the shape and material parameters of the movable structure and needs to be analyzed separately.

**Fig. 13** A schematic of a parallel plate electrostatic actuator



**Fig. 14** Two basic movement styles



### Movement Style

Roughly, the movement style of movable structures can be divided into two categories, the vertical movement style and the lateral movement style. As shown in Fig. 14, the vertical movement means that the motion path of the movable structure is perpendicular to the surface of the substrate, and the lateral movement means that the motion path is parallel to the surface of the substrate.

Vertical movement style is the most commonly used movement style in RF MEMS switches. The reason could be that the RF/MW transmission lines fabricated by surface micromachining process are easier to produce vertical movements. The movable structures of those vertical RF MEMS switches are usually metallic membranes or cantilevers (Goldsmith et al. 1995). The four basic switch configurations in above section are all based on the vertical movement style. The advantages of vertical movement are that the actuating structure is very simple, and the switch size can be small. However, vertical movement structures can only do simple one-dimensional movements of upward and downward.

So far, lateral movement style is rarely used in RF MEMS switches. Because it requires large depth to width ratio structures, which are not compatible with RF/MW transmission lines. Laterally actuated RF MEMS switches can be fabricated by several different micromachining processes, such as high-aspect-ratio electroplating nickel process (Girbau et al. 2007), silicon-on-insulator substrate process based on the deep-reactive-ion etching processing (Tang et al. 2005), and silicon on glass (SOG) process (Park et al. 2009). Compared with vertically actuated devices, moving laterally provides the benefit of flexibility of in-plane complicated actuators design. These complicated actuators can be used to improve some performances of RF MEMS switches.

## Vertical Actuating RF MEMS Switches

In this section, two vertical RF MEMS switches based on different kinds of substrates are introduced. One is a shunt capacitive RF MEMS switch on GaAs substrate, and the other is a DC-contact shunt RF MEMS switch on LCP substrate. Their lumped circuit models could be found in section “[Basic RF Principles](#).” Here, we mainly focus on the fabrication processes, the mechanical designs, and measurements.

### Shunt Capacitive RF MEMS Switch on GaAs Substrate

GaAs substrate has two fundamental advantages over silicon substrate: the faster transistor speed and the smaller dielectric loss. Hence, it is very suitable for high frequency circuit applications. GaAs MMIC (Monolithic Microwave Integrated Circuit) process is a type of integrated circuit (IC) process that fabricates circuits operating at microwave frequencies. Here, a RF MEMS switch is designed and fabricated based on the standard GaAs MMIC process (Zheng et al. 2005).

#### Fabrication

The shunt capacitive RF MEMS switch is fabricated on a 600- $\mu\text{m}$ -thick GaAs substrate. The fabrication process used here is compatible with the standard MMICs. Figure 15 illustrates the detail procedures of the fabrication. The CPW lines are first defined by using a liftoff process through evaporating 800/300/2200 Å layer of AuGeNi/Au. 1000 Å dielectric SiN layer is then deposited and patterned. Following that, a 2  $\mu\text{m}$ -thick sacrificial layer of polyimide is deposited and patterned. The thickness of the polyimide layer determines the initial gap height of the membrane. A 500/1500/300 Å Ti/Au/Ti seed layer is finally sputtered and patterned. To remove the top Ti layer, the membrane and the CPW line are electroplated in a 55° cyanide-based solution, excluding the area underneath and near the beam. The thickness of the electroplated Au layer is 1.5  $\mu\text{m}$ . The sacrificial layer of polyimide is removed using developer, and the alcohol is utilized to get rid of the residual water in the switches.

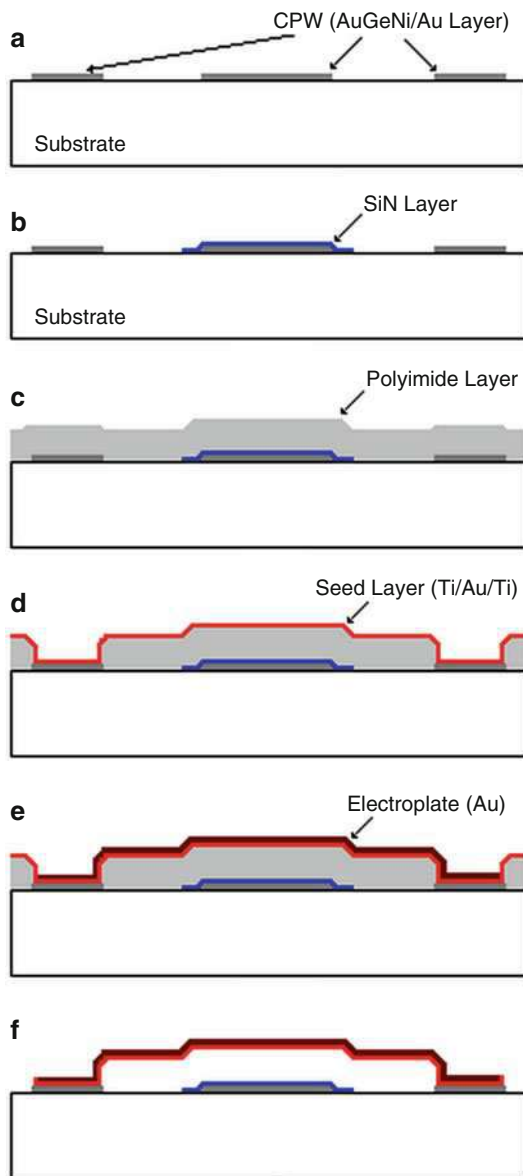
#### Design

Figure 16 shows the schematic diagram of the shunt capacitive RF MEMS switch. The switch is designed based on above fabrication process. The dimensions of the coplanar waveguide (CPW) G/S/G are 84/140/84  $\mu\text{m}$  (50  $\Omega$ ).

To calculate the threshold voltage of the switch, the elastic coefficient of the membrane bridge should be obtained first. A simplified model of the membrane bridge is shown in Fig. 17. According to the theory of stress and strain, the elastic coefficient  $k$  of the membrane bridge could be given as (Yang and Budynas 2002),



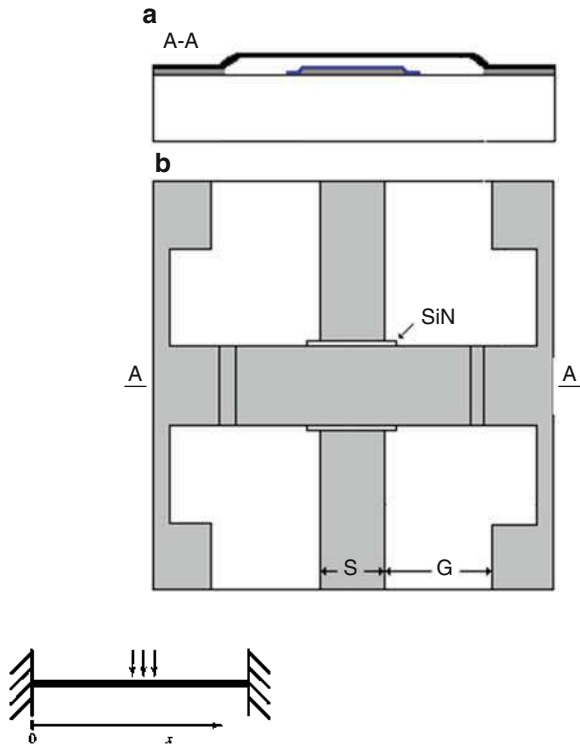
**Fig. 15** The fabrication procedures of the switch (a) CPW patterning, (b) a dielectric SiN layer depositing and patterning, (c) the polyimide layer depositing, (d) a seed layer (Ti/Au/Ti) sputtering, (e) a Au layer electroplating, (f) the sacrificial layer removing



$$k = 32Ew \left(\frac{t}{l}\right)^3 \frac{1}{8(x/l)^3 - 20(x/l)^2 + 14(x/l) - 1} \quad (3)$$

Where  $x$  represents the position along the membrane bridge,  $E$  is Young's modulus of the membrane material,  $t$  is the thickness of the membrane,  $l$  is the length of the membrane, and  $w$  is the membrane width. Substituting Eq. 3 into Eq. 2,

**Fig. 16** Schematic diagram of the switch (a) cross-section view (b) top view



**Fig. 17** A simplified model of the membrane bridge with the force applied at the center of the beam

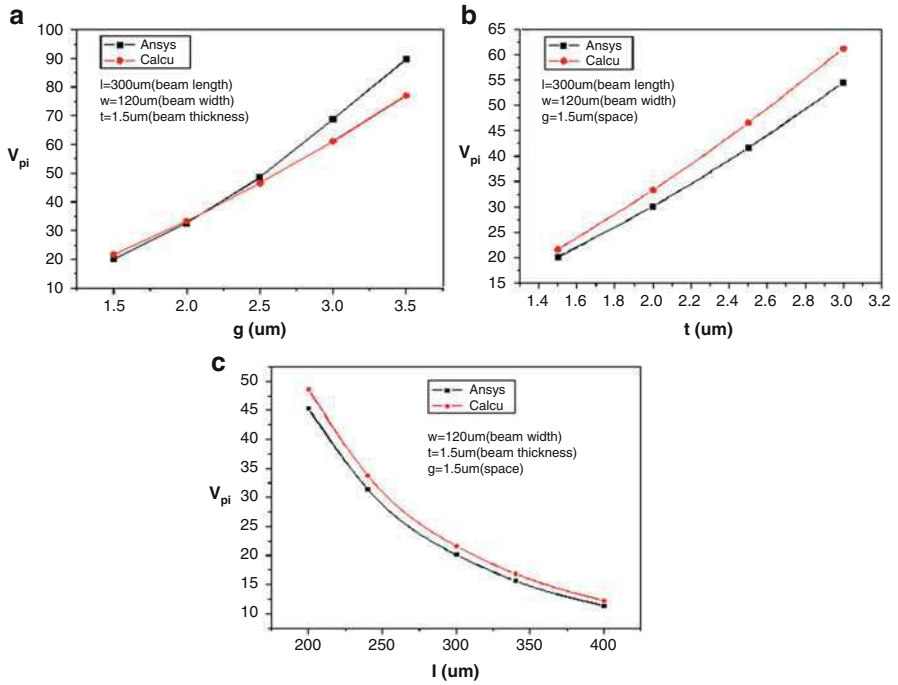
the threshold voltage  $V_{pi}$  of this shunt capacitive RF MEMS switch could be obtained. The mechanical characterization of the switch is also simulated by ANSYS software. The comparisons between model calculation results and the simulation results are shown in Fig. 18. It could be seen that the simulation results and model calculation results are in good agreement.

### Measurement

The top view of the fabricated switch is shown in Fig. 19. The cross-section SEM picture of the fixed-end beam is shown in Fig. 20.

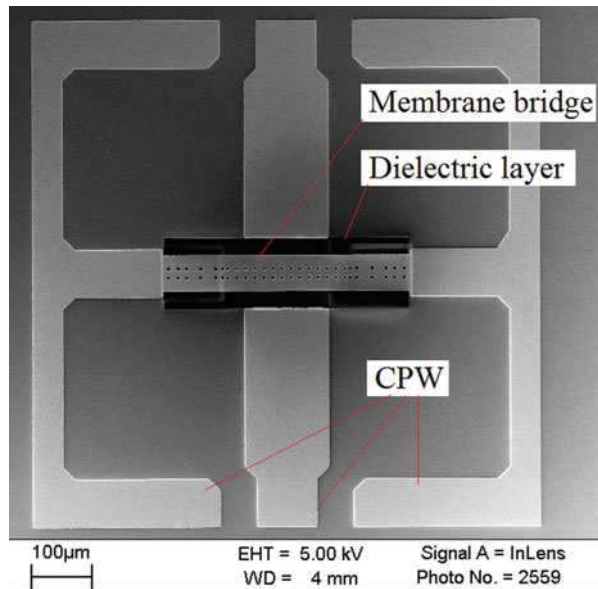
Figure 21 presents the S-parameters of the microwave switch for signals from 500 MHz to 25.6 GHz. The parameters are measured using a HP 8510C vector network analyzer and a Cascade probe station. The insertion loss at the on-state is less than 0.4 dB from 500 MHz to 25.6 GHz. The isolation at the off-state can reach  $-42$  dB at its self-resonant frequency. The self-resonant frequency of the switch is about 24.5 GHz. These measured S-parameters prove that the MEMS switches have superior high-frequency performance.

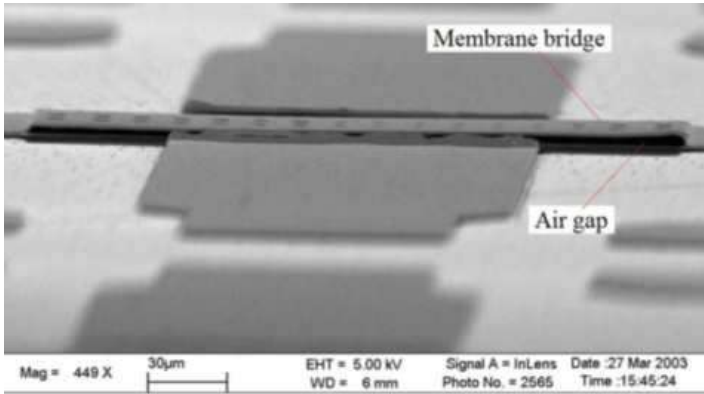
The return loss of the RF MEMS membrane switch in the on-state is also shown in Fig. 21a. This reflected energy is due to the parasitic capacitance caused by the



**Fig. 18** The threshold voltage as a function of beam dimensions (a) space, (b) beam thickness, (c) beam length

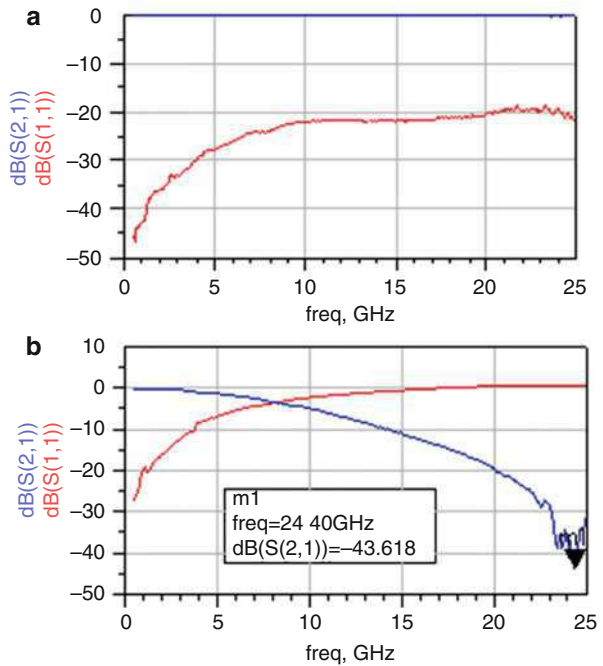
**Fig. 19** The SEM picture of the MEMS membrane switch



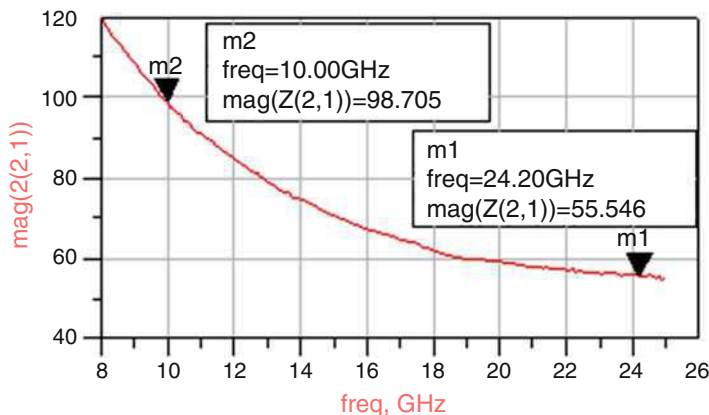


**Fig. 20** The cross-section SEM picture of the switch

**Fig. 21** (a) Up-state and (b) down-state S-parameters of the switch



proximity of the transmission line to the beam suspended above. In Fig. 21a, the return loss is less than  $-19$  dB at the frequency of 0.5–25.6GHz. This shows that the MEMS switch possesses an excellent impedance match to  $50 \Omega$ . The port impedance of the switch as a function of frequency is shown in Fig. 22. As the increment of frequency, the port impedance of the switch decreases and gradually tends to  $50 \Omega$ .



**Fig. 22** The port impedance of the switch as a function of frequency

## Shunt Contact RF MEMS Switch on LCP Substrate

LCP is a kind of flexible substrate. LCP substrate has very low dielectric loss, low cost, low water absorption, and multilayer circuit capability. These make it very appealing for RF devices and RF circuits (Kingsley et al. 2007). Hence, the RF MEMS switches fabricated on LCP substrate have the feature of flexible too (Gao et al. 2015).

### Fabrication

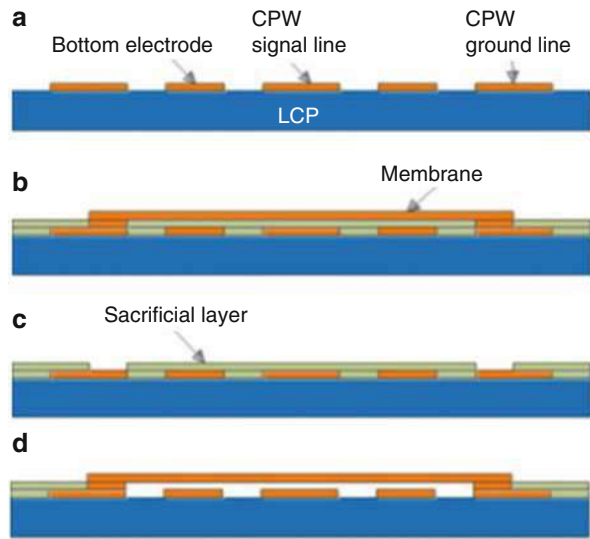
Figure 23 shows the LCP substrate process in detail. Firstly, by evaporating seed layer Cu, then electroplating and patterning Au, and then etching seed layer Cu, CPW transmission line and the bottom electrodes are formed. Secondly, a 3- $\mu\text{m}$ -thick photoresist is spin coated and patterned to define the air gap and two anchors of the membrane. Thirdly, the membrane is formed by electroplating and patterning Au layer again. Finally, the sacrificial photoresist layer is removed with a resist stripper to release the switch.

### Design

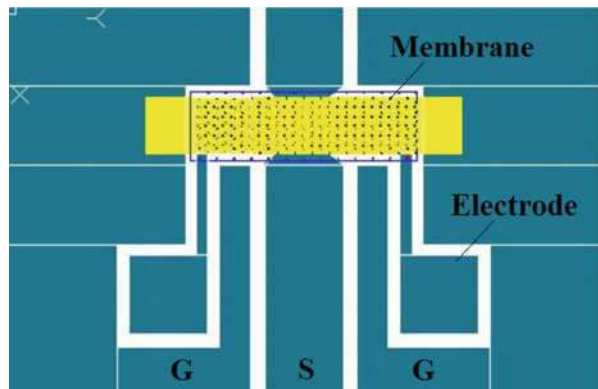
The shunt contact CPW switch with a membrane size of  $600 \times 140 \mu\text{m}$  is shown in Fig. 24 and the thickness of the air gap between the membrane and the bottom electrodes is 3  $\mu\text{m}$ . The switch consists of a clamped-clamped beam as the top electrode and two bottom electrodes between the signal line and the ground lines. The DC voltage which is applied to activate the switch and the RF signal of the shunt contact switch is separated.

A simplified model of the membrane bridge with the force applied at two sides of the beam is shown in Fig. 25. In order to actuate the switch, the bottom electrodes are

**Fig. 23** LCP substrate process steps of the shunt contact CPW switch, (a) electroplate and pattern the first gold layer, (b) form the sacrificial layer, (c) electroplate and pattern the second gold layer, (d) release the switch



**Fig. 24** Shunt contact RF MEMS switch

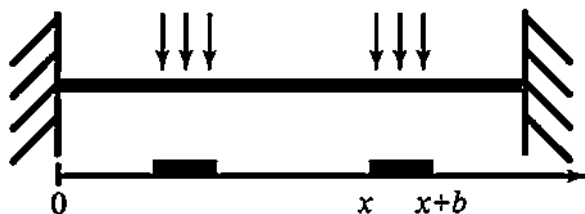


DC biased with respect to the ground. According to the theory of stress and strain, the effective spring constant  $k$  of the membrane is approximately by

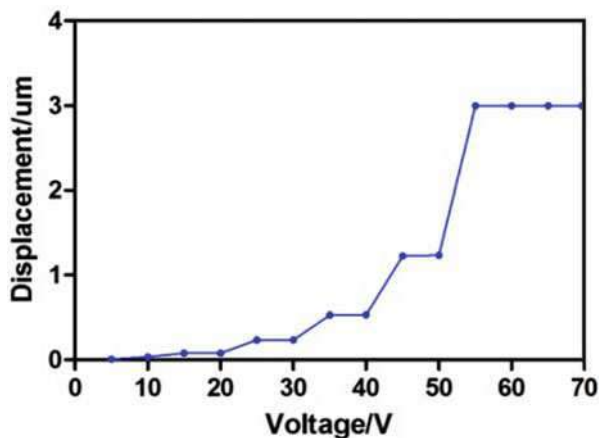
$$k = \frac{4Ewt^3b}{l^3(x+b) - 3l^2(x+b)^2 + 3l(x+b)^3 - (x+b)^4 - l^3x + 3l^2x^2 - 3lx^3 + x^4} \tag{4}$$

in which  $b$  is the bottom electrode width. The mechanical model of the switch is also established and simulated in ANSYS software. According to the simulation of ANSYS, we can find that the pull in voltage of the RF MEMS switch is about 55 V as shown in Fig. 26.

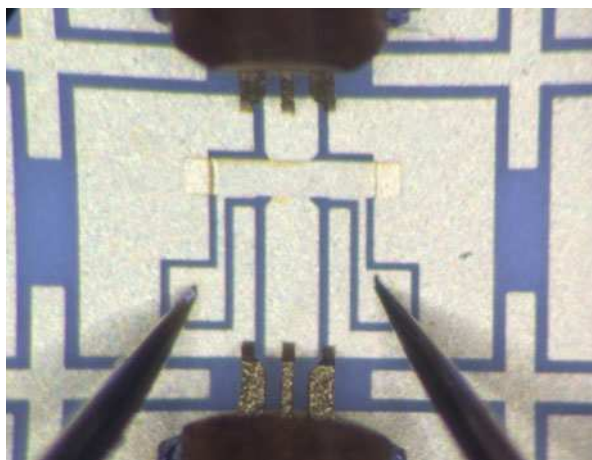
**Fig. 25** A simplified model of the membrane bridge with the force applied at two sides of the beam



**Fig. 26** The plot of displacement with voltage



**Fig. 27** The microscope photo of a fabricated shunt contact switch on LCP substrate

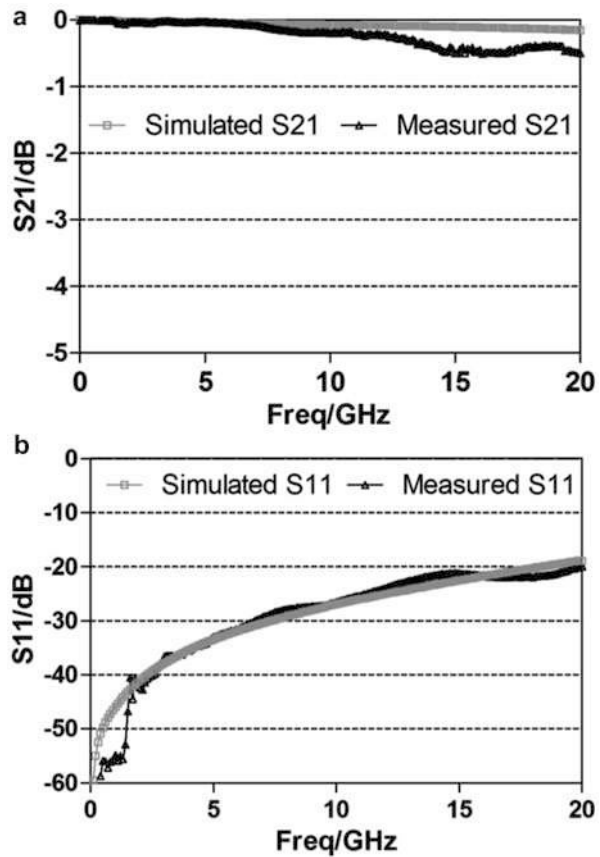


## Measurement

A microscope photo of a fabricated shunt contact switch on LCP substrate is shown in Fig. 27. The membrane size of the switch is  $600 \times 140 \mu\text{m}$ .

The measurements of the clamped-clamped CPW switches based on LCP substrates have been taken using an Agilent N5244A PNA-X network analyzer with

**Fig. 28** Measured (a) S21 and (b) S11 parameters of the clamped-clamped CPW switch on LCP substrate in the UP state



150 $\mu$ m-pitch cascade microtech ground-signal-ground coplanar waveguide probes. Figure 28 shows the experimental results of S21 and S11 parameters of the switch at the up state. Figure 29 shows the experimental results of S21 parameters of the switch at the down state.

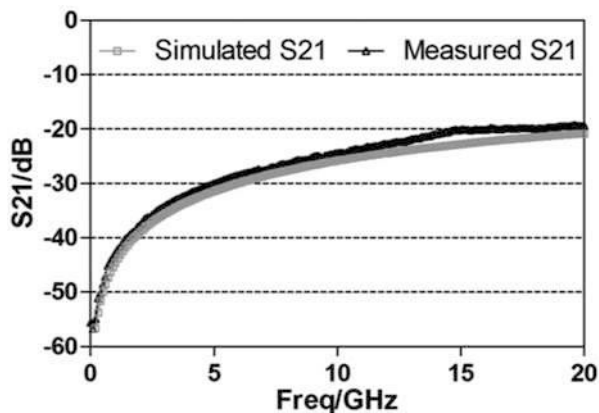
According to the measurement results, it could be seen that the switch shows good microwave performance. When the switch is at the down state, the S21 is around  $-20$  dB up to 20 GHz. When the switch is at the up state, the S21 is around  $-0.50$  dB up to 20 GHz, and the S11 is better than  $-20$  dB up to 20 GHz. The measured S parameters and the simulated S parameters are fitted well.

## Lateral Actuating RF MEMS Switches Based on SOG Process

As introduced in “[Basic Mechanical Principles](#),” lateral movement structures can do two-dimensional movements along the surface of the substrate. Hence, it is easier to design complex actuating structures by using lateral movement, which can be used



**Fig. 29** Measured S21 parameters of the clamped-clamped CPW switch on LCP substrate in the DOWN state



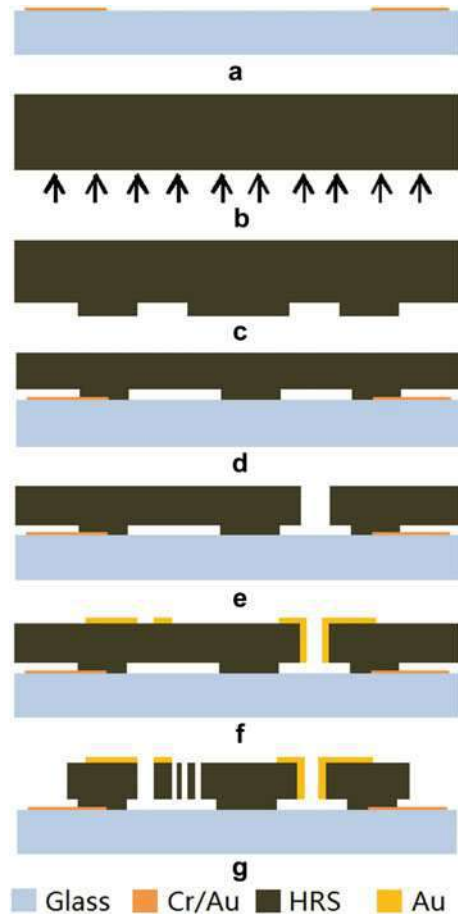
to improve some performances of RF MEMS switches. This section presents two lateral actuating RF MEMS switches, both of them are based on SOG process (Wang et al. 2013a, 2015). To fabricate RF MEMS switches, the silicon on glass (SOG) process should be modified and the SOG based microwave transmission line should be redesigned first.

## Fabrication

The fabrication process for those proposed switches above is based on the SOG process which makes use of an anodic bonding between silicon and glass. Pyrex 7740 Glass is used here as the device substrate. To fabricate microwave devices, some changes have been made to the standard SOG process. (1) High resistivity silicon ( $HRS > 20,000 \Omega \cdot \text{cm}$ ) is used as the substrate of the CPW transmission line to reduce the dielectric loss, and (2) low loss gold film is added to the SOG process. And it is thickened by electroplating to improve the microwave performance especially at low frequency bands.

Figure 30 gives the fabrication process flow. (a) A 500/1500 Å thick Cr/Au layer is sputtered and patterned using a liftoff process on the 500 μm thick Pyrex 7740 Glass substrate. This process forms the DC bias lines of the switch. (b) Boron ions are implanted into the bottom surface of the HRS substrate and annealed. This process is used to improve the metal-silicon contact resistance after anodic bonding. (c) Inductively coupled plasma (ICP) is used to etch the bottom surface of the silicon substrate by about 80 μm deep using a photoresist layer as the mask material. This step forms the anchors of the switch. (d) The silicon and the glass wafer are anodic bonded at 350°C in vacuum for 20 min with 1000 volts applied across the bond. Then the top silicon of bonded wafers is thinned and polished to about 130 μm thick using lapping and chemical mechanical polishing (CMP). So the thickness of the moving part layer is about 50 μm. (e) Contact holes are formed using the ICP etch process. (f) The 500/5000 Å thick Ti/Au plating base layer is sputtered. Gold is electroplated to a nominal thickness of 2 μm into the patterned photoresist stencil. The photoresist stencil used for

**Fig. 30** The modified SOG process flow for prototype switches. (a) DC bias line deposited and patterned, (b) boron ions implanted and annealed, (c) anchor formed by silicon etched, (d) anodic bond and silicon layer thinned, (e) contact holes formed by ICP, (f) CPW and contacts electroplated, (g) comb formed and structure released

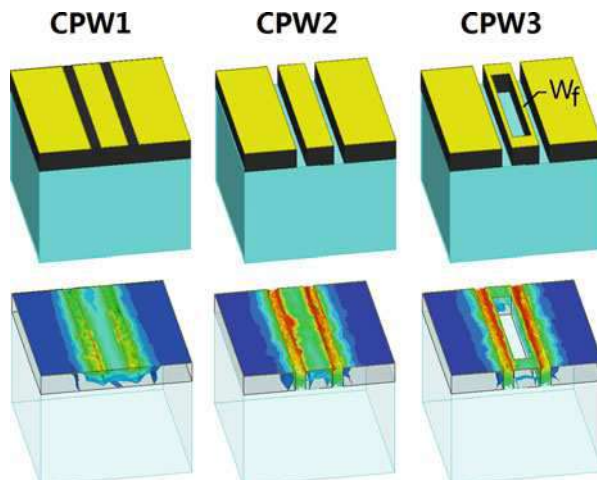


electroplating is AZ4330, and its thickness is around 3.5  $\mu\text{m}$ . Both the transmission line and the sidewall of the contact holes are electroplated. The gap distance of the contact part is finally decided by the thickness of the electroplated gold on its sidewall, so the process conditions of the electroplate process needs to be carefully controlled (including the current density, stirring rate, temperature, PH value, and time). To prevent the microwave signal reflection, actuator structures are not covered with gold. Then ion-beam etch (IBE) is used to remove the plating base metal and pattern the transmission line. (g) Finally, the differential parallel plate is formed and the device is released using the ICP etch process.

### Design and Analysis of SOG Process Based CPW Lines

The SOG process based structures mainly include the glass substrate layer and the silicon device layer that contains anchors and movable structures. If the SOG process

**Fig. 31** Schematic and electric field distribution of the three types of SOG process based CPW lines



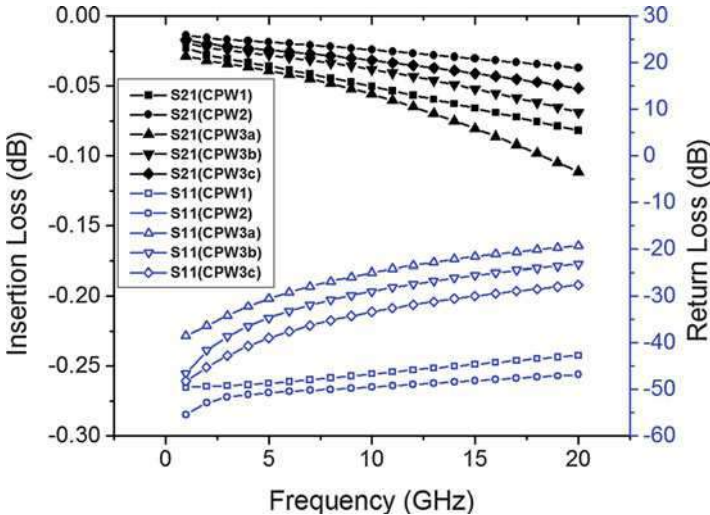
is used to fabricate RF MEMS switches, an extra metal layer will be needed on top of the silicon device layer for microwave transmission. The simplest SOG process based CPW transmission line is CPW1, as shown in Fig. 31, and high resistivity silicon (HRS) is used as the silicon device layer for low dielectric loss. The electric field distribution of CPW1 is also shown in Fig. 31 by HFSS simulation. To make transmission lines become laterally movable, the silicon layer between the signal line and the ground lines should be removed, the obtained transmission line is CPW2 shown in Fig. 31. Because the dielectric loss factor of air (nearly zero) is much less than HRS, it can be inferred that the insertion loss of CPW2 will be less than CPW1. It can be seen in Fig. 31 that both CPW1 and CPW2 have very small electric field strength at the inner part of their signal lines. Therefore, if the inner part of the signal line is removed, the transmission of microwave signals will not be much affected. The benefit of removing the inner part of the signal line is that the actuator of the switch can be placed inside the signal line. The proposed actuator embeddable SOG process based CPW line is CPW3 shown in Fig. 31, and  $W_f$  is the width of the frame of the signal line. The electric field distribution of CPW2 and CPW3 in Fig. 31 is very close to each other, and both of their microwave signals mainly transmit in the air between the signal line and the ground lines.

The microwave performances of CPW1, CPW2, and CPW3 are simulated and compared using HFSS software. To analyze the effect of the frame width  $W_f$ , CPW3 is subdivided into CPW3a with 10  $\mu\text{m}$  frame width, CPW3b with 20  $\mu\text{m}$  width frame, and CPW3c with 30  $\mu\text{m}$  width frame. The parameter settings for simulation are listed in Table 2.

The simulation results and comparison of all CPW lines are shown in Fig. 32. The insertion loss of CPW2 is much less than that of CPW1, which agrees with previous analysis. Due to the discontinuity of the signal line, the return losses of all three

**Table 2** Parameters for the simulation

Parameters	Values	
HRS	Thickness	100 $\mu\text{m}$
	Dielectric constant	11.9
	Loss tangent	0.015
Glass	Thickness	500 $\mu\text{m}$
	Dielectric constant	4.6
	Loss tangent	0.005
Gold	Thickness	2 $\mu\text{m}$
	Conductivity	$4.1 \times 10^7 \text{S/m}$
CPW	G/S/G	50/130/50 $\mu\text{m}$
	Length	800 $\mu\text{m}$
CPW3a	$W_f$	10 $\mu\text{m}$
CPW3b	$W_f$	20 $\mu\text{m}$
CPW3c	$W_f$	30 $\mu\text{m}$



**Fig. 32** Simulation results and comparison of all SOG process based CPW lines

CPW3 are larger than that of CPW1. The tendency is that the narrower the frame is, the larger the return loss is. However, both CPW3b and CPW3c have smaller insertion loss than that of CPW1, because their substrate dielectric losses are less than CPW1. Only CPW3a, the narrowest frame one, has worse insertion loss than CPW1. Therefore, it is recommended that the frame width of signal line should be larger than 10  $\mu\text{m}$  for the design under consideration. Although a concise analytical formula is not available, 3D FEM simulations seem to hint that a minimum acceptable  $W_f$  would be about 1/10 of the signal line length.

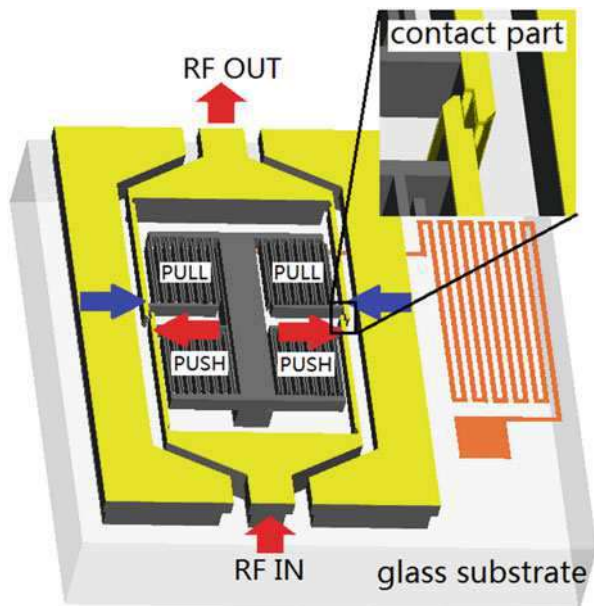
According to the simulation and comparison of SOG process based three types CPW lines, it is feasible to use CPW3 to design laterally actuated inside-driven RF MEMS switches.

### Design A: Push-Pull Type Switch

The first proposed SOG process based RF MEMS switch is a push-pull type switch. The push-pull structure was introduced into RF MEMS switches to increase their isolation (Hah et al. 2000; Cho and Yoon 2010). However, the common shortage of these push-pull switches is that the leverage or toggle structures are not easy to process and the actuation control signal is complex. Here, by using the SOG process and the inside-driven method, the push-pull structure is easy to realize. The schematic view of the push-pull switch is shown in Fig. 33. The switch consists of an actuator embeddable SOG process based CPW, the inside-placed actuators, the contact parts and the DC bias line. The hollow part of the CPW is broadened to accommodate larger actuators. The inside-placed actuators are electrostatic-driven laterally actuated differential parallel plate actuators. The DC bias line connecting to the actuators are located on the surface of the glass substrate. In order to avoid direct connect with CPW transmission lines, the DC bias line connects to the electrodes of the actuators through the recess at the anchor part. There is only one DC bias line, because the DC ground bias line shares with the RF signal line.

As shown in Fig. 33, four differential parallel plate actuators are connected to four ends of the signal lines, respectively, two of them are used to push and two of them

**Fig. 33** A schematic view of the push-pull type RF MEMS switch



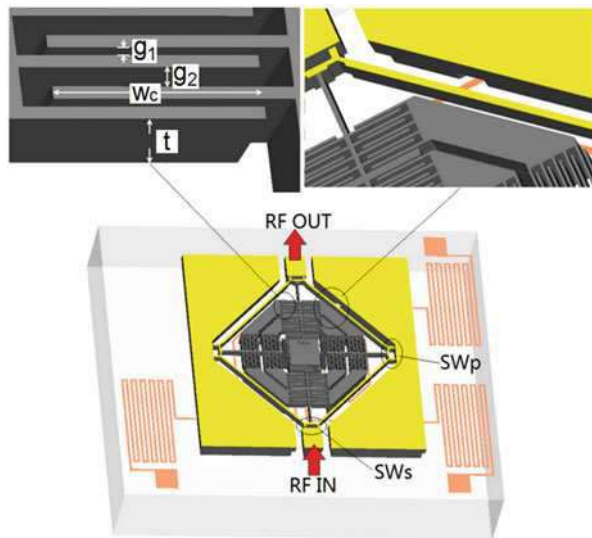
are used to pull. However, all their high potentials are connected to only one common DC bias line. Hence, the two drive groups can accomplish the push-pull action by only one actuation signal. Therefore, the proposed push-pull switch is not only easy to fabricate but also simple to control.

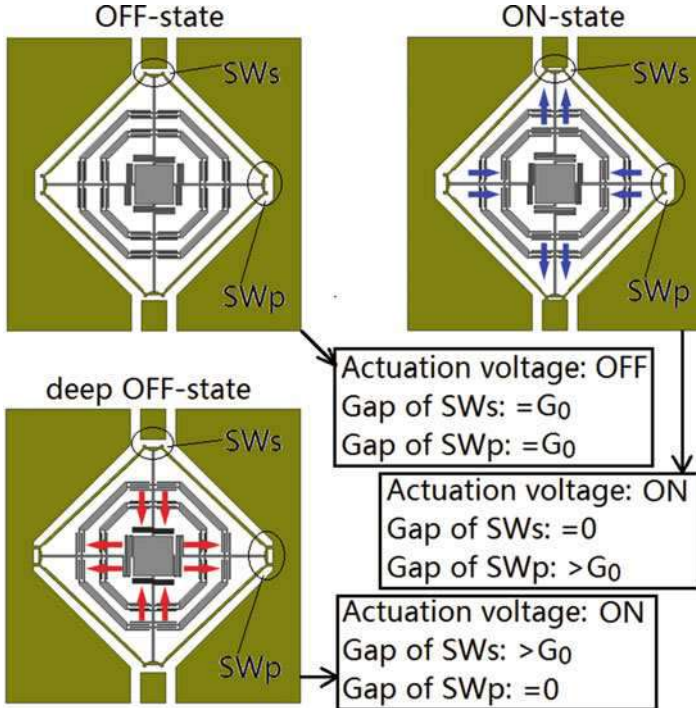
### Design B: Three-State Switch by Rhombic Beam Deformation

The second proposed SOG process based RF MEMS switch is a three-state RF MEMS switch, as shown in Fig. 34. Like the push-pull type switch, the three-state switch mainly consists of an actuator embeddable SOG process based CPW, the inside-placed actuators, the contact parts, and the DC bias lines. The difference is that the hollow part of the CPW is rhombic and the movable signal line of the three-state switch is a rhombic beam structure. Four lateral contacts, i.e., SWs and SWp, are located at the four corners of the rhombic beam. The function of the contact part is to switch the connection state between the resistor connection and the capacitor connection. So the contacts are labeled “SW,” it is an abbreviation for “switch.” Two groups of differential parallel plate actuators are linked with the end of the rhombic beam to precisely control the states of the switch. The three DC bias lines connecting to the two differential parallel plate actuators are located on the surface of the glass substrate.

The rhombic beam is deformable. The switching mechanism of the proposed switch is shown in Fig. 35. The initial state of the switch is at OFF-state since the four contacts may not contact the signal line or the ground line when no DC voltage is applied. The initial gap of SWs and SWp are both  $G_0$ . When an actuation voltage is applied on the outer actuators, the rhombic beam extends along the direction of the

**Fig. 34** Schematic view of a three-state RF MEMS switch by rhombic beam deformation





**Fig. 35** Description of the three states of the proposed switch

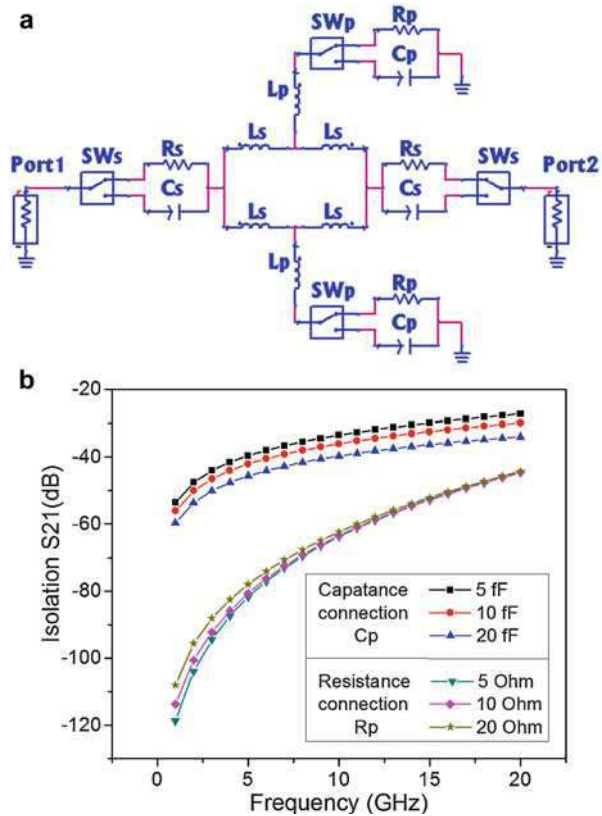
CPW line and shrinks along the vertical direction of the CPW line at the same time. This kind of deformation causes the SWs parts to contact with the signal line and the SWp parts to become farther away from the ground lines, so that the switch is at ON-state. When an actuation voltage is applied on the inner actuators, the rhombic beam deforms oppositely. This kind of deformation causes the SWs parts to become farther away from the signal line and the SWp parts to contact the ground lines. This situation is defined here as the deep OFF-state since the isolation of this state will be larger than the initial OFF-state.

A lumped circuit model for the proposed switch is shown in Fig. 36a. The contacts are modeled as resistors or capacitors, depending on the state of the switch.  $R_s$  and  $C_s$  are the contact resistor and capacitor of the SWs parts, respectively.  $R_p$  and  $C_p$  are the ones of the SWp parts, respectively. SWs is used to switch between  $R_s$  and  $C_s$ , and SWp is used to do between  $R_p$  and  $C_p$ . The rhombic beam is modeled as inductors.  $L_s$  is the series inductor of the rhombic beam, and  $L_p$  represents the parallel inductor between the signal line and the ground line.

Two reasons could make the isolation larger when the switch state changes from OFF-state to deep OFF-state. One is that the series capacitance  $C_s$  becomes



**Fig. 36** (a) The lumped circuit model of the three-state switch; (b) the comparison of isolations between resistor connections (deep OFF-state) and capacitor connections (OFF-state) using the lumped circuit model, while keeping all other parameters fixed as  $C_s = 10\text{fF}$ ,  $L_s = L_p = 0.5\text{nH}$



smaller. This could be explained by the vertical shrink of the rhombic beam. The other is that the parallel capacitance  $C_p$  is switched to parallel resistance  $R_p$ . By lumped circuit model simulation, the comparison of isolations between resistor connections and capacitor connections while keeping all other parameters fixed is shown in Fig. 36b. The results show that the resistor connection between signal line and ground line has larger isolation than capacitor connection during the 0–20 GHz range.

### Mechanical Design

The parameter definitions of the differential parallel plate actuator are shown in Fig. 34, where  $w_c$  is the length of the overlapping of the comb fingers,  $t$  is the thickness of the comb,  $g_1$  and  $g_2$  are the small and large gap of the comb fingers, respectively. The pull-in voltage of the differential parallel plate actuator is rewritten as (Rebeiz 2003),



**Table 3** Parameters of the differential parallel plate actuators and the movable beams

	$g_1(\mu\text{m})$	$g_2(\mu\text{m})$	$w_c(\mu\text{m})$	$t(\mu\text{m})$	$n$	$l_b(\mu\text{m})$	$w_b(\mu\text{m})$
Push-pull	4	10.5	207	50	8	425	15
Three-state	4	9.5	99	50	10	728	10

$$V_p = \sqrt{\frac{8kg_0^3}{27\varepsilon_0w_c t n}} \quad (5)$$

where  $\varepsilon_0$  is the dielectric constant of air,  $k$  is the spring constant of the movable beam,  $n$  is the number of the comb finger pairs, and  $g_0$  is the equivalent air gap. The equivalent air gap  $g_0$  can be approximated as

$$\frac{1}{g_0^2} \approx \frac{1}{g_1^2} - \frac{1}{g_2^2} \quad (6)$$

Parameters of the differential parallel plate actuators and the movable beams of the proposed push-pull switch and three-state switch are given in Table 3, where  $l_b$  and  $w_b$  is the length and the width of the movable beam, respectively. According to the analysis of SOG process based CPW3, larger frame width of the signal line is better for microwave transmission. However, larger frame width of the signal line causes larger spring constant and leads to higher actuation voltage. To reduce their spring constants, the movable beams are notched near their anchors, as shown in Figs. 33 and 34.

By using mechanical FEM simulation, it is very convenient to obtain the spring constant of these movable beams. Figure 37 shows the FEM simulation displacement of the movable beam of the push-pull switch and the deformation of the rhombic beam under 100 $\mu\text{N}$ , respectively. So the spring constant of the two switches are calculated as 75.2 N/m and 73 N/m, respectively. Substituting the simulated spring constant and the actuator parameters into Eqs. 1 and 2, the designed pull-in voltage of the push-pull switch and the three-state switches are calculated as 49.2 V and 65 V, respectively.

Multiplying the simulated spring constant by actuator travel distance, the recovery force  $F_r$  can be obtained. By using the actuator travel distance and the parameters in Table 3 as well as the measured pull-in voltage  $V$  of the switch, the stressed electrostatic force  $F_e$  can be easily calculated. Hence, the contact force  $F_c$  can be obtained as

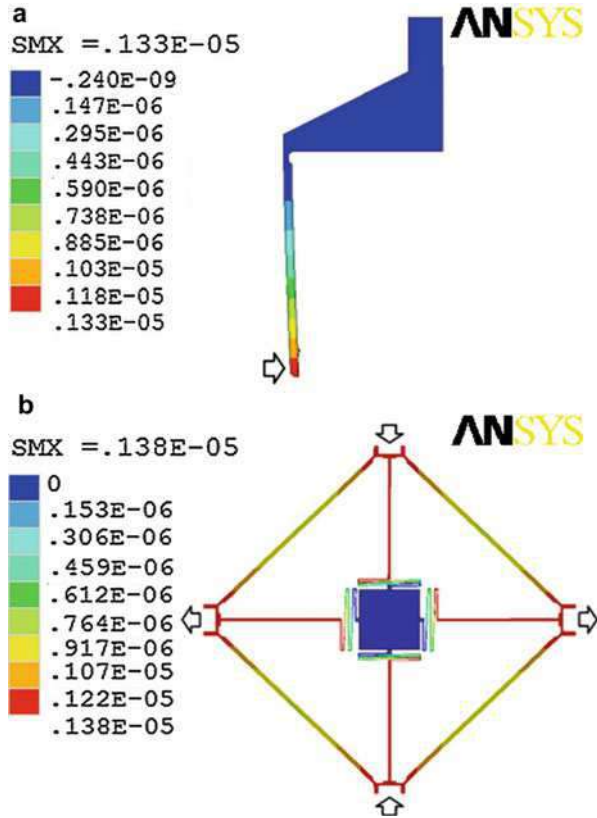
$$F_c = F_e - F_r \approx \frac{\varepsilon_0 w_c t n}{2(g_1 - d)^2} V^2 - \frac{\varepsilon_0 w_c t n}{2(g_2 + d)^2} V^2 - kd \quad (7)$$

where  $d$  is the actuator travel distance.

## Measurement Results and Discussion

The photograph of the fabricated push-pull type switch is shown in Fig. 38. The size of this push-pull type switch is  $1.4 \times 1.4 \text{ mm}^2$ . The SEM photograph of the

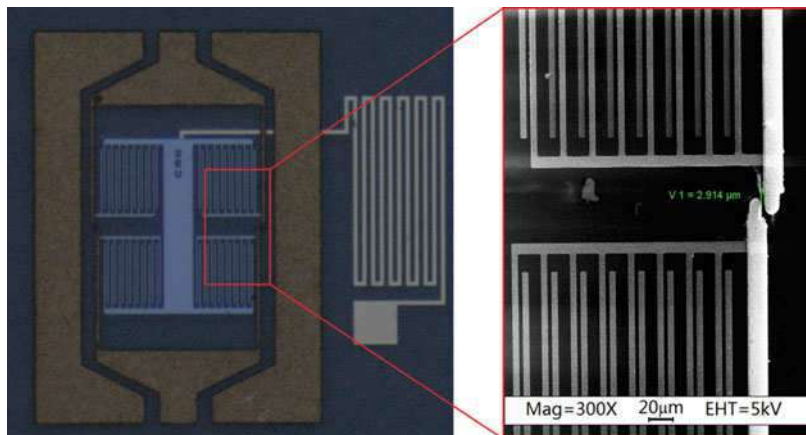
**Fig. 37** FEM simulation of (a) the movable beam of the push-pull type switch; (b) the rhombic beams of the three-state switch



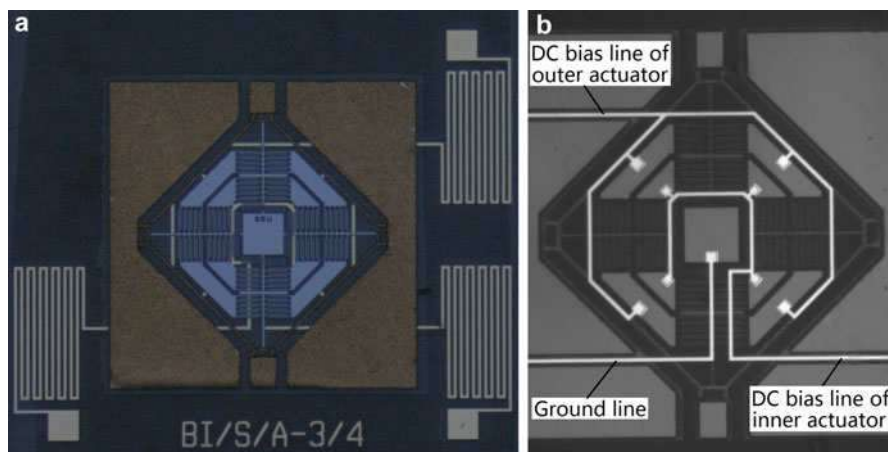
differential parallel plate actuator and the contact part of the fabricated switch are shown in Fig. 38. The measured small gap  $g_l$  of the actuator is  $4.3 \mu\text{m}$ , which is  $0.3 \mu\text{m}$  wider compared with the designed  $4 \mu\text{m}$  gap. This difference implies that the ICP process laterally etch  $150 \text{ nm}$  extra silicon each sidewall. The measured gap distance of the contact part of the push-pull switch is about  $2.9 \mu\text{m}$ .

The photographs of the fabricated three-state switch and its backside view are shown in Fig. 39. The size of the three-state switch is  $2.5 \times 2.5 \text{ mm}^2$ . The measured gap distance of the contact part of the three-state switch is about  $1.5 \mu\text{m}$ . It can be seen that, the fabricated gap distance is  $0.5 \mu\text{m}$  smaller than the designed one. In another word, the electroplated gold thickness is  $0.25 \mu\text{m}$  thicker than we demanded.

The RF characteristics of the fabricated switches are tested using the Agilent 8719ES S-parameter Network Analyzer together with the CASCADE microwave probing station. The mechanical characteristics are tested using a measurement setup as shown in Fig. 40. Both actuation voltage test waveform and switching time test waveform are supplied by the Agilent 33220A Function/Arbitrary Waveform Generator and the NF BA4825 High Speed Bipolar Power Amplifier. Here,  $1 \text{ Hz}$  Sawtooth waves are used for actuation voltage test, and  $1 \text{ kHz}$  square waves are



**Fig. 38** Photographs of the push-pull type switch



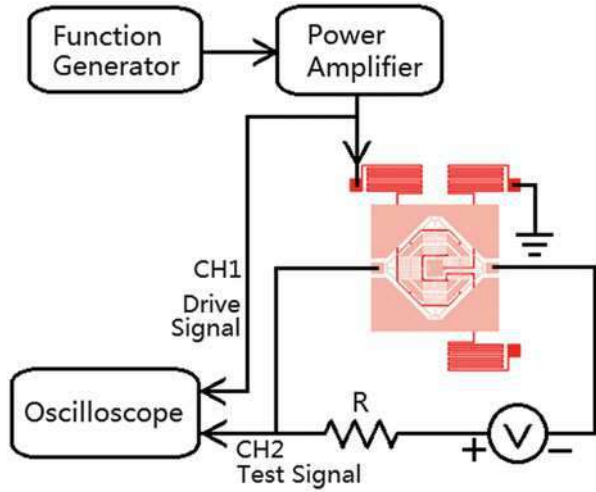
**Fig. 39** (a) A photograph overview of the rhombic switch; (b) the backside view of the switch structure

used for both switching time test and lifetime test. Both drive signal and the test signal are observed and compared by the UNI-T UTD2052CEL Digital Storage Oscilloscope. The temperature and relative humidity of the measurement environment is 25 °C and 40% RH, respectively.

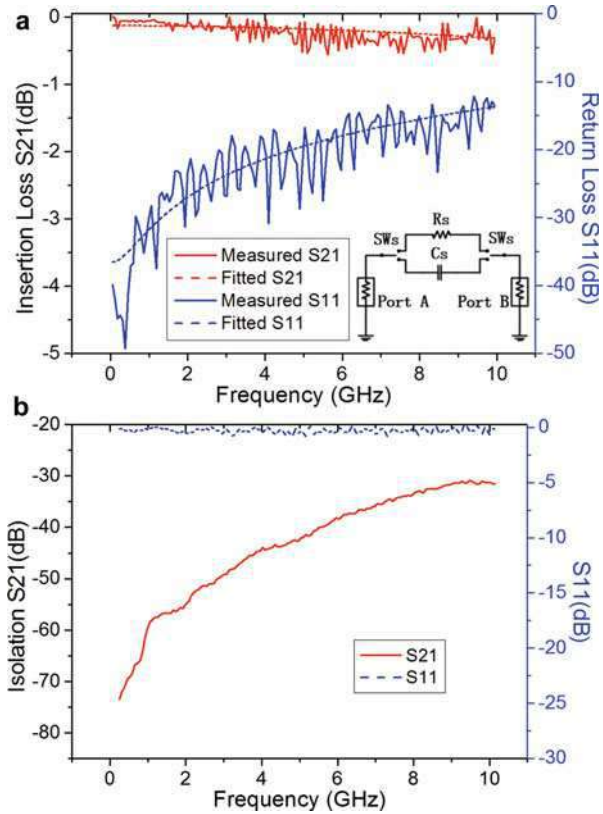
### RF Characteristics

Figure 41 gives the measured insertion loss and isolation of the push-pull type switch in the DC-10 GHz range, among which 6 GHz is one of the typical microwave frequencies in ground wireless communication applications. Figure 41a shows the

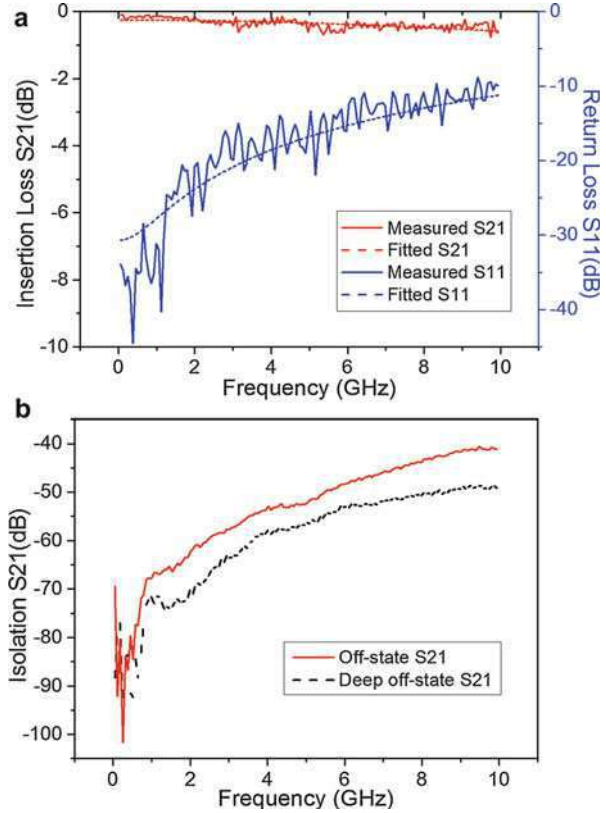
**Fig. 40** Measurement setup for mechanical characteristics testing



**Fig. 41** (a) Comparison of insertion loss between fitted data and measured data of the push-pull type switch; (b) measured isolation of the push-pull type switch



**Fig. 42** (a) Comparison of insertion loss between fitted data and measured data of the three-state switch; (b) the measured OFF-state and deep OFF-state isolation of the three-state switch

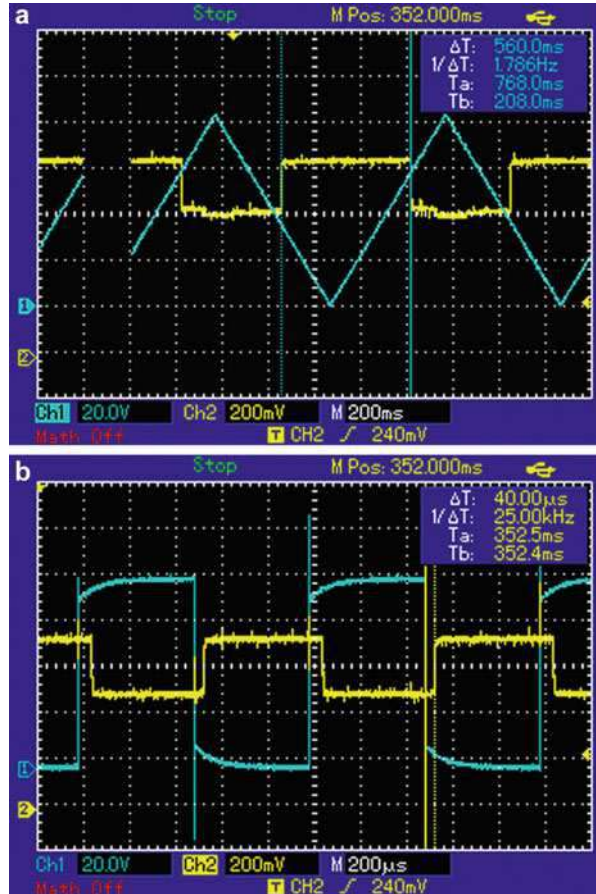


measured insertion loss is  $-0.28$  dB at 6 GHz, and the fitted series resistance  $R_s$  and series inductance  $L_s$  are  $1.5 \Omega$  (the contact area of the push-pull switch is  $4 \times 50 \mu\text{m}^2$ ) and  $0.17\text{nH}$ , respectively. The wiggly return loss may be caused by small parasitic reflections in setup. Figure 41b shows the measured isolation is  $-38.4$  dB at 6 GHz. As a result of the push-pull structure, good isolation value is obtained.

The ON-state S-parameter measurement result of the three-state switch is shown in Fig. 42a. Measured insertion loss of the three-state switch at 6 GHz is  $-0.38$  dB. The lumped circuit model fitted results are also shown in Fig. 42a. The contact resistance of the large switch is fitted as  $1.9 \Omega$  (the contact area of the three-state switch is  $5 \times 50 \mu\text{m}^2$ ), and the series inductance of them is  $0.89\text{nH}$ . The fitted lump model parameters indicate that the large insertion loss of the large rhombic switch is caused by its large series inductance, because return loss rapidly deteriorates with increasing series inductance.

The OFF-state and the deep OFF-state S-parameter measurement results of the three-state switch are shown in Fig. 42b. By using the series-shunt isolation mechanism, the switch exhibit very high isolation. Measured OFF-state and deep OFF-state isolation of the three-state switch at 6GHz are  $-48.3$  dB and  $-53.0$  dB. Both

**Fig. 43** The measurement results of (a) the pull-in voltage of the push-pull switch, (b) the switching time of the push-pull switch



measurement and simulation results proved that the deep OFF-state has a larger isolation than the normally OFF-state as we designed.

### Mechanical Characteristics

The measurement results of the pull-in voltage and the switching time of the push-pull switches are shown in Fig. 43. The measured result of the pull-in voltage of the push-pull type switch is 57 V, which is larger than the designed 49.2 V. The reason is mainly caused by the extra lateral etch of the gap of the actuators during the ICP process. Figure 43b shows the measured switching times of the push-pull switch. The measured switching times of the switch are 56  $\mu$ s for the transition from OFF-state to ON-state and 40  $\mu$ s for the transition from ON-state to OFF-state.

The measured pull-in voltage of the three-state switch is 78 V. The measured switching times of the three-state switch are 72  $\mu$ s for the transition from OFF-state to ON-state and 64  $\mu$ s for the transition from ON-state to OFF-state. The measured

switching times of the three-state switch are slightly larger than that of the push-pull switch.

Moreover, the contact forces of these switches could be calculated by Eq. 3. According to the measured contact gaps of the fabricated switches, the actuator travel distance of push-pull type switch and the three-state switches can be obtained as 1.45  $\mu\text{m}$  and 1.5  $\mu\text{m}$ , respectively. Substituting the actuator travel distances together with the measured pull-in voltages, the contact force  $F_c$  of the push-pull type switch and the three-state switch are obtained as 81  $\mu\text{N}$  and 105.5  $\mu\text{N}$ , respectively.

Finally, preliminary lifetime tests of both push-pull and three-state RF MEMS switches are carried out. To speed up the experiment, 1 kHz square waves are used and the testing time of each sample is set to 3 h. The experimental results show that both push-pull and three-state RF MEMS switches can pass the 3-h continuous lifetime test, which means that the lifetimes of both exceed the magnitude of  $10^7$  cycles.

---

## Conclusions and Further Reading

In this chapter, the configuration principles of RF MEMS switches is first introduced, including four basic EM models and two basic movement styles. According to the two basic movement styles, the vertical movement and the lateral movement, several RF MEMS switches with different features are shown. Two different vertical actuating membrane bridge RF MEMS switches are obtained based on two different substrates. The first one has the advantage of compatibility with GaAs MMIC process, and the second one has the feature of flexibility. Furthermore, another two lateral actuating RF MEMS switches are demonstrated based on the SOG process. By designing lateral actuating structures, some performances of the RF MEMS switches are improved.

The contents about RF MEMS switches here are very basic and limited. The researches about RF MEMS switches also include some other performances, reliabilities, packages, drive circuits, etc. Other important performances are, for example, the switch dynamic characteristic (Rebeiz 2003) and the power handling (Muldavin et al. 2003; Martinez et al. 2007). The reliability problems include the fatigue of the membrane, contact deterioration, and dielectric charging problems (Weiser 1998; Majumder et al. 2001; Wolf and Spengen 2002; Papaioannou et al. 2006; Herfst et al. 2008; Wang et al. 2013b). However, it has been reported that some RF MEMS switches already achieve ten billion cycle times (Czaplewski et al. 2013). The package is also a difficult problem, because the package needs to protect the movable structures and, at the same time, not affect the RF performance (Margomenos and Katehi 2004; Min and Rebeiz 2005; Tsou et al. 2007). Furthermore, the actuation voltage of RF MEMS switches usually higher than the voltage of IC devices, so additional boost circuit is needed (Goggin et al. 2011).



## References

- Adam F, Rick P, Ron N et al (2009) A single-crystal silicon DC–40 GHz RF MEMS switch. In: IEEE MTT-S digest, Boston, 2009
- Chan R, Lesnick R, Becher D et al (2003) Low-actuation voltage RF MEMS shunt switch with cold switching lifetime of seven billion cycles. *J Microelectromech S* 12(5):713–719
- Cho JJ, Yoon E (2010) Design and fabrication of a single membrane push-pull SPDT RF MEMS switch operated by electromagnetic actuation and electrostatic hold. *J Micromech Microeng* 20:1–7
- Cho JJ, Song T, Baek SH et al (2005) A low-voltage and low-power RF MEMS series and shunt switches actuated by combination of electromagnetic and electrostatic forces. *IEEE Trans Microw Theory Technol* 53(7):2450–2457
- Czaplewski DA, Nordquist CD, Patrizi GA et al (2013) RF MEMS switches with RuO<sub>2</sub>–Au contacts cycled to 10 billion cycles. *J Microelectromech S* 22(3):655–661
- Daneshmand M, Fouladi S, Mansour RR et al (2009) Thermally actuated latching RF MEMS switch and its characteristics. *IEEE Trans Microw Theory Technol* 57(12):3229–3238
- Gao XF, Han L, Nie M et al (2015) The study of a RF MEMS switch based on LCP substrate. In: IEEE sensors conference, Busan, 2015
- Girbau D, Pradell L, Lázaro A et al (2007) Electrothermally-actuated RF-MEMS suspended parallel switch. *Microw Opt Technol Lett* 49:2894–2896
- Goggin R, Wong JE, Hecht B, et al (2011) Fully integrated, high yielding, high reliability DC contact MEMS switch technology & control IC in standard plastic packages. In: IEEE sensors conference, Limerick, 2011
- Goldsmith C, Lin TH, Powers B et al (1995) Micromechanical membrane switches for microwave applications. In: IEEE MTT-S digest, Orlando, 1995
- Goldsmith C, Randall J, Eshelman S et al (1996) Characteristics of micromachined switch at microwave frequencies. In: IEEE MTT-S digest, San Francisco, 1996
- Hah D, Yoon E, Hong S (2000) A low-voltage actuated micromachined microwave switch using torsion springs and leverage. *IEEE Trans Microw Theory Technol* 48:2540–2545
- Herfst RW, Steeneken PG, Schmitz J et al (2008) Kelvin probe study of laterally inhomogeneous dielectric charging and charge diffusion in RF MEMS capacitive switches. In: IEEE international reliability physics symposium, Phoenix, 2008
- Kingsley N, Anagnostou DE, Tentzeris M et al (2007) RF MEMS sequentially reconfigurable sierpinski antenna on a flexible organic substrate with novel DC-biasing technique. *J Microelectromech S* 16:1185–1192
- Lee HC, Park JH, Park JY et al (2005) Design, fabrication and RF performances of two different types of piezoelectrically actuated Ohmic MEMS switches. *J Micromech Microeng* 15:2098–2104
- Majumder S, McGruer NE, Adams GG et al (2001) Study of contacts in an electrostatically actuated microswitch. *Sensor Actuat A Phys* 93(1):19–26
- Margomenos A, Katehi LPB (2004) Fabrication and accelerated hermeticity testing of an on-wafer package for RF MEMS. *IEEE Trans Microw Theory Technol* 52(6):1626–1636
- Martinez JD, Blondy P, Pothier A et al (2007) Surface and bulk micromachined RF MEMS capacitive series switch for watt-range hot switching operation. In: Proceedings of 2nd European microwave integrated circuit conference, Munich, 2007
- Min B, Rebeiz GM (2005) W-band low-loss wafer-scale package for RF MEMS. In: Proceedings of European microwave conference, Paris, 2005
- Muldavin J, Boisvert R, Bozler C et al (2003) Power handling and linearity of MEM capacitive series switches. In: IEEE MTT-S digest, Philadelphia, 2003
- Pacheco S, Nguyen CT, Katehi LPB (1998) Micromechanical electrostatic K-band switches. In: IEEE MTT-S digest, Baltimore, 1998



- Papaoannou GJ, Exarchos M, Theonas V et al (2006) Effect of space charge polarization in radio frequency microelectromechanical system capacitive switch dielectric charging. *Appl Phys Lett* 89:103512. (3 pp)
- Park J, Shim ES, Choi W et al (2009) A non-contact-type RF MEMS switch for 24-GHz radar applications. *J Microelectromech S* 18:163–173
- Persano A, Quaranta F, Cola A et al (2010) Alternative materials for RF MEMS switches in III-V technology. In: *Symposium on design test integration and packaging of MEMS/MOEMS (DTIP)*, Seville, 2010
- Petersen KE (1978) Micromechanical voltage controlled switches and circuits. In: *IEEE international electron devices meeting*, Washington, 1978
- Rebeiz GM (2003) *RF MEMS theory, design, and technology*. Wiley Interscience, Hoboken
- Schauwecker B, Mehner J, Strohm KM et al (2004) Investigations of RF shunt airbridge switches among different environmental conditions. *Sensor Actuat A Phys* 114:49–58
- Sedaghat-Pisheh H, Rebeiz GM (2010) Variable spring constant, high contact force RF MEMS switch. In: *IEEE MTT-S digest*, Anaheim, 2010
- Solazzi F, Tazzoli A, Farinelli P et al (2010) Active recovering mechanism for high performance RF MEMS redundancy switches. In: *Proceedings of 40th European microwave conference*, Paris, 2010
- Tang M, Liu AQ, Agarwal A (2005) A compact DC – 20 GHz SPDT switch circuit using lateral RF MEMS switches. In: *Proceedings of Asia-Pacific microwave conference*, Suzhou, 2005
- Tsou C, Li H, Chang HC (2007) A novel wafer-level-hermetic packaging for MEMS devices. *IEEE Trans Adv Packag* 30(4):616–621
- Verger A, Pothier A, Guines C et al (2010) Sub-hundred nanosecond reconfiguration capabilities of nanogap RF MEMS switched capacitor. In: *IEEE MTT-S digest*, Anaheim, 2010
- Wang LF, Han L, Tang JY et al (2013a) Lateral contact three-state RF MEMS switch for ground wireless communication by actuating rhombic structures. *J Microelectromech S* 22(1):10–12
- Wang LF, Tang JY, Huang QA (2013b) Effect of environmental humidity on dielectric charging effect in RF MEMS capacitive switches based on C-V properties. *J Microelectromech S* 22(3): 637–645
- Wang LF, Han L, Tang JY et al (2015) Laterally-actuated inside-driven RF MEMS switches fabricated by SOG process. *J Micromech Microeng* 25:065007. (10pp)
- Weiser J (1998) Switching contacts in relays. In: *19th international conference on electric contact phenomena*, Nuremberg, 1998
- Wolf ID, Spengen WM (2002) Techniques to study the reliability of metal RF MEMS capacitive switches. *Microelectron Reliab* 42:1789–1794
- Yang WC, Budynas RG (2002) *Roark's formulas for stress and strain*, 7th edn. McGraw-Hill, New York
- Yao ZJ, Chen S, Eshelman S (1999) Micromachined low-loss microwave switches. *J Micromech Microeng* 8(2):129–134
- Zheng WB, Huang QA, Liao XP et al (2005) RF MEMS membrane switches on GaAs substrates for X-band applications. *J Microelectromech S* 14(3):464–471



# A Microwave Power Sensor

Zhiqiang Zhang and Xiaoping Liao

## Contents

Overview .....	1078
Classification .....	1079
A Terminal Indirect-Heating MEMS Power Sensor .....	1080
Optimization of Basic Porotype .....	1081
Symmetric Structure .....	1086
Three-Channel Structure .....	1088
High Dynamic Range Structure .....	1089
Wireless Input Structure .....	1090
Sandwich Structure .....	1090
A Terminal Self-Heating MEMS Power Sensor .....	1091
An Inline Coupling MEMS Power Sensor .....	1092
Basic Structure .....	1093
Wideband Structure .....	1099
State Conversion Control .....	1100
Transmission Direction Detection .....	1103
An Inline Capacitive MEMS Power Sensor .....	1105
Tethered Beam .....	1105
Cantilever Beam .....	1107
An Inline Inserting MEMS Power Sensor .....	1108
Extension of MEMS Power Sensor .....	1109
A MEMS Microwave Phase Detector .....	1110
A MEMS Microwave Frequency Detector .....	1111
Summary .....	1112
References .....	1113

---

Z. Zhang (✉) · X. Liao

Key Laboratory of MEMS of the Ministry of Education, Southeast University, Nanjing, China

e-mail: [zqzhang@seu.edu.cn](mailto:zqzhang@seu.edu.cn)

---

**Abstract**

This chapter mainly introduces microwave power sensors based on the micro-electromechanical system (MEMS) technology, in order to achieve the power detection, gain control, and circuit protection. In RF and microwave frequencies, the MEMS power sensors have the advantages of miniaturization, low power, high sensitivity, and compatible with GaAs monolithic microwave integrated circuits (MMIC), etc. In structure, several kinds of the power sensors are described according to different application requirements. They utilize the form of coplanar waveguide transmission lines, with small structural dimensions (generally  $<1 \text{ mm}^2$ ). In theory, they adopt conversion principles of microwave power-heat-electricity or microwave power-force-electricity. In fabrication, they are accomplished with the GaAs MMIC process. In measurement, experiments demonstrate the validity of the proposed design and model. These sensors meet the characteristics of high performance and low cost. They can be used as implant devices and embedded in microwave communication and radar systems, such as the self-detection of the transceiver module and the measurement of leakage power in microwave module circuits. The MEMS microwave power sensors can directly measure the power of below 500 mW. For a higher microwave power measurement, it is usually necessary to couple or extract a portion of the microwave power by some structures. These MEMS power sensors have the ability to extend the frequency and phase measurements of microwave signals, constituting frequency and phase detectors.

---

**Keywords**

Microwave measurement · Power sensor · MEMS · Terminal · Inline · Frequency and phase detection · Seebeck effect

---

**Overview**

In the research of microwave technology, microwave power, also known as the amplitude of microwave signals, is one of three importance characteristics (power, frequency, and phase) of microwave signals. Microwave power measurement is indispensable in various aspects for studying the generation, transmission, and reception of the microwave signals and has become an important part of electromagnetic measurement. For dc and low frequency signals, the power of the signals is often obtained by multiplying the root mean square (RMS) voltage, the RMS current, and their phase angle that are measured on a load. For RF and microwave signals, the voltage or current of each part along a uniform transmission line often is not equal. This is due to the appearance of the standing wave and the change of the impedance. Especially, the transmission line of devices has a similar and shorter length than the wavelength. Thus, the voltage and current of the RF and microwave signals lose their uniqueness, making them difficult to measure. But the RF and microwave power still has a certain value. Therefore, the power to measure becomes a more popular measurement. Microwave power sensors that are key instruments

display an indirect measurement method of the microwave power. In principle, the sensors convert the microwave power into other energy forms (e.g., dc electrical energy, thermal energy, etc.) that are easy to measure. The microwave power sensors are similar to voltmeters or ammeters in low-frequency or dc circuits, and mainly applied in microwave circuit systems for the power detection, gain control, and circuit protection. Performance criteria of the power sensors mainly include: operation frequency, dc and microwave loss, sensitivity, linearity, dynamic range of measurable power, and so on. Furthermore, the frequency and phase of microwave signals can be detected by using the power sensors, namely a frequency detector and a phase detector.

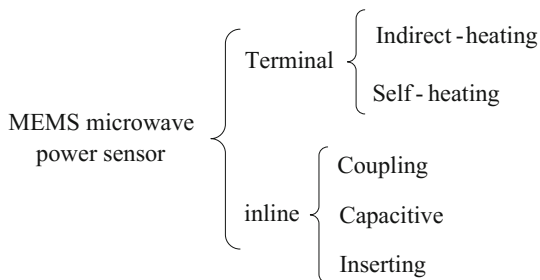
---

## Classification

In general, conventional microwave power sensors include three methods of diode, thermistor, and thermocouple (Daullé et al. 2001; Brush 2007). The former converts the microwave power into a low frequency signal based on the rectification of diodes, which is an active measurement. This method of the sensor requires additional dc power consumption during the power measurement. The latter two convert the microwave power into heat and then to an electrical signal based on the principle of microwave power-heat-electricity conversion, which are a passive measurement. The thermistor senses the change in resistance, and the thermocouple generates the output thermovoltage based on Seebeck effect. It means that there is no dc power consumption during the power measurement. But the thermistor-based sensors are susceptible to ambient temperature. And these thermoelectric sensors suffer from low sensitivity due to thermal losses of the substrate. With the process maturing of MEMS technology, MEMS devices show many advantages, for example, low-power consumption, miniaturization, and easy integration (Rebeiz 2003). Thus, the MEMS technology is able to combine with the conventional power measurement methods to produce high-performance MEMS microwave power sensors, particular suitable for the measurement method of the power-heat-electricity conversion. In addition, it promotes the emergence of new methods of microwave power measurement such as capacitive type.

At present, microwave power sensors based on the MEMS technology are widely investigated due to zero dc power consumption, low reflection loss, and high sensitivity. In terms of measuring signal ways, they are mainly categorized into terminal MEMS microwave power sensors (Dehé et al. 1996; Milanovic et al. 1997, 1998) and inline MEMS microwave power sensors (Han et al. 2007; Fernández et al. 2006; Dehe et al. 1995). In the terminal sensors, the microwave power under test is completely dissipated to generate heat and converted into thermovoltages based on Seebeck effect. In the inline sensors, a portion of the microwave power is extracted through some coupling structures (e.g., MEMS beam and coupler) and measured. For the terminal type, the input microwave signal is unavailable after the microwave power measurement, while for the inline type, most of microwave power is available after the microwave power measurement. According to the measurement principle,

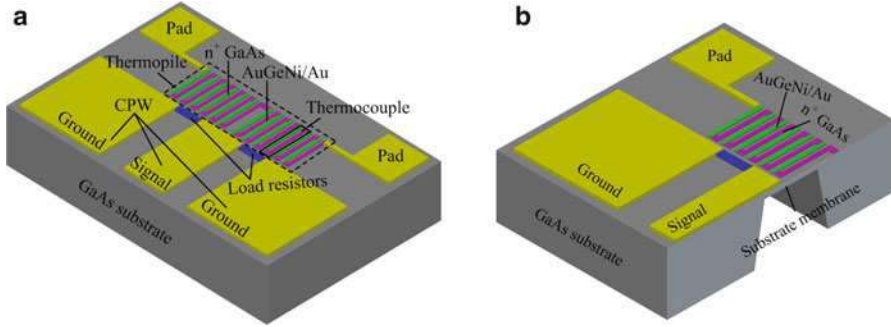
**Fig. 1** Classification form of MEMS microwave power sensors



the terminal sensors can be divided into indirect-heating type and self-heating type, while the inline sensors can be divided into coupling type, capacitive type and inserting type. The indirect-heating, self-heating, coupling and inserting types are based on the conversion principle of microwave power-heat-electricity, and are also called as a thermoelectric type; the capacitive type is based on the conversion principle of microwave power-force-electricity. Fig. 1 shows a classification form of the MEMS microwave power sensors. This chapter is introduced in accordance with the classification. These MEMS microwave power sensors are fabricated with the GaAs MMIC process. The MEMS sensors can directly measure the microwave power of below 500 mW. They are different from coaxial- and waveguide-based sensors that can measure large power (kW and MW levels). However, the coaxial- and waveguide-based sensors are large in volume and can not be integrated with circuit system. In this chapter, the MEMS microwave power sensors show small dimensions and can be embedded into the microwave communication and radar systems for the inline power measurement. If a higher microwave power measurement is required for the MEMS power sensors, they can extract or attenuate the microwave power by some microwave structures (e.g., couplers, attenuators). The structure sizes of these power sensors are still small and are compatible with the standard semiconductor process.

## A Terminal Indirect-Heating MEMS Power Sensor

The section introduces several indirect-heating MEMS power sensors with different applications for some specific purposes. Here, the so-called indirect-heating sensors indicate that they have separate resistors used to dissipate the microwave power. Fig. 2 shows a basic prototype of terminal indirect-heating MEMS microwave power sensors. They mainly consist of a coplanar waveguide (CPW) line, two load resistors, a thermopile, and two pads. The CPW is used to distribute microwave signals, with the characteristic impedance of  $50 \Omega$ . The load resistors are in parallel connected to the CPW, and their total resistance is designed to match with the CPW. A lot of thermocouples are in serial connected to constitute the thermopile for obtaining large output. The ends of the thermopile that is in proximity to and away from the resistors are called hot and cold junctions, respectively. The pads are used to

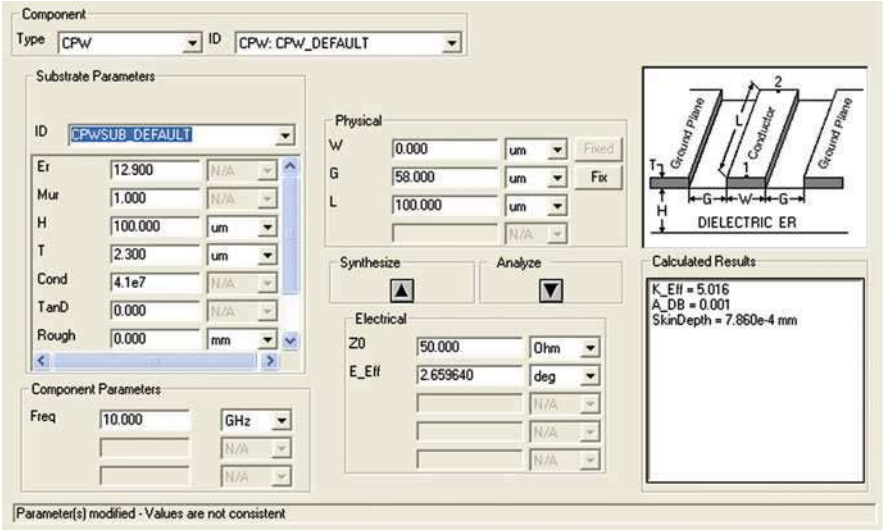


**Fig. 2** Basic prototype of terminal indirect-heating MEMS microwave power sensors. (a) Overall and (b) cross-sectional schematic views

output thermovoltages. The operation principle of the terminal indirect-heating MEMS power sensors is: an input microwave power is transmitted on the CPW and completely dissipated to generate heat by the load resistors; afterward, a heat distribution is produced around the resistors, and a temperature difference is formed along the thermopile; finally, the resulting temperature difference between the cold and hot junctions is converted into an output thermovoltage by the thermopile, based on Seebeck effect. In order to reduce thermal losses of the substrate, the MEMS technology is used to thinning the substrate and form the membrane structure underneath the resistors and the hot junctions. When the substrate in thickness is reduced, the thermal resistance in the region is increased and the temperature is increased. As a result, the thermovoltage and sensitivity of the thermoelectric power sensors are improved.

### Optimization of Basic Porotype

The basic porotype structure is the research foundation of other MEMS microwave power sensors. From the above analysis, we can see that microwave and thermal performances are crucial when the porotype sensors are operating. The microwave performance indicates a matching situation between the CPW characteristic impedance and the resistance of the load resistors. The better the matching is, the more the power dissipated by the resistors is and the less the reflected power is. It can be demonstrated by the reflection loss of the sensors. The thermal performance indicates that the efficiency of the heat transfer from the resistors to the hot junctions of the thermopile. The high efficiency will result in the large temperature difference between the cold and hot junctions and generate the large output thermovoltage. This can be demonstrated by the sensitivity. As a result, the reflection loss and sensitivity are considered as key points of the porotype design. Compared to traditional thermoelectric sensors, the MEMS microwave power sensors can show significant advantages in terms of low loss and high sensitivity. The purpose of this section is to optimize the loss and sensitivity of the porotype MEMS microwave power sensors.

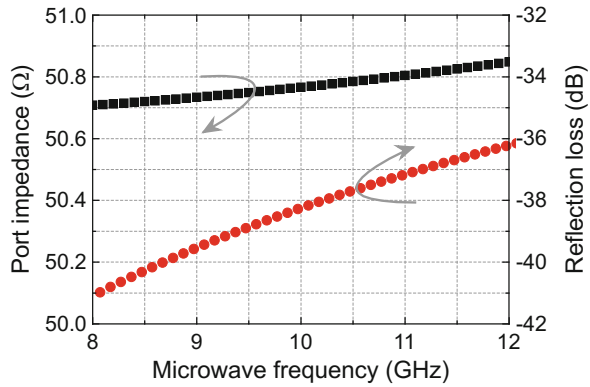


**Fig. 3** Calculated CPW dimensions in the lineCal tool

### Microwave Performance

The microwave performance of the power sensors indicates the matching design between the CPW and the load resistors, namely the design of the microwave power to heat conversion. Generally speaking, the characteristic impedance of the CPW is designed to be  $50 \Omega$ . Some preliminary structural dimensions in the CPW are first calculated by a lineCal tool in the ADS software, and then optimized by a HFSS software. Figure 3 shows the calculated CPW dimensions in the lineCal tool. In Fig. 3, for a  $2.3 \mu\text{m}$  thick CPW that is fabricated on a GaAs substrate with the thickness of  $100 \mu\text{m}$ , when the characteristic impedance is set to be  $50 \Omega$  at  $10 \text{ GHz}$ , the width of the CPW signal line and the gaps between the signal and ground lines are calculated to be  $100$  and  $58 \mu\text{m}$ , respectively. Following that, the resistance of each load resistors is designed to be  $100 \Omega$ . As seen from the microwave input port (Fig. 2a), the two resistors are connected in parallel. In the GaAs MMIC process, the CPW is made of Au while the resistors are made of TaN. Finally, the microwave performance of the prototype power sensors is optimized for accurate results. Figure 4 shows the optimized port impedance and reflection loss of the basic prototype MEMS microwave power sensors in HFSS. It shows good impedance matching between the CPW and the load resistors in the prototype design. In addition, a lumped S-parameter model, taking into account the parasitics of the resistors and the electromagnetic coupling interference between the resistors and the thermopile, is established for the prototype power sensors (Yan and Liao 2016a). It is found that the effects of the distance between resistors and the length of the thermopile on the reflection loss are negligible.

**Fig. 4** Optimized port impedance and reflection loss of the porotype MEMS power sensors in HFSS



**Thermal Performance**

The thermal performance of the power sensors indicates the design of the thermopile, namely the design of the heat to electricity conversion. Based on the Seebeck effect, the thermopile converts the heat caused by the microwave power into the dc thermovoltage. The thermopile is composed of many thermocouples connected in series. And the output thermovoltage ( $V_{th}$ ) is expressed as

$$V_{th} = \alpha_{AB} \sum_i^N (T_h - T_c) \tag{1}$$

where  $\alpha_{AB} = \alpha_B - \alpha_A$  is the difference of Seebeck coefficients between materials A and B,  $T_h$  and  $T_c$  are the hot and cold junctions of the thermopile respectively, and  $N$  is the number of the thermocouples. Thus, the sensitivity ( $S$ ) of the thermoelectric power sensors is represented as

$$S = \frac{V_{th}}{P_{in}} \tag{2}$$

where  $P_{in}$  is the input microwave power. Therefore, as seen by (1), the output thermovoltage is proportional to the difference of the Seebeck coefficients and the temperature difference between the hot and cold junctions; as seen by (2), the sensitivity is proportional to the thermovoltage. The Seebeck coefficients of the thermopile rely on selected materials. The temperature difference can be solved by using a steady-state heat conduction equation. And the heat conduction equation is written as

$$\frac{\partial^2 \Delta T}{\partial x^2} + \frac{\partial^2 \Delta T}{\partial y^2} + \frac{\partial^2 \Delta T}{\partial z^2} = 0 \tag{3}$$

where  $\Delta T = T_h - T_c$  is the temperature difference. It should be noted that the cold junctions are usually designed as a heat sink and their temperature ( $T_c$ ) is regarded as



the environment temperature. A two-dimensional (2D) heat conduction model of the terminal MEMS power sensors has been proposed and verified by using the ANSYS software (Yi et al. 2013). The temperature difference is deduced as

$$T_C - T_0 = \sum_{n=1}^{+\infty} C_n \left( e^{\frac{n\pi}{W}x} - e^{\frac{n\pi}{W}(2L-x)} \right) \sin \frac{n\pi}{W}y = 0 \quad (4)$$

where  $T_0$  is the environment temperature,  $W$  and  $L$  are the width and length of the boundary in the sensors, and the coefficients  $C_n$  is given by

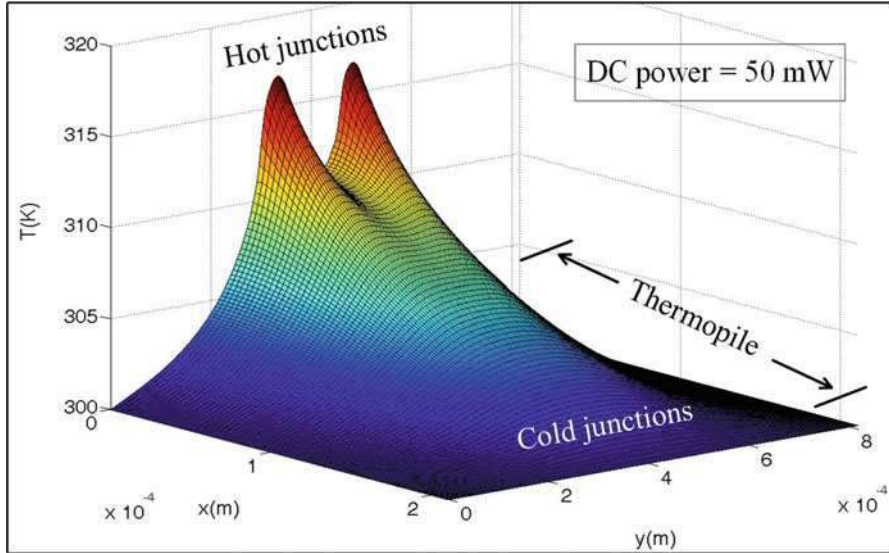
$$C_n = - \frac{2q}{n\pi\lambda \left( 1 + e^{\frac{2n\pi L}{W}} \right)} \left( \int_w^{w+g} \sin \frac{n\pi}{W}y dy + \int_{w+g+s}^{w+2g+s} \sin \frac{n\pi}{W}y dy \right) \quad n=1, 2, 3, \dots \quad (5)$$

where  $q$  is the heat flux density and relates to the microwave power,  $\lambda$  is the thermal conductivity,  $w$  and  $s$  are the width of the CPW signal and ground lines respectively, and  $g$  is the distance of the gap between the signal and ground lines. Figure 5 shows the calculated temperature distribution of the prototype power sensors under a dc power of 50 mW. As can be observed in Fig. 5, the locations of two highest temperatures are the resistors, and the hot junctions of the thermopile near the resistors far higher than the cold junctions. Taking the heat flux density as a function of the thickness of the substrate into account, a 3D heat conduction model is established to better describe the temperature distribution (Yi and Liao 2016a). Relative errors between the models and the measurement are reduced from 7.5% in the 2D model to 3.7% in the 3D model.

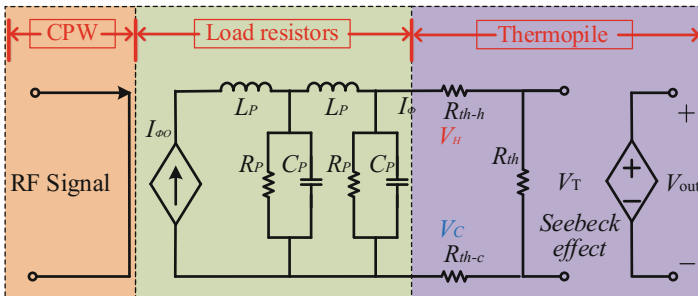
In order to achieve the compatible design of the circuit simulation, an electrical model of the prototype microwave power sensors is developed using an analogy method (Han and Liao 2016a). In the analogy, the thermal resistance, heat flow, and temperature difference are equivalent to the resistance, current, and voltage, respectively. Using the model, the effects of the microwave frequency on the temperature distribution and the sensitivity of the sensors are analyzed. Figure 6 shows the electrical model of the prototype sensors. In Fig. 6, the current controlled current source  $I_{\Phi_0}$  represents the heat generated by the load resistors; the inductance  $L_p$ , capacitance  $C_p$ , and resistance  $R_p$  represent parasitic effects and the resistance of the resistors;  $R_{th-h}$  and  $R_{th-c}$  describe the thermal resistance of the cold and hot junctions, respectively; the voltage  $V_T$  stands for the total temperature difference between the hot and cold junctions. Besides, a voltage controlled voltage source is utilized to account for Seebeck effect through setting control coefficients as the Seebeck coefficients, and the voltage  $V_{out}$  generated in the voltage source is the output thermovoltage of the sensors.

### Optimized Validation

The performances are dependent on the distance between the resistors and the thermopile, the length of the thermopile, the thickness of the substrate membrane, and the suspension length of the thermopile, and so on. These dimensions have been



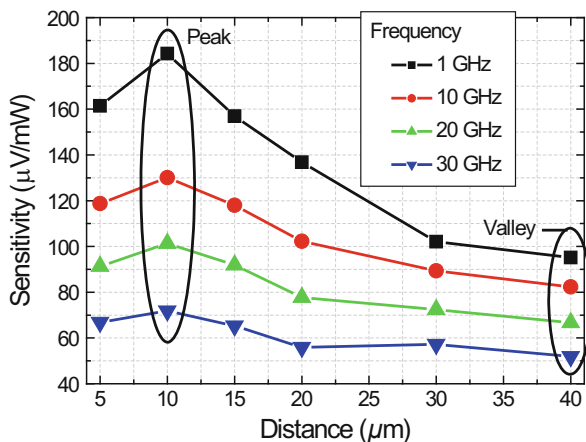
**Fig. 5** Calculated temperature distribution of the prototype power sensors under a dc power of 50 mW



**Fig. 6** Electrical model of the prototype sensors

optimized by the measurement analysis. The effects of the distance between the resistors and the thermopile and the length of the thermopile on the power sensors are investigated in Wang et al. (2012a). When the distance is from 5 to 40  $\mu\text{m}$  and the length is from 50 to 400  $\mu\text{m}$ , the measured reflection loss are less than  $-20$  dB up to 20 GHz. Figure 7 shows the sensitivity with respect to the distance between the resistors and the thermopile. It is found that the maximum sensitivity is at the distance of 10  $\mu\text{m}$ . The longer the thermopile is, the larger the sensitivity is. This is because that the longer thermopile increases the temperature difference between the cold and hot junctions of the thermopile. But the longer thermopile also increases the internal resistance, which leads to a decline in the signal-to-noise ratio (SNR). So the thermopile of 200  $\mu\text{m}$  in length is selected.

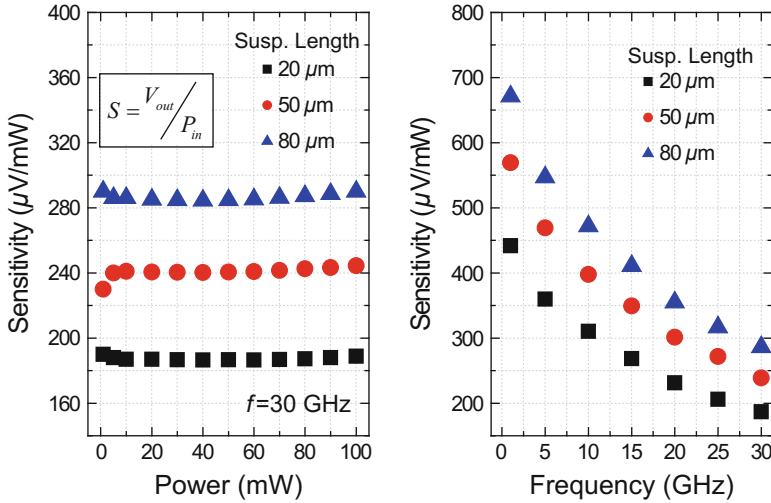
**Fig. 7** Sensitivity with respect to the distance between the resistors and the thermopile



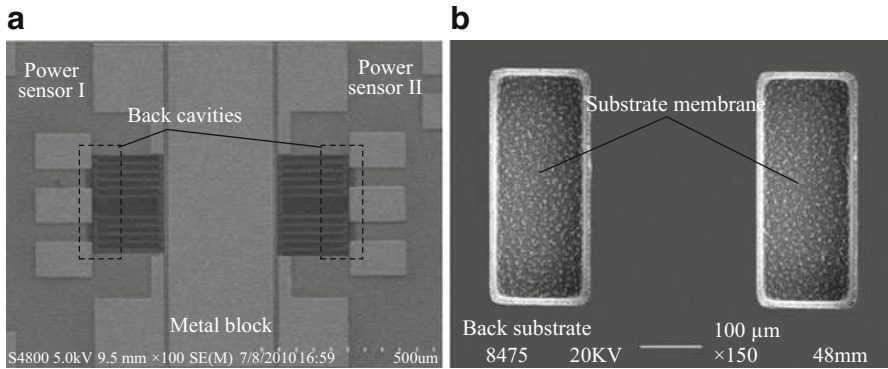
As seen in Fig. 5, heat is spread from the hot spots to the surrounding. The number of thermocouples is carefully designed to compromise the sensitivity and SNR. For the 2, 6, and 14 pairs of the thermocouple distribution, the sensitivities are 36.42, 107.09, and 240.44  $\mu\text{V}/\text{mW}$  at 30 GHz and the measured resistance of the thermocouples is 30.14, 89.84, and 199.30  $\text{k}\Omega$ , respectively (Zhang et al. 2015). The SNR is defined as the ratio of the sensitivity to the noise voltage, where the noise voltage is proportional to the square root of the resistance. Thus, the corresponding SNR is  $1.6 \times 10^6$ ,  $2.8 \times 10^6$ , and  $4.2 \times 10^6/\text{W}$  for the 2, 6, and 14 pairs of thermocouples. It means that about 14 pairs of thermocouples is a preferable choice for the resistors with the width of 58  $\mu\text{m}$ . In order to increase the sensitivity of the sensors, the substrate underneath the resistors and the hot junctions has to be removed and forms the membrane structure. Apparently, the thinner the substrate membrane is, the larger the sensitivity is. But the too thin membrane will bring challenges to the process and the reliability. Furthermore, the size of the back cavity affects the temperature distribution. In a word, the suspension length of the thermopile can improve the sensitivity. For the thermopile with the length of 100  $\mu\text{m}$ , the suspension length of 80  $\mu\text{m}$  shows higher sensitivities than others, as shown in Fig. 8 (Zhang and Liao 2015a). Here, the substrate membrane is fabricated to be 10  $\mu\text{m}$  in thickness. To obtain the output thermovoltages on the order of  $\text{V}/\text{W}$ , an amplification circuit suitable for the microwave power sensors is built (Wang and Liao 2010). For the prototype sensors, the response time is on the order of ms (Yan and Liao 2016b). In addition, the package is studied (Wang and Liao 2012a).

## Symmetric Structure

In order to reduce the temperature drift and the mismatch uncertainty, symmetric MEMS microwave power sensors are presented (Wang and Liao 2009, 2012b). Figure 9 shows the symmetric structure of these sensors. Their structure consists



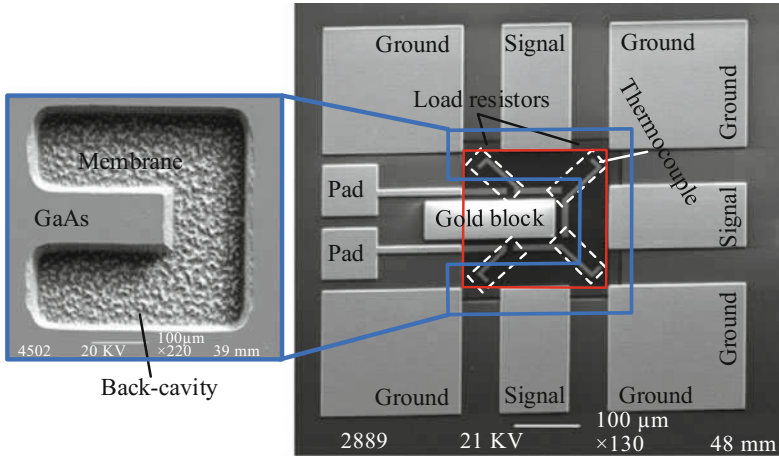
**Fig. 8** Sensitivity with respect to the suspension length of the thermopile



**Fig. 9** Symmetric structure of indirect-heating MEMS microwave power sensors. (a) Front structure and (b) two back-cavities (substrate membrane)

of two identical prototype power sensors, where they are placed symmetrically. One side inputs the microwave power under test while the other inputs a known dc power. When the output thermovoltages of the thermopiles on both sides are equal, the microwave power is known by the dc power. A large metal block is inserted between the two prototype sensors for maintaining the cold junctions as the heat sink.

Due to conductor and dielectric losses of the CPW, parasitic losses of the resistors, and electromagnetic coupling loss of the thermopile in the microwave frequency, the thermovoltages for the constant power and the sensitivity change with the change of the microwave frequency. Using the symmetric structure, the frequency-dependent characteristics can be eliminated to achieve the accurate and



**Fig. 10** Three-channel structure of indirect-heating MEMS microwave power sensor

wideband measurement of the microwave power (Wang et al. 2012b). For example, the left section inputs the microwave power under test at X-band while the right section inputs the known power at the center frequency of 10 GHz as a calibration port. When the output thermovoltages on both sides are equal, the calibration factors are obtained by comparing the measured power at the different frequency with the known power at 10 GHz. The experiments demonstrate that the maximum relative measurement error is 5.9% before the calibration, but is reduced to 0.96% after the calibration.

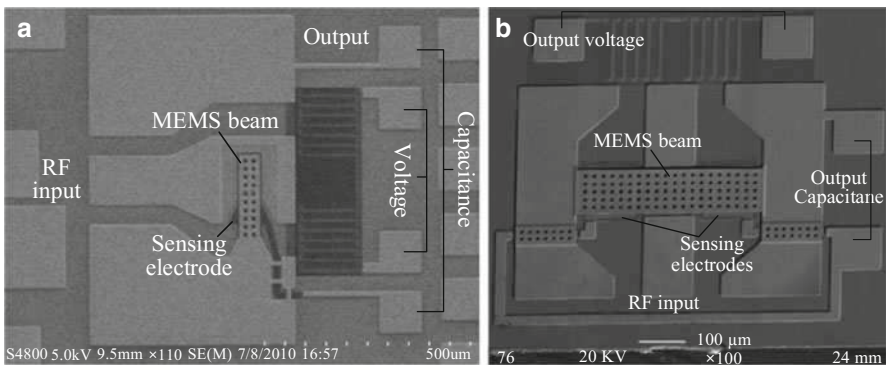
### Three-Channel Structure

In order to meet a multichannel power measurement, a miniaturization thermoelectric MEMS microwave power sensor with three inputs and one output is reported (Zhang and Liao 2012a). This sensor can achieve the power measurement of no less than three input microwave signals. Figure 10 shows this three inputs and one output structure. It mainly includes three CPW lines, six matching load resistors, four thermocouples, a gold block, and two contact pads. The CPW lines share adjacent ground lines to reduce the chip area. Every two resistors are connected in parallel to one CPW output with the matching impedance, which is used to convert the microwave power into heat. The four thermocouples made of  $n^+$  GaAs and AuGeNi/Au are symmetrically placed along two diagonals of a square (red line). In this way, a relative uniform thermal distribution on the thermocouples is obtained when the resistors are heated. The four thermocouples are connected in series to constitute a thermopile. The block is used to maintain the thermal stability of the cold junctions, which contributes to increasing the temperature difference between the cold and hot junctions. Because this power sensor has three CPW-based input

ports, it achieves the power measurement with up to three microwave signals. More importantly, the sensor only has one dc output, with the high integration. The experiments verify the design validity. For the three input ports, less than 1% of the input power is reflected at 8–12 GHz, and the sensitivities for the ports 2 and 3 show relative errors of 6.0% and 10.9% at 12 GHz compared to the port 1, respectively. Such results show that the three-channel sensor has good impedance matching in the process of the microwave power-heat conversion and acceptable thermal isolation in the process of the heat-electricity conversion.

### High Dynamic Range Structure

In order to obtain the power measurement of a high dynamic range, thermoelectric MEMS microwave power sensors with a cascaded capacitive sensor are proposed (Wang et al. 2012c; Yi and Liao 2016b). In the design, the capacitive sensor lies between the microwave port and the thermoelectric sensor. The capacitive sensor mainly consists of a MEMS beam suspended above the CPW and the sensing electrodes beneath the beam. The beam may be a tethered beam or a cantilever beam. When a high power is transmitted and passes through the beam, an electrostatic force is produced between the beam and the CPW signal line. The force causes the beam to bend. Thus, the capacitance change between the beam and the electrodes is detected. In theory, the cantilever beam has smaller stress than the tethered beam, resulting in a large capacitance change. In operation, the thermoelectric sensor is used to measure lower input power, while the capacitive sensor is utilized to measure higher power under test. This is because that the high power enables the thermoelectric sensor nonlinear distortion or burn, and the low power has a little effect on the capacitive sensor. Figure 11 shows two high dynamic range structures based on the tethered and cantilever beams. Using the tethered beam (Wang et al. 2012c), the measured power range is 0.1–100 mW for the thermoelectric sensors



**Fig. 11** High dynamic range structure of indirect-heating MEMS microwave power sensors based on (a) cantilever and (b) tethered beams

and 100–200 mW for the capacitive sensor at 8–12 GHz. Using the cantilever beam (Yi and Liao 2016b), the measured power range is 0.1–100 mW for the thermoelectric sensor and 100–150 mW for the capacitive sensor at 8–12 GHz. In practice, the thermoelectric and capacitive power sensors have larger dynamic range of the power measurement.

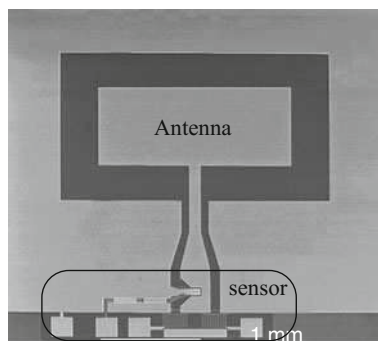
## Wireless Input Structure

In order to achieve the wireless power measurement, a thermoelectric MEMS microwave power sensor with a cascaded CPW-fed antenna is developed (Wang and Liao 2012c). This structure includes two parts: the antenna as a receiving part and the MEMS power sensor with the high dynamic range structure as a detection part. The antenna adopts a CPW feed mode at the port to facilitate the connection of the sensor in the plane. The CPW-fed antenna is used to receive the power of the wireless transmission. After that, the received power is distributed on the CPW and transmitted to the sensors. Finally, the MEMS sensor measures the received power using the principle of the power-heat-electricity conversion. Figure 12 shows the wireless input structure of the indirect-heating MEMS microwave power sensor. In Wang and Liao (2012c), this wireless power sensor is designed to operate at the center frequency of 35 GHz, with a bandwidth of 500 MHz.

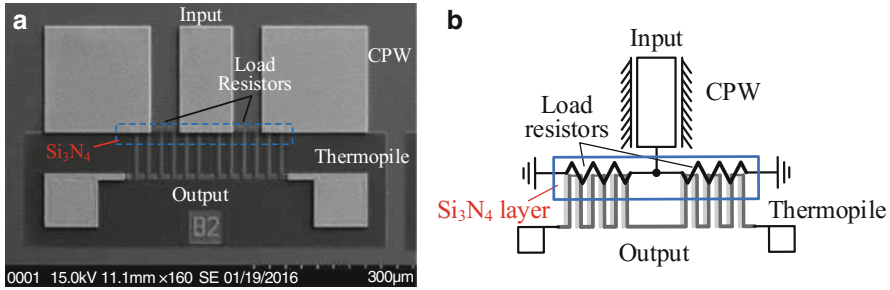
## Sandwich Structure

In order to achieve the low-cost power measurement, a sandwich non-MEMS thermoelectric microwave power sensor is presented (Zhang et al. 2016). For all of the previously reported sensors, there is a distance between the resistors and the thermopile, and the heat is transferred to the hot junctions of the thermopile through the substrate. So the MEMS technology is adopted to fabricate the substrate membrane for reducing thermal losses of the substrate, resulting in the high sensitivity. As you know, every coin has two sides. The MEMS technology increases the complexity and cost of the process compared to the traditional integrated circuit

**Fig. 12** Wireless input structure of indirect-heating MEMS microwave power sensor







**Fig. 13** Sandwich structure of indirect-heating non-MEMS microwave power sensors based on GaAs MMIC. (a) A SEM view and (b) a circuit diagram

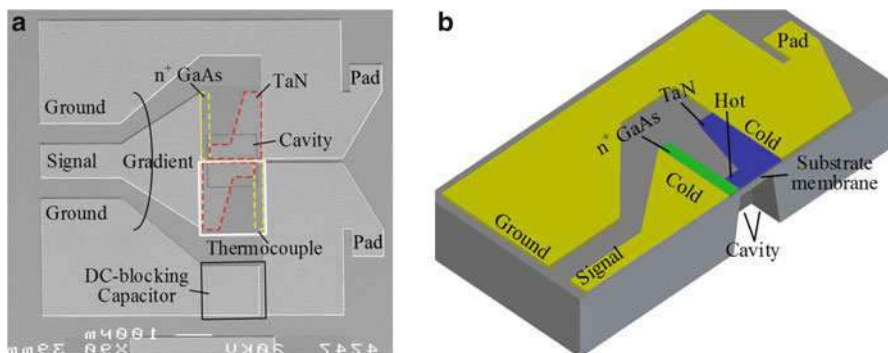
process. It will be preferred if a design meets requirements of both high sensitivity and without the MEMS technology. To achieve this goal, the sandwich non-MEMS thermoelectric microwave power sensor is studied. Figure 13 shows the sandwich structure of the non-MEMS thermoelectric microwave power sensor. In Fig. 13, the bottom layer is the thermopile in the role of heat-electricity conversion, the middle is the  $\text{Si}_3\text{N}_4$  dielectric layer for the isolation, and the top is the resistors in the role of microwave power-heat conversion. It should be noted that there is an overlapping distance between the resistors and the hot junctions of the thermopile. Because the dielectric layer is thin, it shows barely effects on the heat transfer in the vertical direction. As a result, the heat generated by the resistors is passed directly to the hot junctions of the thermopile, which brings the increase of the temperature difference. Therefore, this sandwich sensor does not utilize the MEMS technology but is able to achieve the good sensitivity. Experiments demonstrate that the sandwich sensor without the non-MEMS technology shows higher sensitivity and lower cost than its counterpart with the MEMS technology.

## A Terminal Self-Heating MEMS Power Sensor

The section introduces self-heating MEMS power sensors. Here, the so-called self-heating sensors indicate that they have no separate resistors, where the internal resistance of the thermocouples is used to dissipate the microwave power. In general, the number of the thermocouples is two. Prior to this, some self-heating MEMS power sensors have been reported using nonstandard Si process (Milanovic et al. 1996, 1998; Jackson 1974). Here, we will focus on the design of self-heating power sensor based on the GaAs MMIC process. Compared to Si, the GaAs substrate typically has the higher resistivity and mobility, which are more suitable for microwave devices.

Figure 14 shows schematic views of the self-heating MEMS microwave power sensor based on the GaAs MMIC process. It mainly includes a CPW line, two thermocouples, a dc-blocking capacitor, and two output pads. The output port of the CPW is designed to be much larger than the input port through widening the CPW





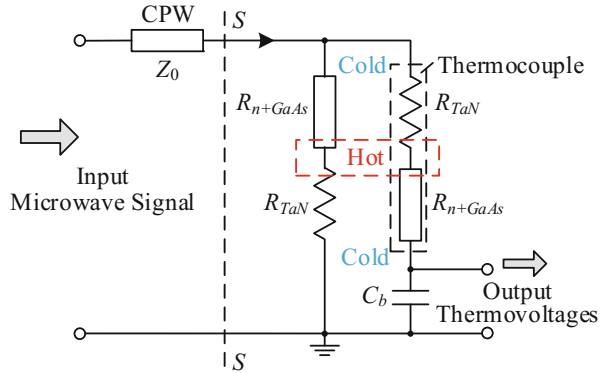
**Fig. 14** Schematic views of the self-heating MEMS microwave power sensor. (a) A structural view and (b) a cross-sectional view

signal line. The larger output port is used to increase the area of the cold junctions of the thermocouples, which serve as a heat sink. The CPW characteristic impedance is  $50 \Omega$ . Each thermocouple is made of TaN and  $n^+$  GaAs and designed to be  $100 \Omega$  in resistance. The two thermocouples are placed centrosymmetrically and their parallel resistance matches with the CPW. The hot and cold junctions locate in the middle and at both ends of the thermocouples, respectively. And the two hot junctions are designed to become tapered for accumulating heat. The pads are connected with the CPW ground lines and used to output dc thermovoltages. The dc-blocking capacitor is connected in series to a loop of the two output pads. It prevents the pads from shorting when the power sensor is operating. Figure 15 shows a lumped equivalent circuit of the self-heating sensor based on the GaAs MMIC process. The GaAs substrate underneath the hot junctions is etched to form a membrane structure, in order to reduce thermal losses. The principle of the self-heating MEMS microwave power sensor is that an input microwave power is dissipated to generate heat and converted into an output thermovoltage by the two thermocouples, based on Seebeck effect. In the sensor, two thermocouples not only are as an element of the microwave power to heat conversion, but also serve as an element of the heat to electricity conversion. Using the theory, the self-heating power sensor is designed at 8–12 GHz, and the measured results verify the design validity, with low reflection loss, good linearity, and high sensitivity (Zhang and Liao 2012b).

## An Inline Coupling MEMS Power Sensor

In the section “A Terminal Indirect-Heating MEMS Power Sensor,” several terminal MEMS microwave power sensors are described. Their largest weakness is that the input microwave signal is not available after the power measurement. That is, the entire microwave signal is used to measure the power. From the beginning of this section, including the following two sections, we will introduce three kinds of inline microwave power sensors (coupling, capacitive, and inserting) based on the MEMS

**Fig. 15** Lumped equivalent circuit of the self-heating MEMS microwave power sensor based on the GaAs MMIC process

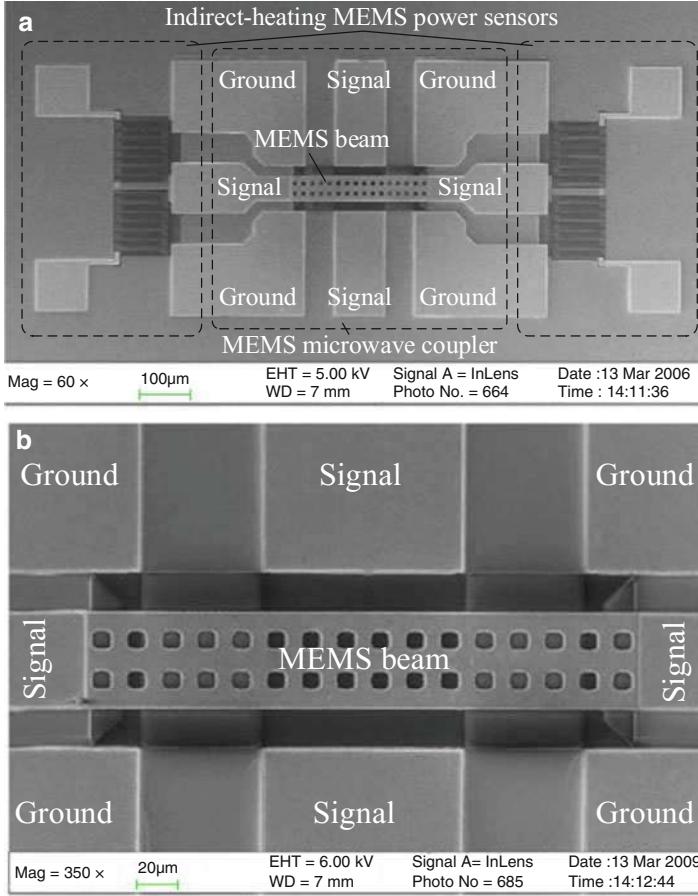


technology. The so-called inline sensors indicate that a small part of the microwave signal is used for the measurement and most of the signal is output to the next level circuit system. In a word, the interference to the signal caused by the measurement is small. These inline sensors can serve as implanted devices and be embedded into modern personal communication and radar systems for inline power monitoring.

This section introduces the inline coupling MEMS microwave power sensors. They are composed of two independent steps: the coupling step and the measurement step. In the coupling step, a certain percentage of the microwave power is coupled by a MEMS beam or a coupler; in the measurement step, the coupled power is measured by the terminal power sensors in section “[A Terminal Indirect-Heating MEMS Power Sensor.](#)” Using the two steps, it becomes possible for the power sensors to design the microwave performance and the sensitivity separately. Here, the microwave performance indicates reflection and insertion losses. It shows that the power sensors can achieve the low reflection and insertion losses and the high sensitivity, with small size and zero dc power consumption.

**Basic Structure**

The coupling MEMS microwave power sensors consist of a microwave power coupler and two indirect-heating microwave power sensors. Figure 16 shows SEM views of a basic coupling MEMS microwave power sensor (Han et al. 2007). The microwave coupler is similar to a capacitive shunt MEMS switch, where a MEMS beam is suspended above the CPW line. Two anchors of the beam are connected to signal lines of additional two CPW lines respectively, rather than the ground lines. The coupler is used to couple a certain percentage of the microwave power into two inputs of the indirect-heating power sensors by the suspended MEMS beam. The operation principle is that a certain percentage of the input microwave power is coupled by the MEMS beam and absorbed to generate heat by the load resistors, and converted into the thermovoltages by the thermopiles based on Seebeck effect.

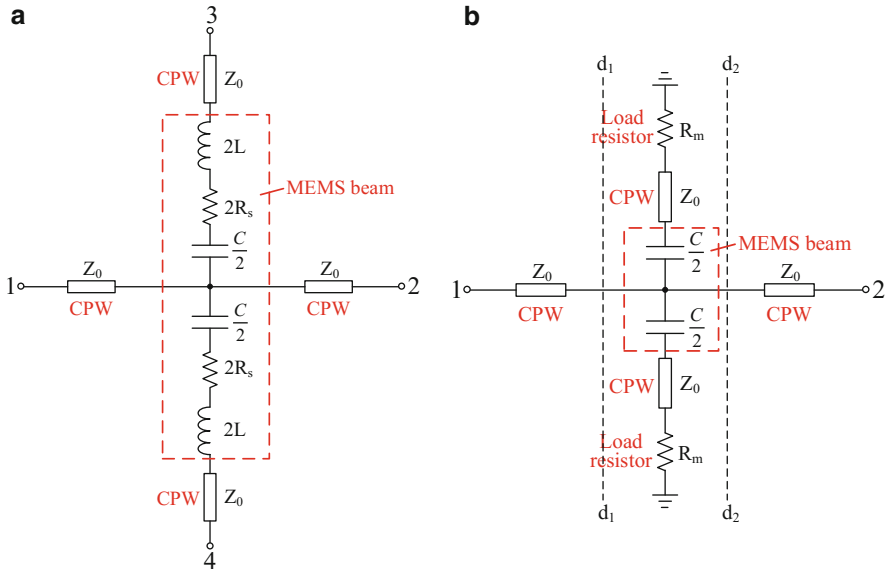


**Fig. 16** Basic structure of coupling MEMS microwave power sensor. (a) Whole view and (b) local view of MEMS beam

**Lumped Equivalent Circuit and S-Parameter**

Figure 17a shows a lumped equivalent circuit model of the microwave power coupler. Ports 1 and 2 are input and output ports and ports 3 and 4 are coupling ports. The impedance  $Z_0$  is the characteristic impedance of the CPW lines. The capacitance  $C$  represents the capacitance between the MEMS beam and the CPW signal line, which couples a certain percentage of the input microwave power. The capacitance  $C$  can be expressed as (Muldavain and Rebeiz 2000)

$$C = \frac{\epsilon_0 b w}{g_0 + \frac{t_d}{\epsilon_r}} + C_f \tag{6}$$



**Fig. 17** Lumped equivalent circuit model of (a) the power coupler and (b) the basic coupling power sensor

where  $b$  and  $g_0$  is the width and the initial height of the beam,  $w$  is the width of the CPW signal line,  $t_d$  is the thickness of the insulation layer,  $\epsilon_0$  and  $\epsilon_r$  are the permittivity of free space and the relative permittivity of the insulation layer, respectively, and  $C_f$  is the fringing field capacitance of the beam, which is about 20–50% of the total capacitance. The inductance  $L_s$  and resistance  $R_s$  indicate parasitic effects of the MEMS beam in the microwave frequency. The parasitic inductance  $L_s$  and resistance  $R_s$  of the beam can be written as (Ulm et al. 2000; Yue and Wong 2000)

$$L_s = 2s \left( \ln \left( \frac{2\pi s}{b} \right) - 1 + \frac{b}{\pi s} \right) \tag{7}$$

$$R_s = \begin{cases} \frac{\rho l}{bt} & (f < 10 \text{ GHz}) \\ \frac{\rho l}{b\delta(1 - e^{-t/\delta})} & (f > 10 \text{ GHz}) \end{cases} \tag{8}$$

where  $s$  is the distance between the edges of the CPW signal line and the anchors of the beam,  $\rho$  is the resistivity of the beam,  $l$  and  $t$  are the length and the thickness of the beam respectively, and  $\delta$  is the skin effect. And  $\delta$  is equal to  $(\rho/\pi\mu f)^{1/2}$ , where  $\mu$  and  $f$  are the permeability and frequency, respectively.

According to the microwave theory, S-parameters of the microwave power coupler can be deduced as

$$S_{11} = -\frac{Z_0}{2Z_0 + 2R_s + j2\left(\omega L - \frac{1}{\omega C}\right)} \quad (9.1)$$

$$S_{21} = \frac{Z_0 + 2R_s + j2\left(\omega L - \frac{1}{\omega C}\right)}{2Z_0 + 2R_s + j2\left(\omega L - \frac{1}{\omega C}\right)} \quad (9.2)$$

$$S_{31} = S_{41} = \frac{Z_0}{2Z_0 + 2R_s + j2\left(\omega L - \frac{1}{\omega C}\right)} \quad (9.3)$$

where  $S_{11}$  is the reflection coefficient seen at port 1 when port 2 is terminated by a matching load ( $Z_0 = 50 \Omega$ );  $S_{21}$  is the transmission coefficient from port 1 to port 2;  $S_{31}$  and  $S_{41}$  are coupling coefficients from port 1 to port 3 and from port 1 to port 4, respectively. The coupling coefficients describe the microwave power coupled by the MEMS beam. As seen in Eq. (9.3), the coupled microwave power is mainly dependent on the capacitance  $C$ . In addition, the coupled microwave power is proportional to the sensitivity. So a desirable sensitivity of the coupling power sensor can be realized by designing the MEMS beam.

### Design of MEMS Beam

When the microwave power is transmitted through the MEMS beam, an electrostatic force is induced on the MEMS beam. The electrostatic force makes the beam deflection, resulting in the change of the coupling capacitance between the beam and the CPW signal line. The effect of the electrostatic force on the MEMS beam can be expressed as (Rebeiz 2003)

$$V = (g_0 - z)\sqrt{\frac{2k_{s\text{-beam}}}{\epsilon_0 b w}} z \quad (10)$$

where  $V$  is the RMS voltage of the microwave signal, and  $z$  and  $k_{s\text{-beam}}$  are the downward deflection and spring coefficient of the beam, respectively. Assuming no residual stress of the beam, the spring coefficient  $k_{s\text{-beam}}$  is expressed as

$$k_{s\text{-beam}} = K_{s\text{-beam}} \frac{E t^3 w}{l^3} \quad (11)$$

where  $E$  is the Young's modulus, and  $K_{s\text{-beam}}$  is a constant and related to the position of the force evenly distributed on the beam. As can be seen from Eqs. (6), (10), and (11), a variation of the MEMS beam in width can change the capacitance between the beam and the CPW line, but can not cause the beam deflection. However, the decrease of the beam in length can greatly reduce the deflection at the certain RMS voltage. Therefore, the width of the beam is

designed to achieve the desirable sensitivity of the sensor, while the length of the beam is designed to be enough short for minimizing the effect of the microwave signal on the beam deflection. The short beam prevents the beam from moving downward due to the electrostatic forces.

**Model with Attenuation and Phase Mechanisms**

Once the coupling power sensors are embedded into the middle of RF circuits for the inline power measurement, the phase caused by the sensors must be considered. In addition to the CPW line, the MEMS beam also produces an additional phase. The phase is undesired for the inline coupling power sensors. On the other hand, based on the distributed theory of a CPW line loaded periodically with some MEMS beams, distributed MEMS phase shifters are designed by changing the height of the beams (Barker and Rebeiz 1998). The capacitance of the MEMS beam is responsible for the effective dielectric constant, which relies on the impedance and propagation velocity of the CPW line. Therefore, for the inline coupling power sensors, the suspended beam not only generates the loss of the microwave power but also results in the undesired phase shift after the power measurement. In order to analyze the phase change, this part provides the lumped model of the coupling power sensors (Zhang and Liao 2013a).

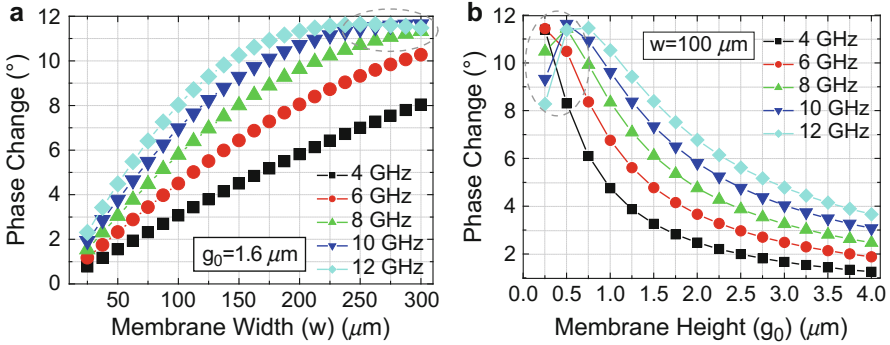
When the operation frequency of the power sensors is much smaller than the series resonant frequency of the MEMS beam, the parasitic inductance and resistance of the MEMS beam can be neglected. Fig. 17b shows a lumped equivalent circuit model of the coupling power sensor. The resistance  $R_m$  of the matching load resistors is added to study the loss and phase of the sensor. As observed in Fig. 17b, the equivalent circuit is described by a two-port lumped network and divided into three parts by the dashed  $d_1-d_1$  and  $d_2-d_2$ . So the normalized  $ABCD$  matrix of the power sensors between port 1 and port 2 can be obtained as

$$\begin{aligned} \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{1-2} &= \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{1-d_1} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{d_1-d_2} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{d_2-2} \\ &= \begin{bmatrix} \cosh(rl) & \sinh(rl) \\ \sinh(rl) & \cosh(rl) \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \frac{2Z_0}{j\omega C} + Z_0 + \frac{R_m}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} \cosh(rl) & \sinh(rl) \\ \sinh(rl) & \cosh(rl) \end{bmatrix} \end{aligned} \tag{12}$$

with

$$r = \alpha + j\beta = (\alpha_c + \alpha_d) + j \frac{2\pi\sqrt{\epsilon_{\text{eff}}}}{c_0} / f \tag{13}$$

where  $r$  is the propagation constant,  $\alpha$  and  $\beta$  are the attenuation constant and the phase constant respectively,  $\omega = 2\pi f$  is the angular frequency,  $\alpha_c$  is the conductor loss of the CPW line,  $\alpha_d$  is the dielectric loss of the GaAs substrate,  $c_0$  is the free space velocity, and  $\epsilon_{\text{eff}}$  is the effective dielectric constant of the CPW line.



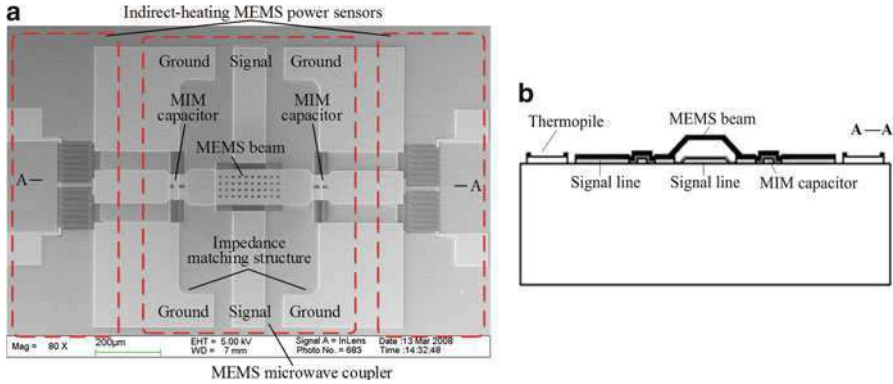
**Fig. 18** Phase change as a function of the height ( $g_0$ ) and the width ( $w$ ) of the MEMS beam at different frequencies

Using the conversion relationship of the  $ABCD$  matrix and the scattering matrix,  $S$ -parameters of the inline MEMS power sensors between port 1 and port 2 can be derived as using Eq. (12)

$$S_{11} = \frac{2\omega CZ_0(\sinh(2rl) - \cosh(2rl))}{\omega C(4Z_0 + R_m) - j4} \quad (14.1)$$

$$S_{21} = \frac{\omega C(2Z_0 + R_m) - j4}{(\omega C(4Z_0 + R_m) - j4)(\sinh(rl) + \cosh(rl))^2} \quad (14.2)$$

In Eq. (14), the derived  $S$ -parameters refer to the loss and phase of the power sensors. Based on the equation, the effects of the MEMS beam on the phase can be studied. Figure 18 shows the phase change as a function of the height and width of the beam at different frequencies. The length of the beam is constant. In Fig. 18a, for the same width, the phase shift that is caused by the beam increases with the decrease of the beam's height. In Fig. 18b, for the same height, the phase shift increases with the increase of the beam's width. It shows that the capacitance of the MEMS beam generates the phase change. However, the slight decrease of the phase change occurs when the capacitance increases to a certain value by changing the height and width of the beam. The higher the microwave frequency is, the earlier the phenomenon appears. If the impedance ( $Z_0$ ) and resistance ( $R_m$ ) in the branches are removed and the capacitance ( $C$ ) is directly grounded in Fig. 17, the phase change will linearly increase with the increase of the capacitance. Therefore, the decrease of the phase change results from the impedance ( $Z_0$ ) and the resistance ( $R_m$ ) connected to the anchors of the beam, as the beam is beyond 0.5 μm in height (Fig. 18a) and 250 μm in width (Fig. 18b).



**Fig. 19** Views of the coupling MEMS microwave power sensor with the MIM capacitors. (a) SEM view and (b) cross-sectional schematic view along A-A

### Wideband Structure

In order to achieve the wideband response, two kinds of capacitance compensating structures are proposed based on the basic power sensors. The one, two metal-insulator-metal (MIM) capacitors are added (Han et al. 2009); the other, four open-circuit transmission lines are utilized to compensate the capacitance (Zhang et al. 2010). In the two structures, the impedance matching method is used to reduce the effect of the beam on the reflection and insertion losses. The method is done by modifying the gap size of the CPW line before and after the MEMS beam.

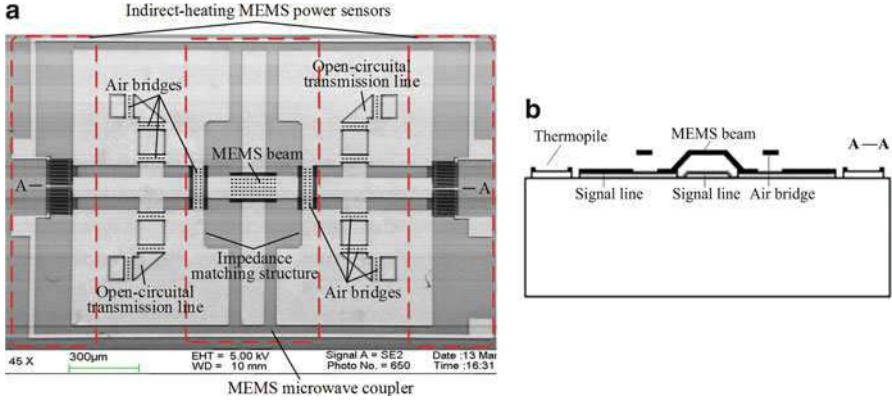
Figure 19 shows views of the coupling MEMS microwave power sensor with the MIM capacitors. In Fig. 19, the two MIM capacitors are same and added after the anchors of the beam. For the MIM capacitors, the bottom layer is connected to the CPW ground line, the top layer is the CPW signal line in the coupling branches, and the middle is the Si<sub>3</sub>N<sub>4</sub> layer. The compensating capacitance ( $C_{MIM}$ ) of the MIM capacitors can be written as

$$C_{MIM} = \frac{\epsilon_0 \epsilon_r b' w'}{t_d} \tag{15}$$

where  $b'$  and  $w'$  are the width of the bottom layer and the length of the top layer, respectively.

Because the compensating capacitance is on the order of  $fF$ , the small capacitance for the MIM capacitors to a certain extent depends on the fabrication process. It is very likely to cause a significant deviation between the fabricated and designed capacitance, resulting in the deterioration of the microwave performance. According to the theory of the transmission line, an open-circuit transmission line can be equivalent to a capacitor. The capacitance compensated by the open-circuit transmission line can be minimally dependent on the process. Figure 20 shows views of the coupling MEMS microwave power sensor with four open-circuit transmission





**Fig. 20** Views of the coupling MEMS microwave power sensor with the open-circuit transmission lines. (a) SEM view and (b) cross-sectional schematic view along A-A

lines. In Fig. 20, the open-circuit transmission lines are placed after the anchors of the beam and connected to the CPW signal lines in the coupling branches. The port impedance ( $Z_{in}$ ) of the open-circuit transmission lines can be written as

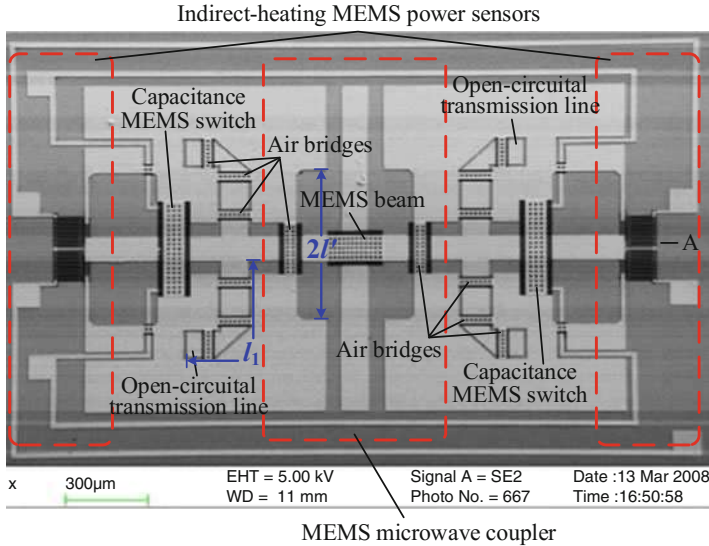
$$\begin{aligned}
 Z_{in}(l_1) &= \frac{Z_1}{j \tan(\beta l_1)} \\
 &= \begin{cases} < 0 & \text{capacitance} & (0 < l_1 < \lambda/4) \\ > 0 & \text{inductance} & (\lambda/4 < l_1 < \lambda/2) \end{cases} \quad (16)
 \end{aligned}$$

where  $Z_1$  and  $l_1$  are the impedance and length of the open-circuit transmission lines, respectively, and  $\lambda$  is the wavelength. As can be seen from Eq. (16), the open-circuit transmission lines are equivalent to capacitors, when the length of the open-circuit transmission lines is less than a quarter wavelength; the open-circuit transmission lines are equivalent to inductors, when the length of the open-circuit transmission lines is between one-quarter wavelength and one-half wavelength.

On the basis of the basic structure in section “Basic Structure,” two lumped equivalent circuit models of the power sensors with the MIM capacitors and the open-circuit transmission lines are established (Han et al. 2009; Zhang et al. 2010). Using the lumped networks, S-parameters of the two improved power sensors can be obtained respectively. Furthermore, experiments demonstrate the design validity of the wideband structures.

## State Conversion Control

In order to achieve the state conversion control, the inline coupling MEMS microwave power sensor with the detection and nondetection function is presented (Zhang et al. 2011). It solves a problem that regardless of whether the sensor needed to



**Fig. 21** SEM view of the coupling MEMS microwave power sensor with the open-circuitual transmission lines and the MEMS switches

measure the power, a part of microwave power always is detected. This results in the unnecessary power loss. In the design, two shunt capacitive MEMS switches are added to the power sensor with the open-circuitual transmission lines.

Figure 21 shows a SEM view of the coupling MEMS microwave power sensor with the open-circuitual transmission lines and the MEMS switches. When the actuation voltage is not applied to the switches, the switches are in the up state and the sensor is in the detection state. It means that a certain percentage of the input microwave power coupled by the MEMS beam is dissipated by the load resistors for the microwave power measurement. When the actuation voltage is applied to the switches, the switches are in the down state and the sensor is in the nondetection state. It means that a certain percentage of the input microwave power coupled by the beam is reflected and not dissipated by the load resistors. Thus, the total microwave power is not almost attenuated. In addition, in order to reduce the effect of the switches on the microwave performance, the impedance matching structures are needed by modifying the CPW gap near the switches. The MEMS switches based on the GaAs MMIC process have been designed (Zheng et al. 2005). In Fig. 21,  $l_1$  and  $l'$  are the lengths of the open-circuitual transmission lines and the uneven CPW line before and after the MEMS beam, respectively.

### Design of MEMS Switches

When an actuation voltage is applied the shunt capacitive MEMS switches, the electrostatic force is induced on the beam. If the actuation voltage is equal to the threshold voltage of the switches, the beam of the switches will be absorbed to

the electrodes beneath the switches. The actuation voltage that makes the switches from the up state to the down state is obtained as (Rebeiz 2003)

$$V_{\text{th}} = \sqrt{\frac{4k_{s\text{-switch}}}{27\varepsilon_0 a_{\text{switch}} w_{\text{switch}}} g_{0\text{-switch}}^3} \quad (17)$$

where  $V_{\text{th}}$  is the applied actuation voltage,  $w_{\text{switch}}$  and  $g_{0\text{-switch}}$  are the width and initial height of the switches respectively,  $a_{\text{switch}}$  is half the width of the CPW signal line under the switches, and  $k_{s\text{-switch}}$  is the spring constant of the beam in the switches. Assuming no residual stress of the membrane, the spring constant  $k_{s\text{-switch}}$  is expressed as

$$k_{s\text{-switch}} = K_{s\text{-switch}} \frac{E t_{\text{switch}}^3 w_{\text{switch}}}{l_{\text{switch}}^3} \quad (18)$$

where  $t_{\text{switch}}$  and  $l_{\text{switch}}$  are the thickness and length of the switches respectively, and  $K_{s\text{-switch}}$  is a constant coefficient and dependent on the location of the evenly distributed force.

As can be observed in Eqs. (17) and (18), the actuation voltage ( $V_{\text{th}}$ ) and the spring constant ( $k_{s\text{-switch}}$ ) of the beam for the switches are a function of the width ( $w_{\text{switch}}$ ), but the actuation voltage ( $V_{\text{th}}$ ) is independent on the width ( $w_{\text{switch}}$ ). This is because that the spring constant ( $k_{s\text{-switch}}$ ) varies linearly with the width ( $w_{\text{switch}}$ ). However, increasing the length of the beam can significantly reduce the actuation voltage. Therefore, for the MEMS switches, the length of the beam is designed to be somewhat long to reduce the actuation voltage, thereby reducing the loss, but not too long to avoid the stiction problem in the releasing process.

## Design Consideration

As for the design of the beams in the coupler and the switches, the following conclusions are drawn. (1) In the microwave power coupler, the MEMS beam must be designed to be “stiff” through appropriately reducing the length of the beam, so as to obtain the constant coupling power. It means that the beam with the smaller length is not easily affected by the electrostatic force induced by the microwave power. (2) In the coupler, the desirable sensitivity of the inline power sensor can be realized through designing the width of the MEMS beam. (3) In the shunt capacitive MEMS switches, the beams are designed to be “soft” through appropriately increasing the length of the beams in the switches, so that the switches are in the closed state. The beams have a small actuation voltage, but not too long to avoid sticking problem.

To show the operation principle and design theory, some simulations of the power sensor are performed in HFSS (Zhang et al. 2011), for example, reflection losses versus uneven lengths of the MEMS beam ( $l'$ ), coupling coefficients versus lengths of the open-circuit transmission lines ( $l_1$ ), and insertion losses in detection and nondetection states. The simulations demonstrate the design validity of the impedance matching structure, compensating capacitance structure, and switch state conversion structure.

## Fabrication and Packaging

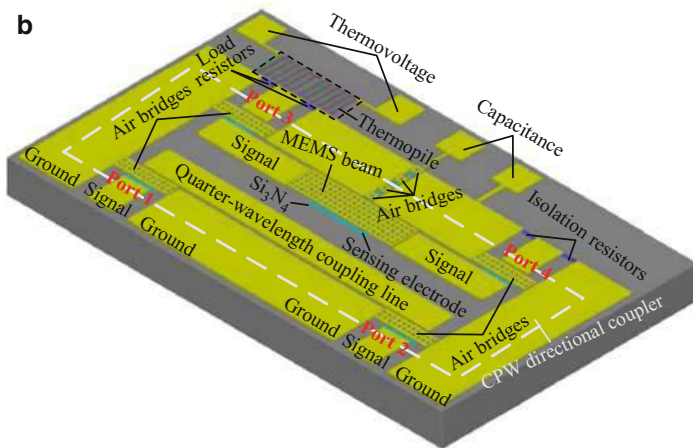
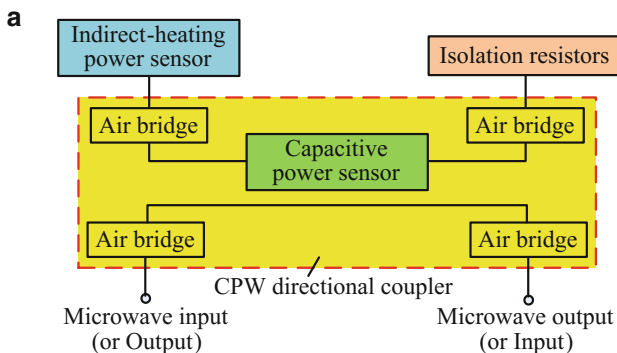
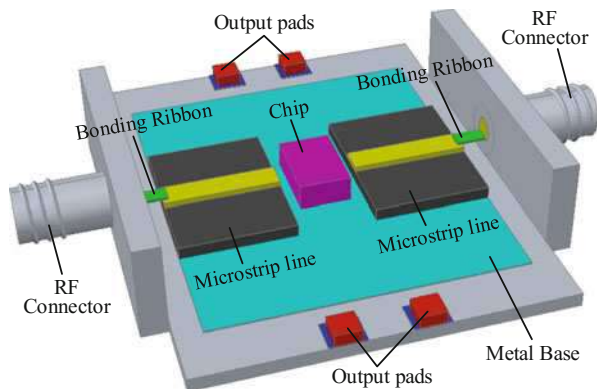
These coupling MEMS microwave power sensors are fabricated by using the GaAs MMIC process. In the sensors, the thermopiles are made of AuGeNi/Au and  $n^+$  GaAs; the beams are made of Au, with the thickness of 2  $\mu\text{m}$ ; the load resistors are made of TaN, with the square resistance of 25  $\Omega/\square$ ; the dielectric layers are made of  $\text{Si}_3\text{N}_4$ , with the thickness of 0.1  $\mu\text{m}$ ; the sacrificial layers are made of polyimide, with the height of 1.6  $\mu\text{m}$ . The fabrication details are discussed (Zhang and Liao 2012c). However, the MEMS beams are subject to different levels of the collapse after the removal of the sacrificial layer, which leads to a performance degradation or failure of the sensor. To solve the problem, the inline power sensor based on the fixed capacitive coupling is developed (Zhang and Liao 2015b). In this sensor, the beam is not required to release for obtaining a fixed capacitance and shows better reliability. Meanwhile, it can perform high power measurement and increase the dynamic range. The material under the beam is PMGI.

In order to realize applications in real-life microwave systems, a packaging-test-fixture method for the microwave power sensors is proposed (Zhang and Liao 2011). The packaging-test-fixture is assembled by the transition of the CPW-port power sensors, two microstrip lines, and two standard connectors. Figure 22 shows a schematic view of the packaging-test-fixture. Measurements show the good microwave performance and sensitivity before and after the packaging. After the packaging, the temperature and humidity effects are studied, and the rise and fall response times (< 10 ms) and the negligible phase noise are obtained for the coupling power sensors (Zhang and Liao 2013b).

## Transmission Direction Detection

In order to detect the transmission direction of microwave power, a directional inline MEMS power sensor is presented (Zhang and Liao 2015c). Figure 23 show a simplified block diagram and schematic view of the directional power sensor. The power sensor is mainly composed of a CPW directional coupler, two isolation resistors, a capacitive power sensor, and an indirect-heating power sensor. The coupler is used to extract a certain percentage of the microwave power, with the directivity. If the ports 1 and 2 are input and output ports, the ports 3 and 4 will be coupling and isolation ports, respectively. The indirect-heating sensor is connected to the port 3 of the coupler and the isolation resistors are matched to the port 4, in order to measure the transmission direction of the known power level. The capacitive sensor based on a MEMS beam is added to the coupling section in the detection branch, in order to improve the high power measurement range. The cooperation of the indirect-heating and capacitive sensors can detect the transfer direction of the unknown signal. In operation, the two sensors does not interfere whit each other. The air bridges are used to achieve the interconnection of separated ground lines instead of bonding lines. The air bridges generate the additional capacitance, so the effects of the four bridges at the corners on the performance of the coupler must be taken into account.

**Fig. 22** Schematic view of the packaging-test-fixture



**Fig. 23** (a) Simplified block diagram and (b) schematic view of the directional MEMS microwave power sensor

## An Inline Capacitive MEMS Power Sensor

Capacitive MEMS power sensors have been already mentioned in the previous sections. The sensors are based on the principle of microwave power-force-electricity conversion. They are different from thermoelectric sensors, independent of temperature. In structure, they are similar to MEMS shunt capacitive switches, where a MEMS beam is suspended over a CPW. The beam may be a tethered beam or a cantilever beam. In the sensors, the sensing electrodes are placed under the beam. Their principles are that: the microwave power transmitted on the CPW passes through the MEMS beam, and an electrostatic force is induced on the beam; the electrostatic force makes the beam deflection, resulting in the capacitance change between the beam and the sensing electrodes; thus, the microwave power can be obtained by measuring the capacitance. It should be noted that most of the microwave power are output and available after the measurement. So the capacitive sensors are also called inline devices. Fernandez et al. presented the capacitive MEMS microwave power sensors on the glass substrate (Fernández et al. 2003, 2005, 2006). This section mainly introduces the capacitive MEMS microwave power sensors based on the GaAs MMIC technology.

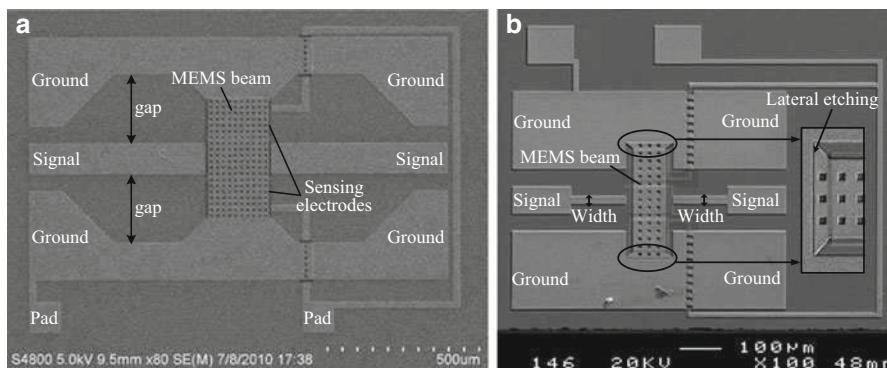
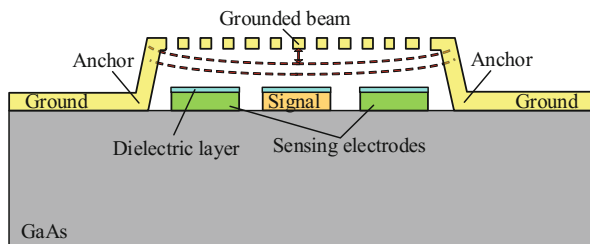
### Tethered Beam

In the capacitive sensors, the beam suspended over the CPW is the fixed-fixed beam. There are two connection methods for two anchors of the beam. One is that the anchors are connected to the CPW ground lines; the other is that the anchors are connected to floating metal blocks. For the two methods, there are different requirements of capacitance testing equipment (Fernández et al. 2006). In addition, the initial capacitance of the beam affects the characteristic impedance of the CPW. So the impedance matching structure has to be designed by modifying the width of the CPW gap or signal line before and after the beam.

### Grounded Anchors

Figure 24 show a cross-sectional view of capacitive MEMS power sensors based on the grounded anchors. Figure 25 shows SEM of the capacitive MEMS power sensors based on the grounded anchors. They consist of a CPW, a beam, two sensing electrodes, and a  $\text{Si}_3\text{N}_4$  dielectric layer. The CPW, beam, and electrodes are made of Au. The dielectric layer is fabricated on the sensing electrodes to avoid direct metal-metal contact. To demonstrate the linear relationship between the microwave power and the capacitance change, a 2D model of the sensitivity is established (Cui and Liao 2012). In Fig. 25a, in order to reduce the effect of the beam on the reflection losses, the impedance matching structure is designed by increasing the width of the CPW gap before and after the beam. In Fig. 25b, the impedance matching structure is achieved by reducing the width of the CPW signal line (Han and Liao 2014). In Fig. 25b, a lateral etching technique is used on the anchors to reduce the spring constant of the beam and improve the sensitivity.

**Fig. 24** Cross-sectional view of capacitive MEMS power sensors based on the grounded anchors



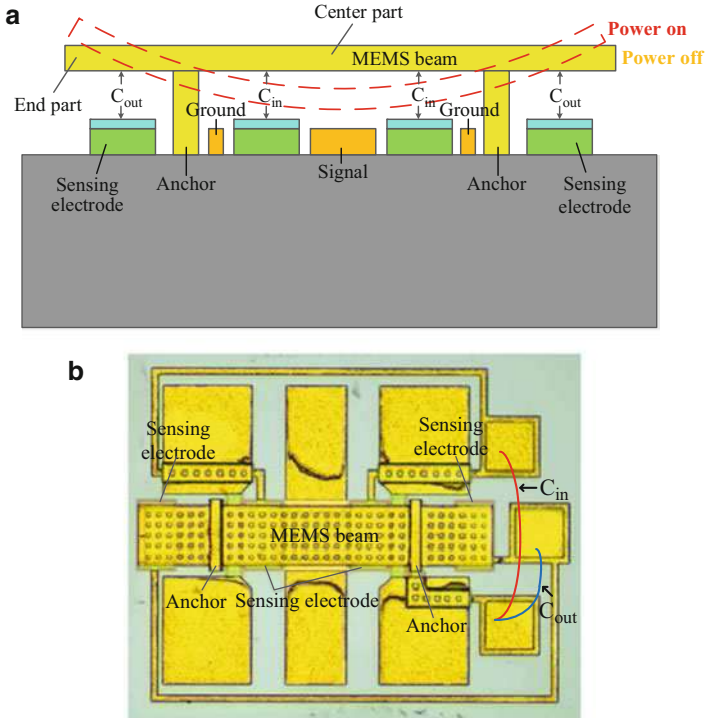
**Fig. 25** SEM of capacitive MEMS power sensors based on the grounded anchors. (a) Increasing the CPW gap and (b) reducing the width of the signal line

It is worth noting that when the small microwave power is applied to the capacitive sensors, there is the approximately linear relationship between the capacitance change and the microwave power. However, if the high microwave power is applied, the power will make the beam move into the nonlinear region or even generate the self-excitation phenomenon. Experiment shows that the relationship between the capacitance change and the microwave power is divided into three regions: linear region (60–200 mW), transitional region (200–640 mW), and saturated region (640–1000 mW) (Yan and Liao 2015). In the capacitive sensors, the movement of the MEMS beam leads to the change in capacitance. Such structure will generate a nonlinear effect of the intermodulation distortion. Measured results show small intermodulation distortions (Han and Liao 2015). For example, when two input microwave signals are the power of 6 dBm, the third-order intercept point is about 29.3 dBm at  $\Delta f = 100$  Hz ( $f_1 = 10$  GHz).

### Floating Anchors

The structure of the capacitive power sensors based on the floating anchors is similar to Fig. 26a. In order to achieve the accurate measurement of the capacitance using the balanced-bridge impedance technique, the anchors of the MEMS beam are in the floating state and not connected to the CPW ground lines. Because the beam is not designed to be too long for avoiding collapse, the two anchors are inside the CPW





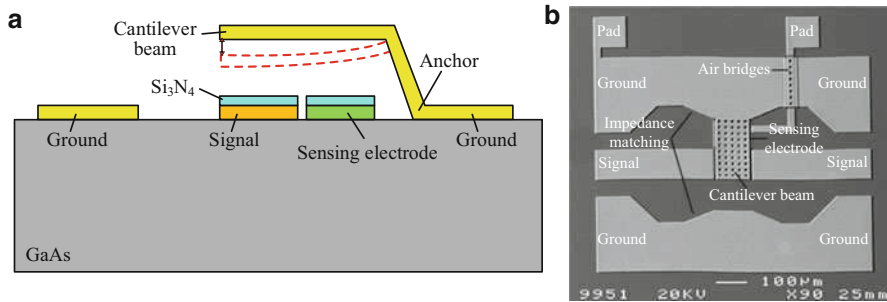
**Fig. 26** (a) Cross-sectional view and (b) SEM of the seesaw-type capacitive MEMS power sensor based on the floating anchors

ground lines. The length of the beam is 400  $\mu\text{m}$ . Each anchor is divided into two parts along the middle. Such anchors can reduce the residual stress of the beam and increase the sensitivity, but also facilitate the design of connection lines between the sensing electrodes and the pads (Zhang and Liao 2014). Following that, a seesaw-type capacitive MEMS X-band power sensor is proposed based on the floating anchors (Yi et al. 2016). Figure 26 shows a cross-sectional view and SEM of the seesaw-type capacitive MEMS power sensor. When the microwave power is applied, the center part of the MEMS beam is pulled down, but two ends of the beam are lifted up. Thus, the capacitance changes include two parts:  $C_{in}$  and  $C_{out}$ . So the sensitivity of the power sensor is the sum of the two parts.

### Cantilever Beam

Here, the MEMS beam in the capacitive power sensor is a cantilever beam. For the cantilever beam, one end is fixed and the other is free. Compared to the tethered beam, the spring constant in the cantilever beam is not affected by the residual stress. In this way, the capacitive power sensor using the cantilever beam is easier to obtain





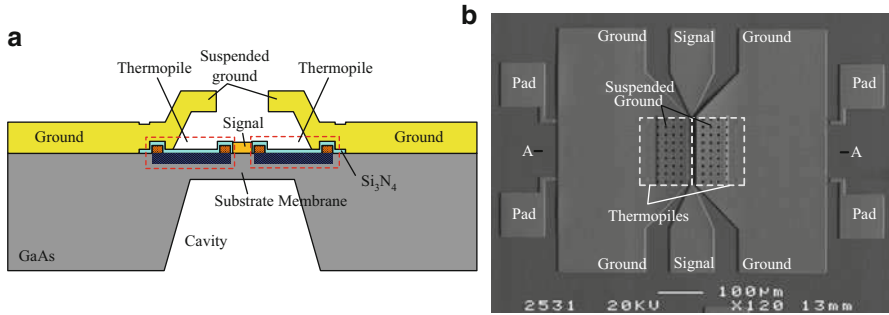
**Fig. 27** (a) Cross-sectional and (b) SEM views of the capacitive MEMS power sensor using the cantilever beam

higher sensitivity in theory. The capacitive power sensor using the cantilever beam has the same principle as that using the tethered beam. Furthermore, they show similar structures. Figure 27 shows a cross-sectional view and SEM of the capacitive MEMS power sensor using the cantilever beam (Yi and Liao 2013). The capacitive sensor with the cantilever beam generates small intermodulation distortion when the frequency difference of two microwave signals exceeds 200 kHz (Yi and Liao 2014).

## An Inline Inserting MEMS Power Sensor

The inline inserting MEMS sensors are based on the thermoelectric principle, but the part of the microwave power that is used to measure is different from that of the coupling sensors in section “An Inline Coupling MEMS Power Sensor.” The inserting sensors utilize a fraction of the unavoidable power caused by the intrinsic ohmic loss of the CPW, while the coupling sensors actively couple a portion of the power through some structures for the measurement. The inserting power sensors were initially presented by Dehé et al. (1995, 2002). Following that, their optimization is further studied based on the GaAs MMIC technology to achieve high sensitivity and low reflection losses (Zhang and Liao 2015d). Because the thermopiles are cleverly placed inside the CPW, the inserting sensors occupy smaller chip areas compared to the coupling sensors. This conforms to the development of miniaturization. The principle of these inserting MEMS power sensors is that a certain amount of the input microwave power is dissipated to generate heat due to the ohmic loss of the CPW, resulting in the temperature distribution, and converted into the thermovoltages by the thermopiles based on Seebeck effect. Figure 28 shows SEM and cross-sectional views of the improved inserting MEMS microwave sensor. The microwave signal is input from one port of the CPW and output from the other end.

In Fig. 28, the CPW have the characteristic impedance of  $50 \Omega$  in interested microwave frequencies. The signal line of the CPW in middle is designed to be thin and narrow to appropriately increase the ohmic loss. The two ground lines in middle



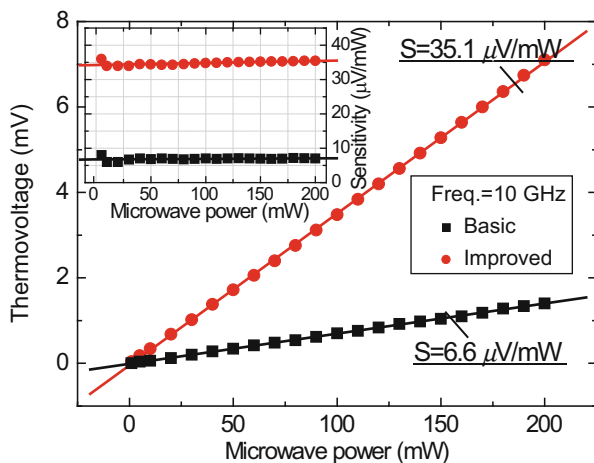
**Fig. 28** (a) Cross-sectional view along the A-A direction and (b) SEM view of the improved inserting MEMS microwave power sensor

of the CPW are designed as suspension structures, beneath which the two thermopiles are placed respectively. In the design of the thermopiles, the cold junctions are covered with the CPW ground planes to serve as a hot sink while the hot junctions lie below the CPW signal line to increase their temperature. The design of the thermopiles increases the temperature difference of the cold and hot junctions, resulting in the improvement of the sensitivity. The pads are used to output thermovoltages. The thermopiles and pads are connected by the connection lines. The thermopiles and the CPW, as well as the connection lines and the CPW ground, are separated by a  $Si_3N_4$  dielectric layer. To reduce thermal losses of the substrate, the GaAs substrate underneath the hot junctions and the narrow CPW signal line is etched to form a back cavity. Figure 29 shows measured results of two inserting MEMS microwave sensors. For the basic sensor, the hot junctions of the relevant thermopiles are along the fringe on each side of the suspended ground lines. It means that there is some distance between the hot junctions and the signal line. For the improved sensor, the hot junctions lie below the signal line. It means that the hot junctions contact with the signal line. Other structures and parameters are same. In Fig. 29, the good linearity between the thermovoltages and the microwave power is obtained. The inset figure shows approximate horizontal linear relationships between the sensitivity and the power. As observed, the improving sensor results in higher thermovoltages and sensitivities than the basic sensor, which verifies the design validity.

### Extension of MEMS Power Sensor

As is well known, the complete information of a microwave signal includes power, phase, and frequency. Moreover, the measurement of microwave phase and frequency is of great importance to phase-controlled radar, electronic warfare, and other applications. Based on the previously described microwave power sensors, we can design fully passive MEMS microwave phase and frequency detectors. These phase and frequency detectors can measure the microwave phase and frequency respectively, being compatible with the GaAs MMIC technology.

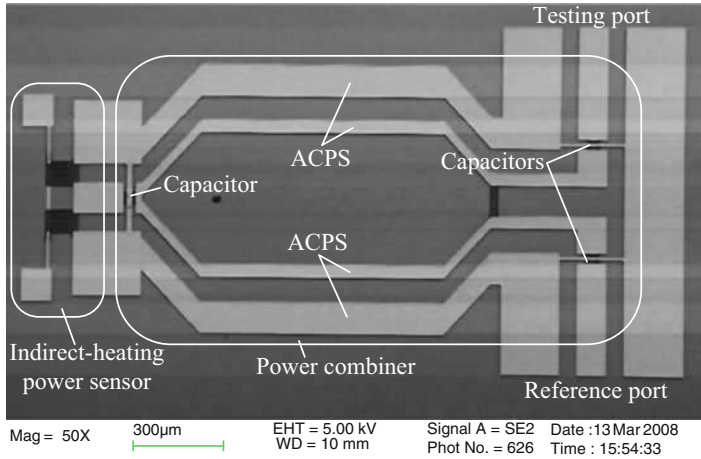
**Fig. 29** Measured thermovoltages and sensitivities versus the microwave power for the inserting MEMS microwave power sensors



## A MEMS Microwave Phase Detector

In order to achieve the microwave phase measurement, a single-channel MEMS microwave phase detector is firstly presented based on the theory of vector synthesis (Hua et al. 2011). The phase detector mainly consists of a reference signal port, a testing signal port, a microwave power combiner, and an indirect-heating power sensor. Figure 30 shows a SEM view of the single-channel MEMS microwave phase detector. When the reference and testing signals with the same power and different phases are applied to the respective ports, the two signals are combined into a synthetic signal by the power combiner. The microwave power of the synthetic signal that is converted into the thermovoltage is measured by the indirect-heating power sensor. Based on the vector synthesis, the proportional relationship between the synthetic power and the phase shift is derived. The phase shift is the phase difference between the reference and testing signals. Therefore, the microwave phase of the testing signal is known by measuring the thermovoltage. To reduce the size of the power combiner with quarter-wavelength length, three compensation capacitors are loaded at ends of asymmetrical coplanar striplines (ACPS) (Hua et al. 2013). Because the bandwidth limit of the power combiner, the phase detector is a narrow-band device. Following that, a broadband microwave phase detector is proposed using the two cascaded ACPS (Han and Liao 2016b). The broadband detector operates over the entire X-band. These single-channel phase detectors achieve the  $-180^\circ \sim 180^\circ$  phase of symmetrical measurements. It means that they only realize a half-cycle measurement and can not distinguish the phase lead and lag.

To solve the problem of the single-channel detectors, a dual-channel microwave phase detector is proposed at X-band (Yan et al. 2016). Figure 31 shows a SEM view of the dual-channel MEMS microwave phase detector. The detector has two measurement channels, and the synthetic power at each channel is measured by indirect-heating and capacitive MEMS power sensors. The  $45^\circ$  leading and lagging phase

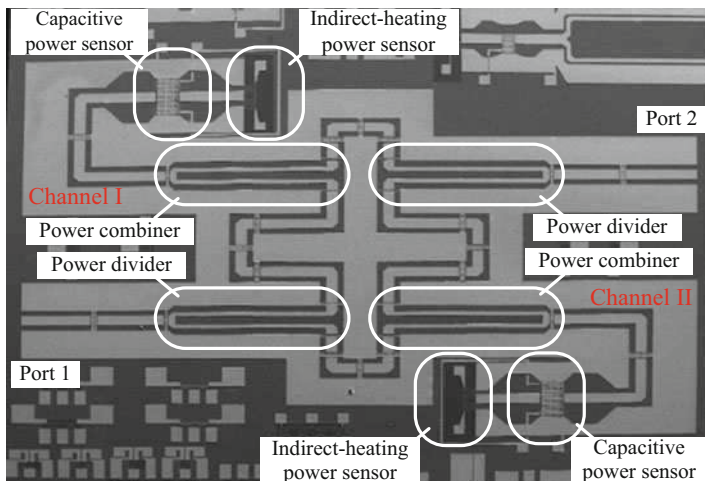


**Fig. 30** SEM view of single-channel MEMS microwave phase detector

shifters are added to the phase difference between the reference and testing signals. The design of the dual channels and the  $45^\circ$  leading and lagging phases can expand the measurable phase range to the entire cycle ( $-180^\circ \sim 180^\circ$ ) by using two inverse trigonometric functions. The capacitive power sensors are placed in front of the indirect-heating power sensors. Two types of the power sensors can enhance the dynamic range of power measurement.

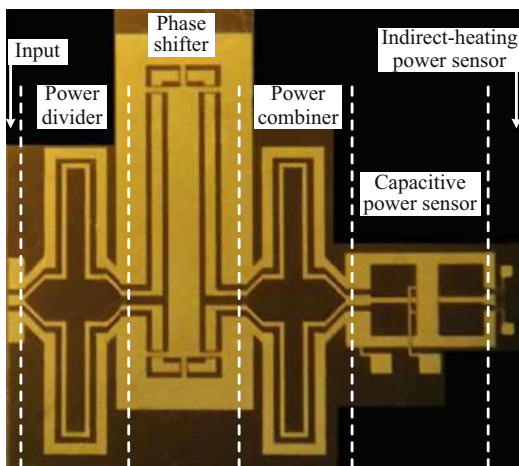
## A MEMS Microwave Frequency Detector

In order to achieve the microwave frequency measurement, a MEMS microwave frequency detector is presented (Yi and Liao 2012; Hua et al. 2009). The frequency detector is similar to the phase detector, and mainly composed of a Wilkinson power divider, a power combiner, a  $90^\circ$  phase shifter, and indirect-heating and capacitive MEMS power sensors. Figure 32 shows a photograph of the MEMS microwave frequency detector. When a microwave signal is input, the signal is evenly divided into two signals with the equal power and phase by the power divider. Next, one divided signal passes through the phase shifter more than the other. As a result, a phase difference is generated between the two separated signals. After that, the two signals with the phase difference are synthesized by the power combiner. Finally, the microwave power of the synthesized signal is measured by the indirect-heating and capacitive power sensors. It should be noted that the synthesized power relates to the phase difference. According to the relationship between the wavelength and frequency, the phase difference caused by the shifter is proportion to the frequency. At this time, the frequency measurement can be converted into the phase measurement. Therefore, a linear relationship between the synthetic power and the microwave frequency is obtained.



**Fig. 31** SEM view of dual-channel MEMS microwave phase detector based on MEMS power sensors

**Fig. 32** Photograph of MEMS microwave frequency detector based on MEMS power sensors



## Summary

This chapter presents an overview of MEMS microwave power sensors to characterize the power of microwave signals. The topics include various representations of classification, terminal indirect-heating and self-heating power sensors for all power measurements, inline coupling, capacitive and inserting power sensors for portion power measurements, and finally extension applications of the power sensors for the measurement of the phase and frequency of microwave signals. These MEMS

microwave sensors are passive structures and implemented based on the GaAs MMIC technology, with good performances. Through studying this chapter, we believe that we can gain a deeper understanding of the current development of MEMS microwave sensors.

---

## References

- Barker NS, Rebeiz GM (1998) Distributed MEMS true-time delay phase shifters and wide-band switches. *IEEE Trans Microw Theory Tech* 46(11):1881–1890
- Brush AS (2007) Measurement of microwave power. *IEEE Instrum Meas Mag* 10(2):20–25
- Cui Y, Liao X (2012) Modeling and design of a capacitive microwave power sensor for X-band applications based on GaAs technology. *J Micromech Microeng* 22(5):055013. (10pp)
- Daullé A, Xavier P, Raully D (2001) A power sensor for fast measurement of telecommunications signals using substitution method. *IEEE Trans Instrum Meas* 50(5):1190–1196
- Dehe A, Krozer V, Fricke K, Klingbeil H, Beilenhoff K, Hartnagel HL (1995) Integrated microwave power sensor. *Electron Lett* 31(25):2187–2188
- Dehé A, Krozer V, Chen B, Hartnagel HL (1996) High-sensitivity microwave power sensor for GaAs-MMIC implementation. *Electron Lett* 32(23):2149–2150
- Dehé A, Fricke-Neuderth K, Krozer V (2002) Broadband thermoelectric microwave power sensors using GaAs foundry process. *IEEE MTT-S international microwave symposium digest*. Institute of Electrical and Electronics Engineers Inc., Seattle, pp 1829–1832
- Fernández LJ, Visser E, Sesé J, Wiegierink R, Flokstra J, Jansen H, Elwenspoek M (2003) Radio frequency power sensor based on MEMS technology. *Proceedings of IEEE Sensors*. Institute of Electrical and Electronics Engineers Inc., Toronto, pp 549–552
- Fernández LJ, Sesé J, Wiegierink R, Flokstra J, Jansen H, Elwenspoek M (2005) Radio frequency power sensor based on MEMS technology with ultra low losses. *Proceedings of IEEE International Conference on Micro Electro Mechanical Systems*. Institute of Electrical and Electronics Engineers Inc., Miami Beach, pp 191–194
- Fernández LJ, Wiegierink RJ, Flokstra J, Sesé J, Jansen HV, Elwenspoek M (2006) A capacitive RF power sensor based on MEMS technology. *J Micromech Microeng* 16(7):1099–1107
- Han J, Liao X (2014) A 0.1–40 GHz broadband MEMS clamped-clamped beam capacitive power sensor based on GaAs technology. *J Micromech Microeng* 24(6):065024. (7pp)
- Han J, Liao X (2015) Third-order intermodulation of an MEMS clamped-clamped beam capacitive power sensor based on GaAs technology. *IEEE Sensors J* 15(7):3645–3646
- Han J, Liao X (2016a) Analytical and electrical modeling of a MEMS thermoelectric microwave power sensor. *J Micromech Microeng* 26(9):094001. (7pp)
- Han J, Liao X (2016b) A compact broadband microwave phase detector based on MEMS technology. *IEEE Sensors J* 16(10):3480–3481
- Han L, Huang Q-A, Liao X-P (2007) A microwave power sensor based on GaAs MMIC technology. *J Micromech Microeng* 17(10):2132–2137
- Han L, Huang Q-A, Liao X-P, Su S (2009) A micromachined inline-type wideband microwave power sensor based on GaAs MMIC technology. *J Microelectromech Syst* 18(3):705–714
- Hua D, Liao X, Zhang J, Jiao Y (2009) MEMS frequency detector at X-band based on MMIC technology. *TRANSDUCERS – International Conference on Solid-State Sensors, Actuators and Microsystems*. IEEE Computer Society, Denver, pp 1405–1408
- Hua D, Liao X-p, Jiao J-c (2011) X-band microwave phase detector manufactured using GaAs micromachining technologies. *J Micromech Microeng* 21(3):035019. (7pp)
- Hua D, Liao X-p, Liu H-c (2013) A micro compact coplanar power divider at X-band with finite-width ground plane based on GaAs MMIC technology. *Microsyst Technol* 19(12):1973–1980
- Jackson WH (1974) A thin-film/semiconductor thermocouple for microwave power measurement. *Hewlett-Packard J* 26(1):16–18

- V. Milanovic, M. Gaitan, M. E. Zaghoul: Micromachined thermocouple microwave detector in CMOS technology, Midwest Symp. Circuits Syst. (IEEE, Piscataway, NJ, United States, Ames 1996), pp. 273–276
- Milanovic V, Gaitan M, Bowen ED, Tea NH, Zaghoul ME (1997) Thermoelectric power sensor for microwave applications by commercial CMOS fabrication. *IEEE Electron Device Lett.* 18(9):450–452
- Milanovic V, Gaitan M, Zaghoul ME (1998) Micromachined thermocouple microwave detector by commercial CMOS fabrication. *IEEE Trans Microw Theory Tech* 46(5):550–553
- Muldavin JB, Rebeiz GM (2000) High-isolation CPW MEMS shunt switches-Part1: modeling. *IEEE Trans. Microw. Theory Tech.* 48(6):1045–1052
- Rebeiz GM (2003) *RF MEMS Theory, Design, and Technology*. Wiley, Hoboken
- Ulm M, Walter T, Mueller-Fiedler R, Voigtlaender K, Kasper E (2000) K-band capacitive MEMS-switches. *opical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*. Institute of Electrical and Electronics Engineers Inc., Garmisch, pp 119–122
- Wang D-b, Liao X-p (2009) A novel symmetrical microwave power sensor based on GaAs monolithic microwave integrated circuit technology. *J Micromech Microeng* 19(7):125012. (8pp)
- Wang D-b, Liao X-p (2010) A terminating-type MEMS microwave power sensor and its amplification system. *J Micromech Microeng* 20(7):075021. (8pp)
- Wang DB, Liao XP (2012a) Package solution of indirectly-heated type thermoelectric power sensors for RF application. *Electron Lett* 48(2):102–103
- Wang D-b, Liao X-p (2012b) A novel MEMS double-channel microwave power sensor based on GaAs MMIC technology. *Sens. Actuators A Phys.* 188:95–102
- Wang D-b, Liao X-p (2012c) A 35 GHz wireless millimeter-wave power sensor based on GaAs micromachining technology. *J Micromech Microeng* 22(6):065025. (8pp)
- Wang DB, Liao XP, Liu T (2012a) Optimization of indirectly-heated type microwave power sensors based on GaAs micromachining. *IEEE Sensors J* 12(5):1349–1355
- Wang D-B, Liao X-P, Liu T (2012b) A thermoelectric power sensor and its package based on MEMS technology. *J Microelectromech Syst* 21(1):121–131
- Wang DB, Liao XP, Liu T (2012c) A novel thermoelectric and capacitive power sensor with improved dynamic range based on GaAs MMIC technology. *IEEE Electron Device Lett* 33(2):269–271
- Yan H, Liao X (2015) The high power up to 1 W characteristics of the capacitive microwave power sensor with grounded MEMS beam. *IEEE Sensors J* 15(12):6765–6766
- Yan J, Liao X (2016a) Equivalent lumped circuit model and S-parameter of indirect-heating thermoelectric power sensor. *Sens. Actuators A Phys.* 240:110–117
- Yan J, Liao X (2016b) Research on the response time of indirect-heating microwave power sensor. *IEEE Sensors J* 16(13):5270–5276
- Yan H, Liao X, Hua D (2016) An X-band dual channel microwave phase detector based on GaAs MMIC technology. *IEEE Sensors J* 16(17):6515–6516
- Yi Z, Liao X (2012) An 8-12 GHz microwave frequency detector based on MEMS power sensors. *J Micromech Microeng* 22(3):035005. (8pp)
- Yi Z, Liao X (2013) A capacitive power sensor based on the MEMS cantilever beam fabricated by GaAs MMIC technology. *J Micromech Microeng* 23(3):035001. (10pp)
- Yi Z, Liao X (2014) Measurements on intermodulation distortion of capacitive power sensor based on MEMS cantilever beam. *IEEE Sensors J* 14(3):621–622
- Yi Z, Liao X (2016a) A 3D model of the thermoelectric microwave power sensor by MEMS technology. *Sensors* 16(6):921
- Yi Z, Liao X (2016b) A cascaded terminating-type and capacitive-type power sensor for –10-to 22-dBm application. *IEEE Electron Device Lett* 37(4):489–491
- Yi Z, Liao X, Wu H (2013) Modeling of the terminating-type power sensors fabricated by GaAs MMIC process. *J Micromech Microeng* 23(8):085003. (9pp)
- Yi Z, Yan H, Yan J, Liao X (2016) Fabrication of the differential microwave power sensor by seesaw-type MEMS membrane. *J Microelectromech Syst* 25(4):582–584

- Yue CP, Wong SS (2000) Physical modeling of spiral inductors on silicon. *IEEE Trans Electron Devices* 47(3):560–568
- Zhang Z, Liao X (2011) Packaging-test-fixture for in-line coupling RF MEMS power sensors. *J Microelectromech Syst* 20(3):1231–1233
- Zhang Z, Liao X (2012a) A three-channel thermoelectric RF MEMS power sensor for GaAs MMIC applications. *Sens Actuators A Phys* 182:68–71
- Zhang Z, Liao X (2012b) A thermocouple-based self-heating RF power sensor with GaAs MMIC-compatible micromachining technology. *IEEE Electron Device Lett* 33(4):606–608
- Zhang Z, Liao X (2012c) GaAs MMIC fabrication for the RF MEMS power sensor with both detection and non-detection states. *Sens Actuators A Phys* 188:29–34
- Zhang Z, Liao X (2013a) A lumped model with phase analysis for inline RF MEMS power sensor applications. *Sens Actuators A Phys* 194:204–211
- Zhang Z, Liao X (2013b) Characterization of packaged inline-type radio frequency micro-electromechanical systems power sensors. *Sens Actuators A Phys* 201:294–301
- Zhang Z, Liao X (2014) Inline capacitive RF power sensor based on floating MEMS beam for GaAs MMIC applications. *Electron Lett* 50(18):1292–1294
- Zhang Z, Liao X (2015a) Suspended thermopile for microwave power sensors based on bulk MEMS and GaAs MMIC technology. *IEEE Sensors J* 15(4):2019–2020
- Zhang Z, Liao X (2015b) An inline RF power sensor based on fixed capacitive coupling for GaAs MMIC applications. *IEEE Sensors J* 15(2):665–666
- Zhang Z, Liao X (2015c) A directional inline-type millimeter-wave MEMS power sensor for GaAs MMIC applications. *J Microelectromech Syst* 24(2):253–255
- Zhang Z, Liao X (2015d) An insertion thermoelectric RF MEMS power sensor for GaAs MMIC-compatible applications. *IEEE Microwave Compon Lett* 25(4):265–267
- Zhang Z, Liao X, Han L (2010) A coupling RF MEMS power sensor based on GaAs MMIC technology. *Sens Actuators A Phys* 160(1–2):42–47
- Zhang Z, Liao X, Han L, Cheng Y (2011) A GaAs MMIC-based coupling RF MEMS power sensor with both detection and non-detection states. *Sens Actuators A Phys* 168(1):30–38
- Zhang Z, Liao X, Wang X (2015) Research on thermocouple distribution for microwave power sensors based on GaAs MMIC process. *IEEE Sensors J* 15(8):4178–4179
- Zhang Z, Guo Y, Li F, Gong Y, Liao X (2016) A sandwich-type thermoelectric microwave power sensor for GaAs MMIC-compatible applications. *IEEE Electron Device Lett* 37(12):1639–1641
- Zheng W-B, Huang Q-A, Liao X-P, Li F-X (2005) RF MEMS membrane switches on GaAs substrates for X-band applications. *J Microelectromech Syst* 14(3):464–471





# A Micromachined Reconfigurable Attenuator

Zewen Liu and Xin Guo

## Contents

Overview .....	1120
Structure Design .....	1124
Reconfigurable Attenuator System Topology Design .....	1124
Resistive Network Module Design .....	1125
Microwave Compensate Structure Design .....	1130
RF MEMS Switch Design .....	1131
Simulation of the 3-Bit Reconfigurable Attenuator .....	1133
Fabrication .....	1134
Characterization .....	1138
Structural Observation .....	1138
DC Electrical Characterization .....	1139
RF Measurement .....	1140
Reliability Test .....	1145
Conclusion .....	1145
References .....	1145

## Abstract

As the essential power control and adjustment component, microwave power attenuator has been widely applied in spectrum analyzer, network analyzer, receivers, and other microwave instrument systems. There has been significant interest in developing miniaturized microwave devices with small geometric dimension for commercial microwave test systems. Traditional attenuator with large size is difficult to integrate with IC due to its cumbersome mechanical relay or switches. Radio frequency microelectromechanical system (RF MEMS) switches present many advantages such as less insertion loss, higher isolation, better linearity, and lower power consumption and more importantly, its small

---

Z. Liu (✉) · X. Guo  
Tsinghua University, Beijing, China  
e-mail: [liuzw@tsinghua.edu.cn](mailto:liuzw@tsinghua.edu.cn); [guoxinrainbow@gmail.com](mailto:guoxinrainbow@gmail.com)

size. It has absorbed great attention for the purpose of miniaturization of attenuator. Recent research shows step attenuator based on RF MEMS switches meet well the state-of-the-art requirements of miniaturized reconfigurable attenuation devices with high precision and broadband performance. In this chapter, we introduce a compact 3-bit step attenuator based on RF MEMS switches with 0 ~ 70 dB attenuation at 10 dB intervals up to 20GHz. The on-chip attenuator consists of 12 ohmic MEMS switches, 3  $\pi$ -type resistive attenuation networks, and microwave compensate structures. To optimize the attenuation characteristics within the broadband, theoretical analysis and 3D modelling were performed. The device was obtained using MEMS process combined with polysilicon integrated circuit (IC) process.

In section “[Structure Design](#),” the structure of the step attenuator was proposed and three modules were designed independently including: (1) Resistive attenuation network. Two approaches are presented and discussed. (a) Polysilicon thin film resistor with symmetric topology is applied to realize the high-precision  $\pi$ -type resistive attenuation modules (10, 20, 40 dB). (b) Distributed TaN single thin film resistor (STFR) equivalent to  $\pi$ -type resistive network was designed with a more compact structure (2/3 compared to traditional network). It can realize 5 ~ 30 dB attenuation with less parasite effects. (2) RF MEMS switches. Toggling part is the determining factor to reduce the footprint of the entire attenuator. SP2T switch including 2 ohmic Au contact single-cantilever MEMS switches was customized for the attenuator. (3) CPW and microwave compensate structures. Precise design of the transmission path connecting each module is significant to suppress the inherent insertion loss and improve the matching performance. Right-angle bends and T-junctions were analyzed and designed. Based on the results of these three blocks, the overall attenuator was assembled and optimized with consideration of matching performance.

In section “[Fabrication](#),” we describe the fabrication method of RF MEMS attenuators. The 3-bit attenuator was manufactured with surface micro-machining process and polysilicon IC techniques on 4 inch wafers and features 11 lithography steps. Key processes are emphasized including high-precision resistance and excellent ohmic contact between polysilicon and Au CPW was realized with experimental research. The photosensitive PI was used as sacrificial layer and the patterning process was simplified. The CPW and upper electrodes of the switch which were manufactured by low stress Au electroplating process have small roughness and without evident buckling after the sacrificial layer released. Based on these process studies, the 3-bit attenuator is successfully fabricated.

In section “[Characterization](#),” characterization of the obtained RF MEMS attenuator is presented. The measurement results show that the driven voltage of RF MEMS switch is 32 ~ 42 V and it can operate more than  $3 \times 10^8$  times (cold switched mode). When toggling between the different transmission paths using MEMS switches, the 3-bit attenuator can realize target attenuation in range of

0 ~ 70 dB at 10 dB intervals up to 20 GHz. The accuracy and error of all the attenuation states is better than  $\pm 1.88$  dB and 2.11 dB. The return loss of the 3-bit switched attenuator is better than 11.95 dB. Micromachined reconfigurable attenuator based on RF MEMS switch is demonstrated and can be further applied in particular microwave systems.

---

### Keywords

3-Bit reconfigurable attenuator · DC electrical characterization · Micromachined reconfigurable attenuator · Radio frequency (RF) measurement · Radio frequency micro-electromechanical system (RF MEMS) switches · Single thin film resistor (STFR) · Single-pole double-throw (SPDT) · TaN single thin film resistor (TaN STFR) attenuation module

---

### List of Abbreviations

Al	Aluminum
Au	Gold
CPW	Co-planar waveguide
Cu	Copper
FEA	Finite element analysis
HRS	High-resistivity silicon
IC	Integrated circuit
ICP	Inductively coupled plasma
LPCVD	Low pressure chemical vapor deposition
MEMS	Micro-electro-mechanical system
MESFET	Metal–semiconductor field-effect transistor
MMIC	Monolithic microwave integrated circuit
N <sub>2</sub>	Nitrogen
O <sub>2</sub>	Oxygen
PECVD	Plasma enhanced chemical vapor deposition
pHEMT	Pseudomorphic high electron mobility transistor
PI	Polyimide
PIN	Positive intrinsic-negative
PolySi	Polysilicon
RF	Radio frequency
SEM	Scanning electron microscope
Si <sub>3</sub> N <sub>4</sub>	Silicon nitride
SiO <sub>2</sub>	Silicon dioxide
SPDT	Single pole double throw
STFR	Single thin film resistor
TRL	Through-reflect-line
VCA	Voltage controlled attenuator
VSWR	Voltage standing wave ratio

## Overview

Microwave attenuator reduces the input signal level by absorbing its power with linear network, and it has two main types: active and passive, depending on the function principles of its components. Operating frequency band, attenuation value, accuracy, return loss (or VSWR), and power capacity are the key parameters to characterize a microwave attenuator.

As an essential device for controlling and adjusting input microwave power, programmable attenuator is used in a variety of equipment, including frequency spectrometers, network analyzers, noise factor analyzers, and radio frequency communication front-end modules (Hakan et al. 2008; Yong-Sheng et al. 2007; Hyunchul et al. 2008). As the core block, multibit digital reconfigurable attenuators with switches can also be applied in automatic radar testing system, playing a role for imitation of receiving signals with different levels.

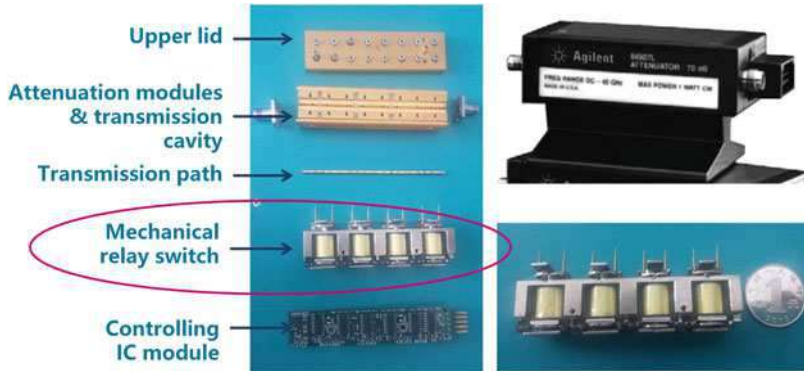
The traditional reconfigurable attenuators are based on mechanical switches, as shown in Fig. 1 as an example. The commercial product 84907 L, programmable attenuator produced by Keysight can operate within DC ~ 40GHz and it offers a repeatability of better than 0.03 dB, and an attenuation range of 0–70 dB in 10 dB steps (<http://www.keysight.com/en/pd-1000001949%3Aeps%3Apro-pn-84907L/programmable-attenuator-40-ghz-70db-10db-steps?nid=-32726.536879322&cc=GB&lc=eng>). Similar products are also provided by Willtron, Narda, Weinschel, etc. The life time of this kind of mechanical switched attenuator is usually limited around five million switching cycles.

It is noted that the size and weight of these devices are in order of 10 cm and 100 g, which set limits for the traditional attenuator to be integrated with other IC controlling modules in programmable systems (Zheng et al. 2016). There has been significant interest in developing miniaturized microwave devices with small geometric dimension for commercial microwave test systems, such as phase shifters, delay lines, and filters (Sanghyo et al. 2004; Kim et al. 2001a; Sekar et al. 2011; Lee et al. 2004; Sukomal et al. 2015; Diego Correias et al. 2014). These devices are also promising to be used in future 5G communication systems (Osseiran et al. 2014).

The disassemble graph of a traditional mechanical attenuator is shown in Fig. 2. It reveals that the obstacle factor for the attenuator to scale down is the mechanical relay switch. Miniaturized reconfigurable attenuators have been intensively investigated



Fig. 1 Reconfigurable attenuator products based on mechanical relay switches



**Fig. 2** Main components inside the traditional attenuator based on mechanical relay switches

using various switches, including PIN, FET, pHEMT MMIC, and more recently the Radio Frequency Micro-Electro-Mechanical System (RF MEMS) switches (Stukach 2006; Marcus et al. 2011; Bae et al. 2017; Avago).

The integrated attenuator with traditional semiconductor switch often works on relatively low frequency domains, such as the attenuator based on PIN diode switch and T-type attenuation network developed by Stukach (2006). The device operates within a narrow band of DC ~ 3GHz with an attenuation of 0.2 ~ 15 dB. The attenuator has an average of 0.3 deg./dB phase shift up to 2 GHz.

In 2010, Marcus Granger-Jones et al. (2011) proposed a high-linearity voltage controlled attenuator (VCA) based on SOI CMOS technique. The VCA attenuators design is based on the classical passive FET  $\pi$  and T resistor network structures, and the stacked FET techniques was used to improve the signal handling capability. The device with FET switches has attenuation of 34 dB@ DC ~ 5GHz with IIP3 over 47 dBm. From the results given by the authors, it is found that the attenuation performance of the device is deteriorated by its high inherent insertion loss ( $-2.4$  dB@1 GHz).

Reconfigurable attenuator based on active switch also draws attention from industry. For instance, AMMC-6650 based on pHEMT MMIC switch provided by Avago. It can realize attenuation up to 20 dB within DC ~ 40GHz. Although it can operate in a wider frequency band, the attenuation accuracy and flatness have to be improved for practical applications. Furthermore, attenuator with pHEMT switch is still to be developed if higher attenuation is desired.

The current studies on reconfigurable attenuator presented in literatures reveal that the attenuators based on traditional active switches still have many problems for high-performance applications that require a large dynamic range and high attenuation precision, due to their inherent high power consumption and low linearity, particularly in high-frequency range.

As one of the fundamental components in RF/microwave system, RF MEMS switch has been intensively researched worldwide during the last two decades in USA, Europe, and China (Petersen 1979, 1982; Zavracky et al. 1997; Rebeiz et al.

2001, 2003, 2012, 2015; Lucyszyn 2010; Zhihao et al. 2009; Ling et al. 2013). Using RuO<sub>2</sub>-Au contacts, the reliability of MEMS switches has improved dramatically to a magnitude of 10 billion times (David et al. 2013). Nowadays, RF MEMS switches have been regarded as one of the most promising devices to be applied for future RF/microwave communication systems thanks to its high linearity, high isolation, low insertion loss, low power consumption, and high reliability (Sanghyo et al. 2004; Marcus et al. 2011; Vikram et al. 2011). Specially, its excellent linearity has drawn attention in 5G communication systems in both mobile front-ends and base-stations. Thanks to its outstanding microwave performance, RF MEMS switch are also regarded as an important building block for miniaturized reconfigurable attenuators for wide-band applications in RF test instruments and wireless communication system.

The first demonstration of this application was by Iannacci et al. (2009). A 6-bit step attenuator based on an RF MEMS switch and polysilicon resistor network was proposed and fabricated, with attenuation of 1.8 ~ 19.7 dB @ 1 ~ 13.5 GHz. This work presented a good example of integration of RF MEMS switches with polysilicon resistor networks. Thereafter, different reconfigurable attenuators working on different frequency bands have been steadily developed in FBK (Italy).

To further investigate RF-MEMS as an enabling solution to address the increasing demands for the upcoming 5G communication system, the same group had developed several novel multibit reconfigurable attenuators operating within a wide range from low frequency to 110 GHz (Iannacci et al. 2016a, b, c, 2017). In their latest published work (Iannacci et al. 2017), electrostatically actuated MEMS switches and resistive resistors both in series and shunt configuration are used for an 8-bit attenuator. The equivalent lumped resistive network illustrated can implement 256 network configurations by controlling the MEMS switches. From 10 MHz to 110 GHz, the device can realize attenuation of 10 ~ 60 dB and shows flatness from 3 to 15 dB. These works are significant demonstration of the potential of RF-MEMS technology in realization of high performance reconfigurable passive components. Nevertheless, it is found that when a large attenuation values (S<sub>21</sub>) required, instead of power consumption on the network, the RF signal is directly reflected at the input port due to large VSWR. The mismatching problem may be a challenge at large attenuation and further optimization is required in future device design.

Sun Junfeng et al. (2016) from Southeast University designed a 3-bit attenuator based on RF MEMS switches in 2016. The MEMS switch has three contact dimples to control the transmission direction of microwave signal. The measured attenuation range is 0 ~ 35 dB @DC ~ 20GHz. It presents a good insertion loss less than 1.7 dB and the attenuation error is less than ±5%.

In summary, miniaturized reconfigurable attenuator is an important device used in both microwave measurement instrument and wireless communication front end. The technology progress of RF MEMS switches allows us to realize integrated reconfigurable attenuators with high performance and surely low cost. A comparison of attenuators based on different switching techniques is listed in Table 1. It can be found that the main features of RF MEMS-based reconfigurable attenuator are:

**Table 1** Comparison between reconfigurable attenuators based on different switches

	GaAs FET	Mechanical relay	RF MEMS
Life	High	Low	High
Integration	√	×	√
Operating band	DC ~ 6GHz	DC ~ 50GHz	DC ~ 110GHz
Attenuation range	0 ~ 130 dB, step 5 dB	0 ~ 110 dB	0 ~ 70 dB
Accuracy	~ 2 dB	< ±1 dB	~0.5 dB
Switching time	< 20 μs	> 10 ms	100 μs
Reference	USA Agilent Inc. Microwave technology Center	USA WSTF (White Sands Test Facility) Calibration Laboratory	Italy FBK Research Center China Tsinghua University Southeast University

1. Lower insertion loss, higher accuracy and higher reliability. The actuation of RF MEMS switches can reach over several billion times (cold switched mode).
2. Wide frequency operating band and possibility of design optimization. RF MEMS devices show excellent performance at a wide frequency up to more than 100 GHz. Furthermore, the structure of RF MEMS switch is similar with other RF MEMS tuning devices (such as capacitor, etc.), and it can be utilized in optimization of resistive matching to improve the attenuation accuracy of the system.
3. Fabrication compatibility with IC process. Traditional attenuator using mechanical switches is hard to be integrated with controlling circuits. RF MEMS switches fabrication process is possibly compatible with IC process, and the reconfigurable attenuator with its driving circuits can be fabricated or integrated in one single chip.

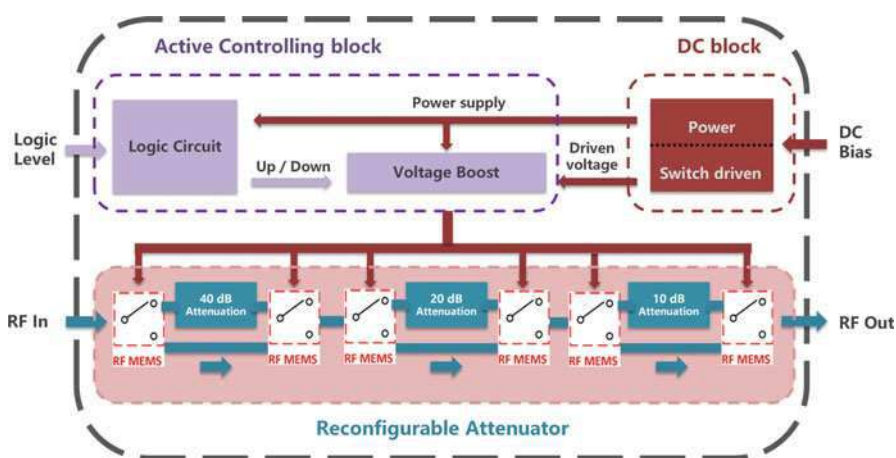
To provide a single-chip broadband attenuator with wide attenuation range and better performance, a miniaturized 3-bit reconfigurable attenuator based on compact RF MEMS single-pole double-throw (SPDT) switches and a resistive attenuation module is presented in this chapter. In the rest part of this chapter, we describe structure design, fabrication, and measurement of the reconfigurable attenuator. The switches were utilized to toggle between each attenuation module path and a reference path to realize the target signal attenuation. To obtain broadband flatness and high precision, the resistive attenuation module and compensate structures are optimized using theoretical calculation and simulations. Symmetric topology is used in the overall device structure design. The total dimension of the 3-bit device is  $2.45 \times 4.34 \times 0.5 \text{ mm}^3$ . Measurement results show that the fabricated device has an excellent attenuation performance between 10 ~ 70 dB @ DC and 20 GHz with an accuracy better than  $\pm 1.88 \text{ dB}$  at 60 dB and error less than 2.22 dB at 10 dB. The device return loss is better than 11.95 dB for all eight testing states (Xin et al. 2016).

## Structure Design

Normally, a programmable attenuator consists of two main blocks, controlling block and microwave reconfigurable attenuator, as shown in Fig. 3. The red lines and blue lines indicate the DC and RF signal paths, respectively. In the controlling block, attenuation value is determined by the input signal (“0”/ “1”) of the Logic Circuit, the driven voltage (20 ~ 80 V) of the RF MEMS switching voltage is provided by a voltage boost chip. The reconfigurable attenuator block reacts to the controlling block and required attenuation is realized on the RF signal path. In this section, we present the overall microwave reconfigurable attenuator structure design and simulations based on careful analysis of individual RF MEMS switch, compensate structure, and resistive attenuation modules.

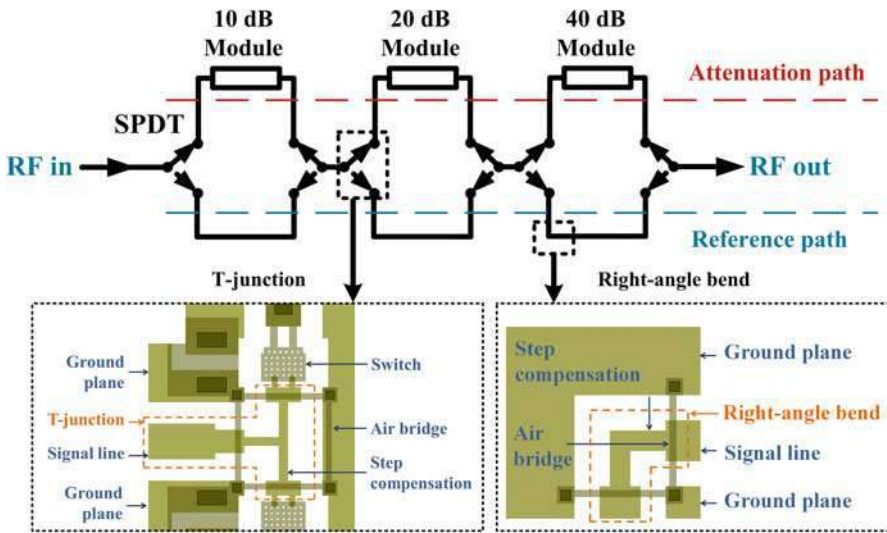
## Reconfigurable Attenuator System Topology Design

The 3-bit reconfigurable RF MEMS attenuator consists of three cascaded 1-bit attenuation units. Each 1-bit unit is designated with 10-, 20-, and 40-dB attenuation modules, as shown in Fig. 4 (Zhihao et al. 2009). With consideration of high isolation and matching performance, each 1-bit unit comprises one electrostatically activated SPDT MEMS switches and two transmission paths (the reference path and the attenuation path) based on Au CPW. The reference path is a transmission line with minimum power consumption. The attenuation path absorbs signal power at 10-, 20-, and 40-dB, respectively according to the switch reconfiguration. Resistive networks are applied to implement attenuation at each module. T-junction, right-angle bend, and other compensational structures are designed to address the mismatching problems; therefore, the microwave performance can be optimized. By



**Fig. 3** The schematic view of a programmable attenuator system





**Fig. 4** Schematic overview of the proposed 3-bit reconfigurable attenuator based on RF MEMS SPDT switches. Compensation structures are designed to reduce the inherent transmission loss

routing the signal through the selected transmission path using SPDT switches, attenuation of 0 ~ 70 dB with an interval of 10 dB (eight states) is realized.

## Resistive Network Module Design

The attenuation module is the essential component for the performance of the 3-bit attenuator. Resistive attenuation network based on microfabrication techniques has been widely applied thanks to its smaller size, possible longer operating life, and capability to be integrated in on-chip system (Simon et al. 2009). “T,” “ $\pi$ ,” and mix “T” are the most frequently used topologies of the resistive network.

To achieve high accuracy for on-chip systems in microwave application scenarios, several substrate and resistor materials have been studied such as TaN thin film on BeO, quartz (Jiang et al. 2010; Cano et al. 2010; Yeh et al. 2016), as well as graphene flakes (Yasir et al. 2017). As a widely utilized material in both CMOS and MEMS devices, polysilicon thin film resistor has also been intensively investigated in RF/microwave applications for its low cost, tunable resistivity, and little parasitic capacitance (Hiromi et al. 2004), such as load resistors in static RAM (Ohzone et al. 1985), structural layer in pressure sensor (Malhaire et al. 2003), and high-Q resonator (Jing et al. 2004). Iannacci et al. (2009) used polysilicon resistor to realize on-chip attenuator with attenuation of 1.8 ~ 19.7 dB from 1 ~ 13.5 GHz. It was probably the first trial to use polysilicon as attenuation resistor, but the matching performance of the attenuator is far from requirements, in which only one single resistor was applied in the attenuation path. Further improvement should be carried out with

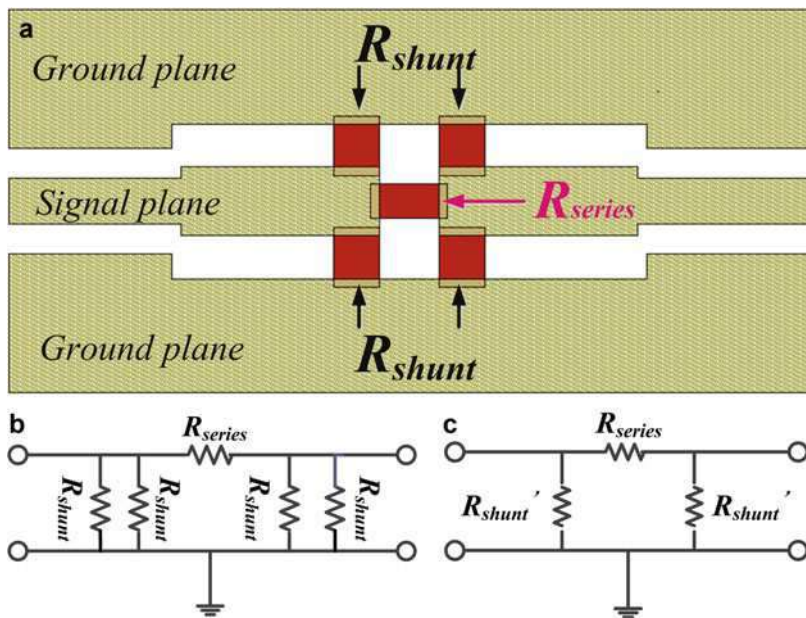
regards of attenuator resistive network design, microwave signal matching and the high-precision resistivity process control.

In this section, we introduce the resistive network design. Traditional “T” type or “ $\pi$ ” type occupies a large amount of more device space. To improve the device size and reduce the parasitic effect, distributed single thin film resistor (STFR) based on TaN is also discussed.

### Polysilicon Resistive Network

To reduce parasite effects, a  $\pi$ -type polysilicon network with symmetric topology was designed with high-precision attenuation properties. The attenuation module comprised both series and shunt resistors: a series resistor  $R_{series}$  is in the center of the signal line and four shunt resistors  $R_{shunt}$  are symmetrically allocated to ground, as illustrated in Fig. 5a. The equivalent circuit of the structure is drawn in Fig. 5b. By replacing the two equal parallel  $R_{shunt}$  with one  $R_{shunt}'$  ( $R_{shunt}' = 2R_{shunt}$ ), a so-called  $\pi$ -type topology was formed, as shown in Fig. 5c.

The value of  $R_{shunt}$  and  $R_{series}$  can be calculated according to the attenuation amplitude  $A$ , as described by Eqs. 1 and 2 (Simon et al. 2009):



**Fig. 5** (a) Design layout of the resistive attenuation network. (b) Equivalent circuit of the structure. (c) Simplification of the equivalent circuit

**Table 2** Design parameters and dimensions of the attenuation modules

		$R_{series}$	$R_{shunt}$	CPW (center)	CPW (port) <sup>a</sup>
10 dB	Design value ( $\Omega$ )	68	190	50 $\Omega$ at 20GHz	50 $\Omega$ at 20GHz
	L/W ( $\mu\text{m}/\mu\text{m}$ )	35/60	75/50		
20 dB	Design value ( $\Omega$ )	228	120	–	–
	L/W ( $\mu\text{m}/\mu\text{m}$ )	105/60	75/80	75/120/200	52/80/243
40 dB	Design value ( $\Omega$ )	213	18	–	–

<sup>a</sup>Matched to microprobe for detection purpose

$$R_{shunt} = \frac{(10^{\frac{A}{20}} - 1) \cdot Z_0}{10^{\frac{A}{20}} + 1 - 2 \cdot 10^{\frac{A}{20}}} \quad (1)$$

$$R_{series} = \frac{Z_0}{2} \cdot \frac{10^{\frac{A}{20}} - 1}{10^{\frac{A}{20}}} \quad (2)$$

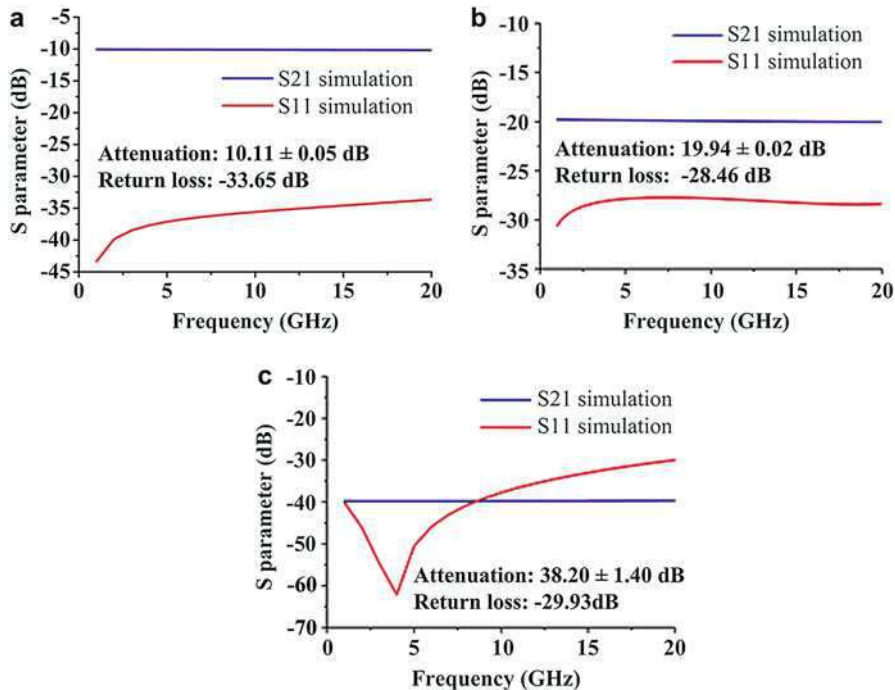
where  $Z_0$  is the characteristic impedance of the port (50  $\Omega$  at 20 GHz). The calculated  $R_{shunt}$  and  $R_{series}$  for the 10-, 20-, and 40- dB modules are listed in Table 2. As shown in the table, the values of  $R_{shunt}$  and  $R_{series}$  in the 10 and 20-dB modules are similar and they were realized and calibrated with same sheet resistance and a compact layout. However, there is a large geometric disparity (almost 10 times) between the  $R_{shunt}$  and  $R_{series}$  values for the 40-dB module, which can lead to serious RF performance deterioration due to the parasitic effect results from the significant size variation. It is found the 40-dB module performance can be largely improved when using two cascaded 20 dB module structure instead of one single 40 dB attenuator.

The sheet resistance of all the resistors in attenuator modules was optimized with the commercial Finite Element Analysis (FEA) electro-magnetic simulator HFSS. To obtain target length/width ratio of 0.5 ~ 2 for all the resistors in the network, the optimized sheet resistance was determined to be 127  $\Omega/\square$  and the corresponding structure parameters are summarized in Table 2. The CPW transmission structure dimensional parameters used in the simulations are also shown in the same table.

The simulation results demonstrate that the three modules present desired attenuation performance at the 10, 20, and 40 dB, as shown in Fig. 6 and Table 3. The error was less than 4.94% for all three modules. The matching performance was better than  $-28.46$  dB over the entire operating band (up to 20 GHz).

### TaN STFR (Single Thin Film Resistor) Attenuation Module

The distributed resistive attenuator embedded in CPW with single TaN thin film resistor and equivalent  $\pi$ -type resistive network (Zhong and Liu 2016) is shown in Fig. 7. Instead of three resistors in Fig. 13a, one single film resistor is used as shown in Fig. 13b. Equivalently, the  $R_P$  and  $R_S$  in Fig. 7a are divided into two parallel  $2R_P$  and  $2R_S$  in Fig. 7b, respectively.  $l$  and  $h$  are the length and width of the STFR, respectively.  $h$  has been set to 280  $\mu\text{m}$  in the CPW design.



**Fig. 6** 3D FEA EM simulation results of polysilicon resistive network (a) 10 dB module; (b) 20 dB module; (c) 40 dB module. *S21*: Attenuation value. *S11*: Return loss

**Table 3** Simulation results of the three attenuation modules

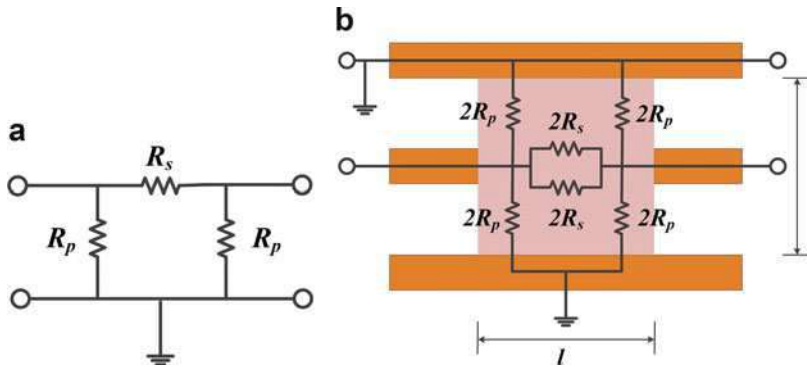
	Attenuation S21 (dB)	Return loss S11 (dB)	Error (%)
10 dB	$10.11 \pm 0.05$	33.65	4.94
20 dB	$19.94 \pm 0.02$	28.46	1.00
40 dB	$38.20 \pm 1.40$	29.63	3.66

To determine the STFR dimensions with given resistivity  $\rho$  according to the requirement of attenuation, one can divide the STFR into two same parts, as shown in Fig. 8a.

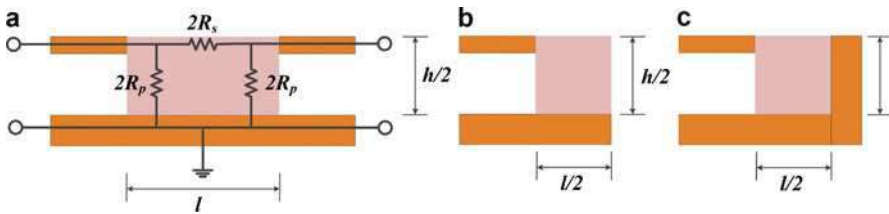
Supposing  $R_{sc}$  and  $R_{oc}$  are the short circuited and open circuited resistance of STFR, respectively, the relationship between  $R_p$ ,  $R_s$ , and  $R_{sc}$  and  $R_{oc}$  can be written as following equations:

$$R_s = \frac{R_{sc}\sqrt{R_{oc}}}{2\sqrt{R_{oc} - R_{sc}}} \tag{3}$$

$$R_p = \frac{R_{oc} + \sqrt{R_{oc}^2 - R_{oc}R_{sc}}}{2} \tag{4}$$



**Fig. 7** Equivalent circuit of STFR. (a)  $\pi$ -type network. (b) equivalent circuit



**Fig. 8** Equivalent circuit of half of STFR. (a) half STFR. (b)  $R_{och}$  (c)  $R_{sch}$

**Table 4** Length and sheet resistance of STFR

Attenuation (dB)	10	15	20	30
Length ( $\mu\text{m}$ )	90.96	140.21	189.65	288.58
Sheet resistance ( $\Omega/\square$ )	91.64	92.08	92.24	92.30

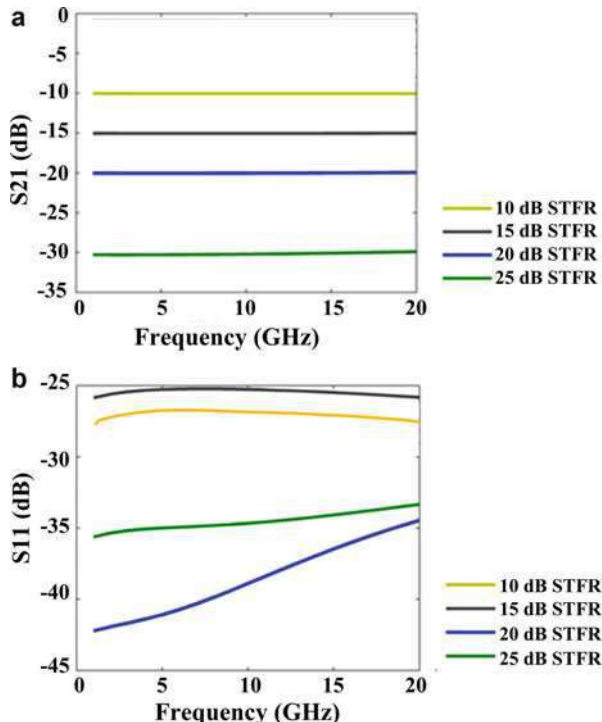
To get  $R_{oc}$  and  $R_{sc}$ , the half of STFR is further divided into two same parts defined as open circuit half resistor (OCHR) with a resistance of  $R_{och}$  as illustrated in Fig. 8b. To get an accurate resistance, another parameter defined as short circuit half resistor (SCHR) is necessary, which has an extended electrode as illustrated in Fig. 8c. According to Bartlett’s theory (Mingchun 2010),  $R_{sc}$  and  $R_{oc}$  can be calculated by equation:

$$R_{sc} = \frac{2R_{och}R_{sch}}{R_{och} + R_{sch}} \tag{5}$$

$$R_{oc} = \frac{R_{och} + R_{sch}}{2} \tag{6}$$

The value of  $R_{och}$  and  $R_{sch}$  can be obtained with conformal mapping approach developed in (Van Caekenbergh 2009). The length and sheet resistance of the STFR for different attenuation (10, 15, 20, and 30 dB) are listed in Table 4.

**Fig. 9** Simulations results of the TaN STFRs. (a) Attenuation value S21. (b) Return loss S11

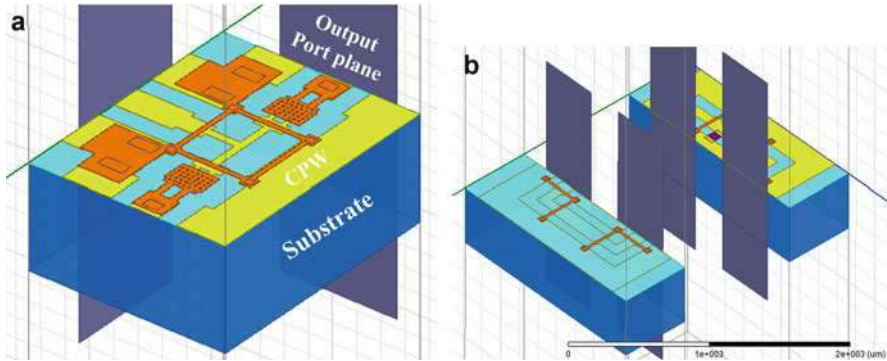


Compared to the traditional  $\pi$ -type resistive network, the parasitic inductance introduced by the resistors located between the signal plane and ground (four  $R_{shunt}$  in Fig. 5b) is significantly reduced. The shorter length of STFR also leads to less parasitic inductance. The simulation results of the TaN resistive module are shown in Fig. 9. The attenuation of the 4 STFR is  $10.02 \pm 0.05$  dB,  $15.29 \pm 0.02$  dB,  $20.05 \pm 0.14$  dB, and  $29.95 \pm 0.35$  dB, respectively, and the return loss is better than 25 dB. The distributed STFR shows excellent RF performance as well as smaller footprint.

## Microwave Compensate Structure Design

As a power-absorbing device, a broadband microwave attenuator is sensitive to the inherent signal loss introduced by parasitic issues, especially at high frequencies. As illustrated in Fig. 4, the CPW configuration of the attenuator comprises of multiple discontinuities such as right-angle bends and T-junctions, which require compensation structure design to suppress parasitic effects (Weller et al. 1998). As shown in Fig. 4, compensate structures for right-angle bends and T-junctions are designed carefully to improve the matching performance and accuracy of attenuation.

Using 500  $\mu\text{m}$  low-loss high-resistivity silicon (HRS) as substrate, the fundamental Au CPW structure is designed for 50  $\Omega$  matching at 20 GHz. To mitigate



**Fig. 10** Modelling of T-junction and right-angle bend. (a) T-junction (b) right-angle bend

the coupled slot-line mode signal in laterally asymmetric structures, air-bridges are placed near CPW discontinuities in both right-angle bends and T-junctions (Weston et al. 1997). However, unwanted capacitance was introduced by air-bridges consequently. It can be further improved using step compensation techniques (Weller et al. 1998).

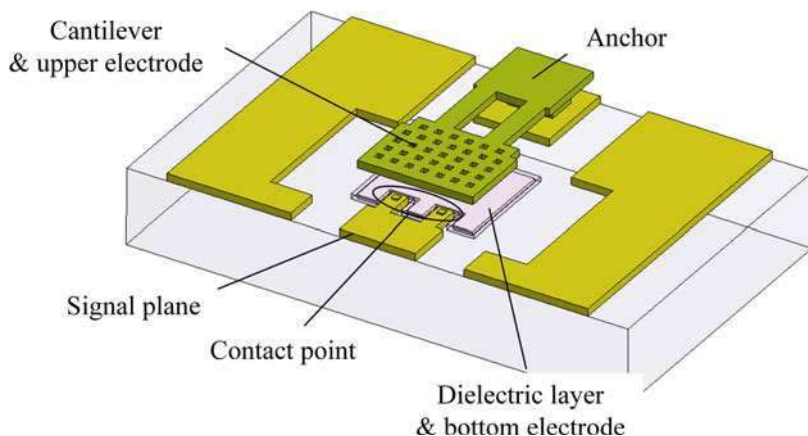
Furthermore, T-junction is a signal-divider structure. When the RF signal transmits through one path, the other path will be left as an open-stub. The capacitive open-stub can also lead to RF performance degradation (Kim et al. 2001b). A narrower conductive signal line at the input port of the T-junction is applied to reduce this effect.

The structure was designed and simulated using HFSS, as shown in Fig. 10. In Fig. 10, The light blue part indicates the high-resistivity silicon substrate, the yellow part is the microwave transmission line, the orange part is the upper cantilever of the MEMS switch and the air bridge, and the dark blue part is the simulation 2-port plane. The simulation results show that the utility of the T-junction and right-angle bends greatly improve the matching performance with a return loss of  $-28$  dB (@ 20 GHz) and insertion loss less than 0.18 dB.

## RF MEMS Switch Design

The RF MEMS switch is the critical component for the miniaturization of reconfigurable attenuator. In the last decade, Tsinghua University focuses on investigation of both parallel capacitive and series resistive switches with various structure configurations (David et al. 2016; Bansal et al. 2017). The Au–Au contact series resistive RF MEMS switch with a single-cantilever structure is applied in to the 3-bit attenuator, as shown in Fig. 11. The details of the switch were described in earlier publication (Ling et al. 2013). When a DC-driven voltage is added between the upper cantilever and the bottom electrode, the cantilever is electrostatically actuated





**Fig. 11** Modelling of the RF MEMS switch

**Table 5** Simulation summary of each module

		Insertion loss (dB)		Return loss (dB)	Isolation (dB)
CPW		-1.03		< -35.20	-
RF MEMS switch		-0.07		< -23.08	15.40
SP2T switches		-0.15		< -28.25	16.74
Right-angle bend and T junction		-0.19		< -20.67	-
Attenuation network	10 dB	-10.12	$\pm 0.05$	< -33.65	-
	20 dB	-19.94	$\pm 0.03$	< -28.46	-
	40 dB	-39.72	$\pm 0.08$	< -29.93	-

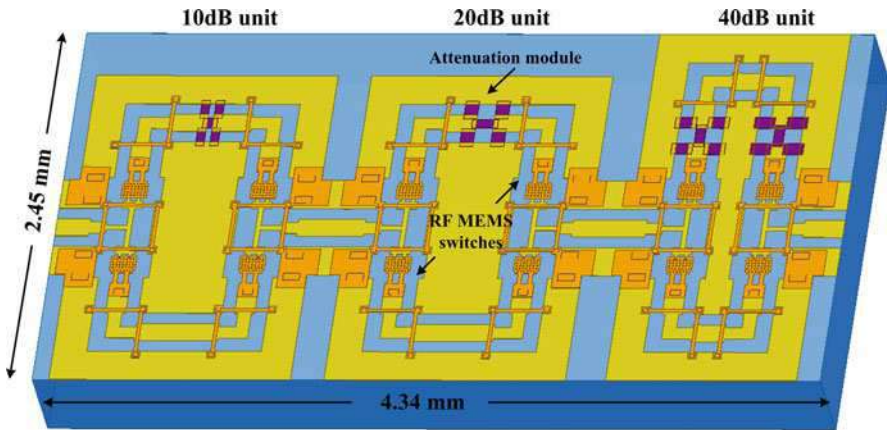
and move downwards to the bottom layer until contact with the dimples on the electrode. When the voltage is off, the cantilever will recover to its original position. The holes in the cantilever are for sacrificial layer releasing purpose and reducing the air damping during the actuation.

By careful structure optimization, the RF MEMS switch presents good RF performance with a relative low electrostatic actuation voltage between 20 ~ 30 V. From DC to 40 GHz, the insertion loss of the switches is better than 0.17 dB at “down” state and isolation is better than 23.08 dB at “up” state, respectively.

For reconfigurable attenuators, SPDT switch is prerequisite, which can be realized by placing a pair of RF MEMS switches at two sides of the signal path connected by a T-junction. To analyze the SPDT characteristics, it is 3D modelled and simulated. The insertion loss and isolation of SPDT is better than 0.15 dB and 16.74 dB, respectively. Compared to the single RF MEMS switch, the matching performance is improved thanks to the utility of T-junction structure.

Simulation results of each module in this section are listed in Table 5.





**Fig. 12** 3D FEA model of 3-bit attenuator based on MEMS switches

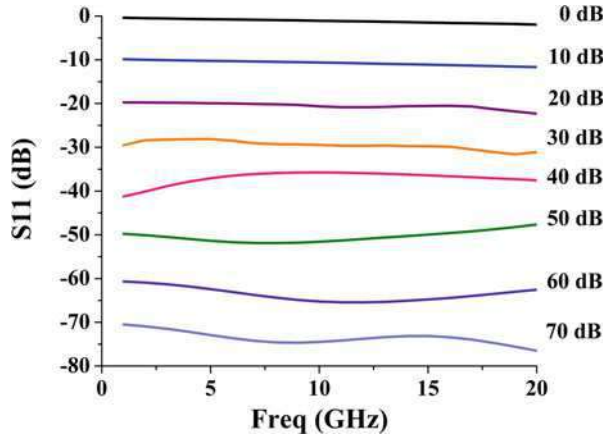
### Simulation of the 3-Bit Reconfigurable Attenuator

The 3-bit reconfigurable attenuator is shown in Fig. 12. It consists of three polysilicon resistive attenuation modules (10, 20, and 40 dB) and 6 SPDT RF MEMS switches including 12 single switches and compensate structures. With the signal path configured by switching off/on the MEMS switches in different status, it achieved 0 ~ 70 dB attenuation at intervals of 10 dB. The topology of the 40-dB unit is optimized by redistributing the two 20-dB attenuation modules using a more compact layout (the length of the entire device was reduced by 300  $\mu\text{m}$ ) and improved transmission loss in the reference path.

Based on the geometrical parameters obtained in the analytical calculation presented in section “[Structure Design](#),” the 3-bit attenuator is further modeled using HFSS. In the simulations, the substrate material is set to HRS (500  $\mu\text{m}$ ). Au is used as main functional material for CPW, air bridges, and RF MEMS actuation cantilever. Al is used for the bottom electrode and driving pads. The thickness of CPW, air-bridge membrane, and RF MEMS switch cantilever are supposed to be 2  $\mu\text{m}$  according to the process capacity and design experience. The air bridge height is as same as the gap between the top electrode and bottom electrode of the switches. A polysilicon thin film with a thickness of 300 nm and sheet resistance of 127  $\Omega/\square$  was used for all of the resistor networks. The corresponding full-wave simulation results are shown in Fig. 13 and Table 6.

As illustrated in Fig. 13, the FEA simulation results corresponding to the 3-bit MEMS attenuator show outstanding RF performance up to 20 GHz. The attenuation accuracy of all the states is within 6.87% ( $\pm 2.75$  dB at 40 dB) with an error less than 5.02% (2.41 dB at 60 dB). The results show an excellent matching network, with return loss better than 23.75 dB (VSWR < 1.14) for all eight states. The 60-dB state had the best attenuation accuracy ( $< \pm 0.89$  dB) and the 10-dB state presents the best return loss (better than 37.70 dB).

**Fig. 13** Simulated attenuation results of the 3-bit attenuator based on RF MEMS switches



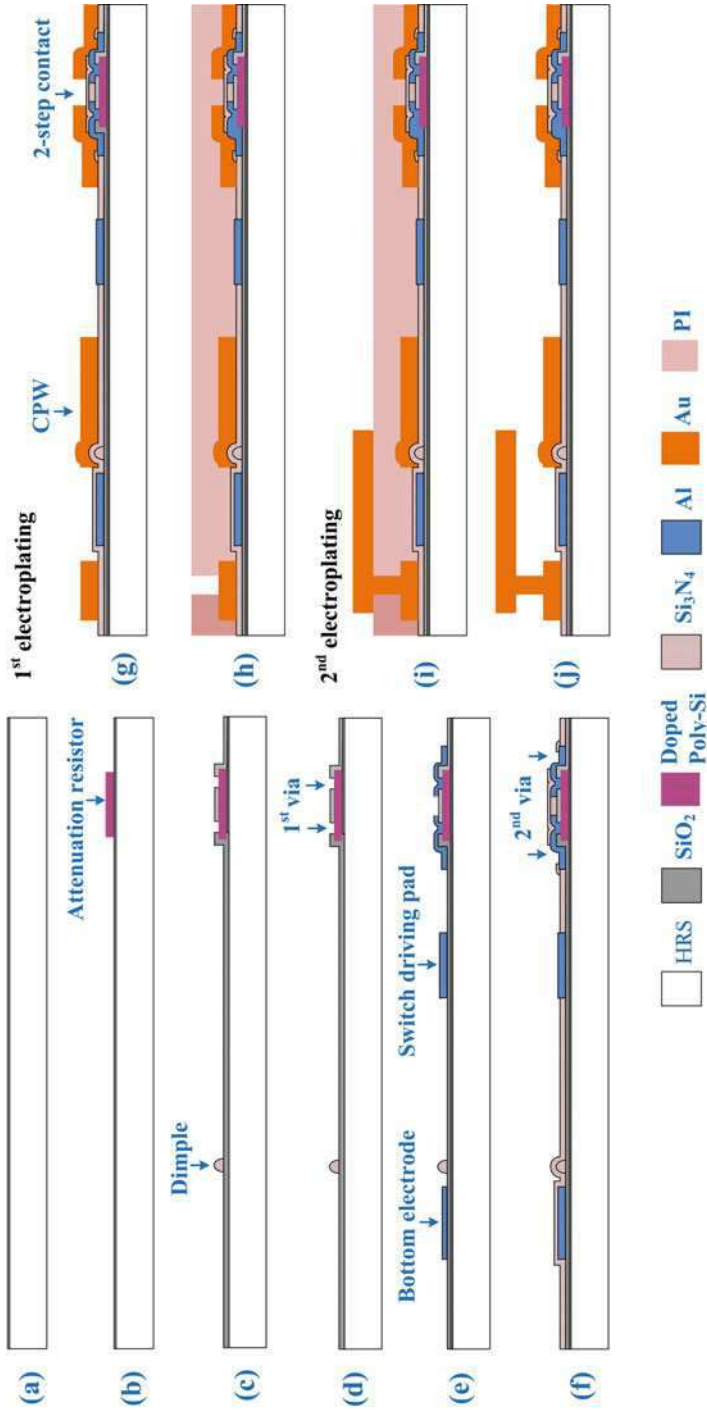
**Table 6** Simulation performance of the 3-bit attenuator

State	Target (dB)	Attenuation (dB)			Return loss (dB)	VSWR
		Center	accuracy	error		
000	0	1.14	$\pm 0.79$	1.14	24.33	1.13
001	10	10.76	$\pm 0.89$	0.76	33.05	1.05
010	20	21.00	$\pm 1.29$	1.00	29.44	1.08
011	30	29.85	$\pm 1.75$	0.15	36.23	1.03
100	40	38.49	$\pm 2.75$	1.51	26.44	1.10
101	50	49.77	$\pm 2.11$	0.03	37.70	1.03
110	60	63.01	$\pm 2.41$	3.01	23.75	1.14
111	70	72.56	$\pm 2.09$	2.56	31.46	1.06

## Fabrication

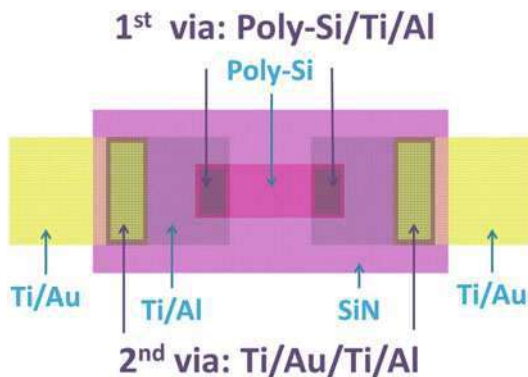
The 3-bit attenuator is manufactured with surface micromachining process and polysilicon IC techniques on 4 in. wafers and features 11 lithography steps. Polysilicon resistive network is chosen to be fabricated simultaneously with other components. The process flow is schematically depicted in Fig. 14.

The process started from a N-type (100)-oriented high-resistivity silicon ( $>5000 \Omega \text{ cm}$ ) wafers (Fig. 14a). Following the thermal oxidation, a 300-nm-thick polysilicon film used for attenuation network and isolation resistors was deposited onto a 240-nm-thick  $\text{SiO}_2$  layer at  $630^\circ \text{C}$  using a low-pressure chemical-vapor-deposition (LPCVD) system in  $\text{SiH}_4$  atmosphere. The polysilicon films were patterned with inductively coupled plasma (ICP) (Fig. 14b), followed with ion-implantation with Boron at 33 keV with a dose of  $2.9 \times 10^{15} \text{ cm}^{-2}$  and then annealed in a  $\text{N}_2$  atmosphere at  $1050^\circ \text{C}$  for 30 min. To obtain high-accuracy attenuation modules,



**Fig. 14** IC-compatible fabrication process of the 3-bit attenuator

**Fig. 15** Schematic view of the proposed 2-step via between the polysilicon thin film and gold CPW



the ion implanting and annealing condition was determined from intensive experimental studies and theoretical calculations for high-precision polysilicon thin film preparation process under a wide temperature range, as discussed in our previous work (Xin et al. 2013).

After that, a 300-nm-thick silicon oxide layer and a 300-nm-thick silicon nitride layer were plasma-enhanced chemical-vapor-deposited (PECVD) onto the resistor. The silicon nitride layer was etched to define the down-dimples of the MEMS switches (Fig. 14c), afterwards the silicon oxide is etched to open the first via (Fig. 14d).

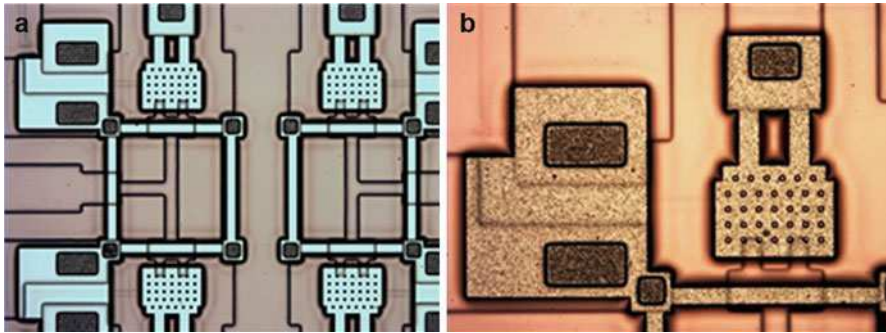
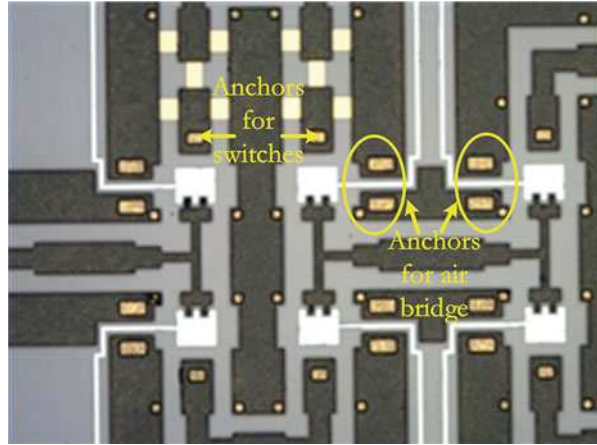
To realize perfect ohmic contact and complete isolation between the Au CPW and polysilicon thin film, the transition structure is carefully designed with a two-step via approach using Ti/Al/Ti as transition layer, as shown in Fig. 15.

Then the Ti/Au was electroplated onto the opening window of second Ti/Al via. Ti was used as adhesion layer. Ohmic contact with multimetall was realized using this methodology. The multilayer structure was simulated to optimize the attenuation performance. A 100-nm/500-nm-thick Ti/Al layer was patterned as the bottom electrode for switch and resistor metal pads in one evaporation and lift-off step. The polySi–Ti–Al contact (first via) was alloyed using an RTA process at 420 °C for 20 min (Fig. 14e). After the polysilicon thin film resistor was fabricated, another 300-nm-thick SiN was deposited as a passivation layer. The second via was opened with a SiN window etching (Fig. 14f). A seed layer of Ti/Au (50 nm/150 nm) was sputtered onto the surface and 2- $\mu\text{m}$ -thick Au layer, used as the CPW transmission lines, was electroplated and the ohmic contact with multimetall was realized.

The Au electroplating was performed with low stress approach using in-house made system. The wafer is electroplated in 60 °C Au plating bath with speed of 0.28  $\mu\text{m}/\text{min}$ .

Following the electroplating process, the second via was covered with Ti/Al/Ti/Au and resistors were electrically connected to the transmission line (Fig. 14g). A 3- $\mu\text{m}$ -thick photosensitive polyimide, which was used as a sacrificial layer, was spun and etched to form the anchor area (Fig. 14h). Normally, photoresist has to be used when patterning traditional PI layer, which requires a single-layer exposure and two-membrane (photoresist is above PI layer) etching and may induce a controlling

**Fig. 16** Microscope image of attenuator fabrication result after PI patterning



**Fig. 17** Microscope image of the MEMS switch upper electrode structures (a) before (b) after electroplating

issue to obtain the required PI layer shape with acceptable size and edge. An alternative solution is to use photosensitive PI as the sacrificial layer and pattern it directly. As illustrated in Fig. 16, the device is covered well with PI except the windows with lighter color, which are the anchors for MEMS switch cantilever. The windows indicate that the patterned PI keeps well the shape size with the designed layout and the errors at the edge keeps very small.

To form the top electrode of the switch and the air bridges (Fig. 14i), an additional 2- $\mu\text{m}$ -thick Au layer was then electroplated onto the sacrificial layer. Figure 17a gives the image after Ti/Au seed layer sputtering, afterwards the Au cantilever was obtained with Au electroplating, as shown in Fig. 17b.

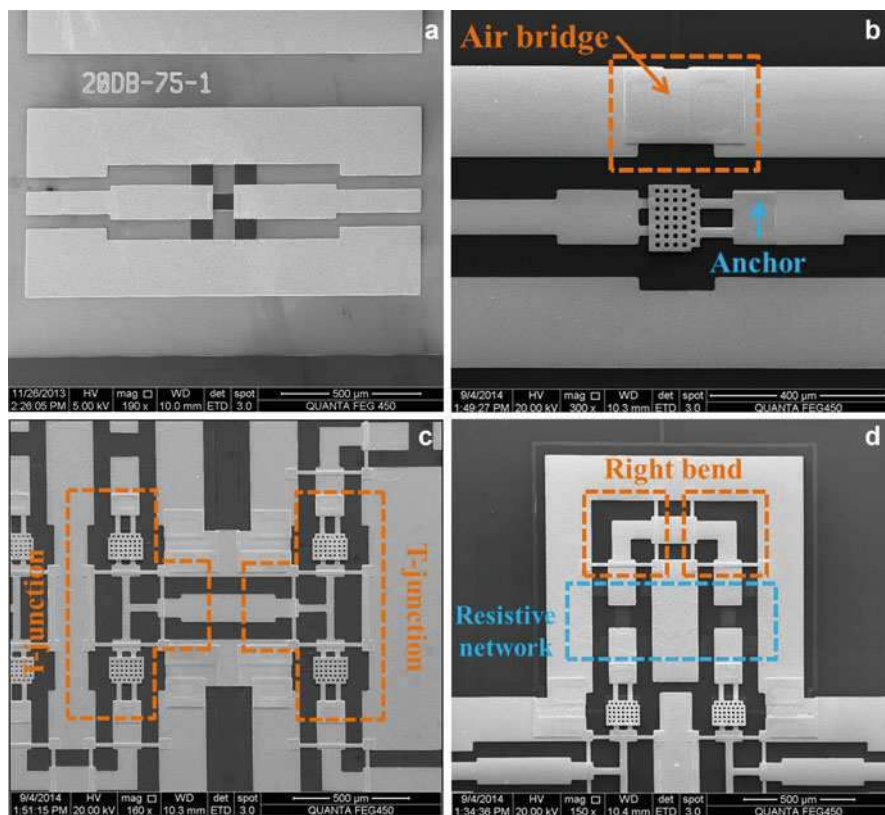
Finally,  $\text{O}_2$  plasma was used to remove the polyimide sacrificial layer, and the MEMS switches and air bridge structures were released (Fig. 14j). The obtained results show the MEMS process for switch fabrication is compatible with main steps (implantation) of the standard integrated circuit (IC) techniques.

## Characterization

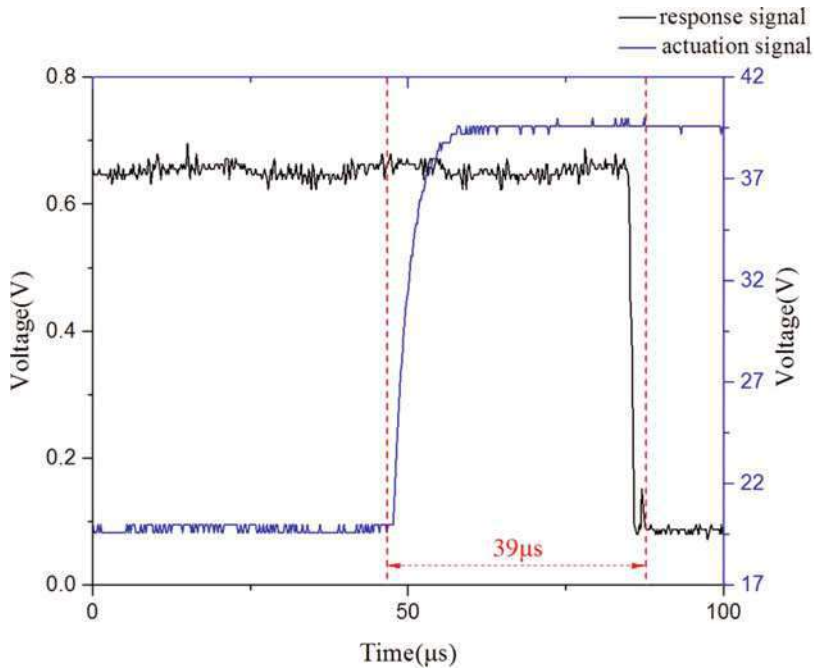
The 3-bit attenuator using polysilicon resistive networks was successfully fabricated. Characterization of the devices was carried out using SEM and network analyzer to evaluate the attenuation, matching performance within the whole desired operating band. The RF MEMS switches and polysilicon resistive network are also independently measured and discussed in this section.

## Structural Observation

The details of the fabricated 3-bit reconfigurable attenuator were observed by using SEM as shown in Fig. 18. The images demonstrated that the attenuation module, RF MEMS switches, T-junction, and right-angle bend compensate structures were well



**Fig. 18** SEM images and photos of fabricated reconfigurable attenuator. (a) 20 dB module. (b) RF MEMS switch. (c) T-junction. (d) right-bend



**Fig. 19** Waveform of the RF MEMS switch under DC test

patterned and over the CPW (Fig. 18a–d, respectively). Figure 18a presents test structures of 20 dB polysilicon attenuation network. Figure 18b shows that the switch was sufficiently released with small roughness and without evident buckling.

## DC Electrical Characterization

### SPDT Switch

The switch used in reconfigurable attenuator is characterized by applying a square wave modulated DC voltage, as shown in Fig. 19. It is found that the driven voltage of the RF MEMS switches is in a range of 32 ~ 42 V. It is higher than the designed values and different from each other. This is due to the process dispersion and lateral geometric errors produced during the patterning and vertical geometric errors produced in resist preparation and electroplating process.

### Polysilicon Resistor

The polysilicon resistance is critical to the accuracy of the attenuation performance. DC measurements were performed using an Agilent B1500A to characterize the resistivity of resistors in attenuation network.

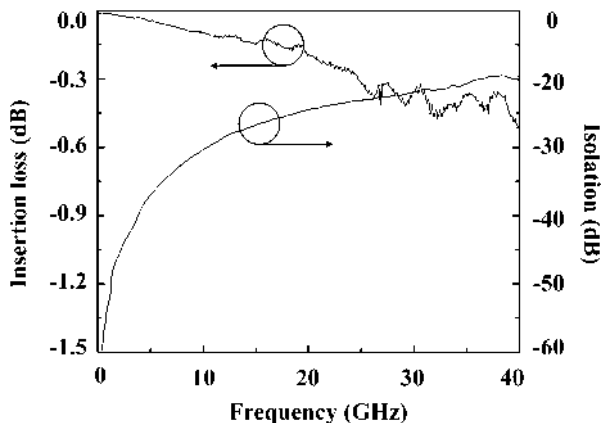
The average values of  $R_{series}$  and  $R_{shunt}$  in the 10- and 20-dB modules are summarized in Table 7. The measured resistances agree with the designed values



**Table 7** Designed value and measured value of attenuation modules at different dimensions

	10 dB		20 dB	
	$R_{series}$	$R_{shunt}$	$R_{series}$	$R_{shunt}$
L/W ( $\mu\text{m}/\mu\text{m}$ )	35/60	75/50	105/60	75/80
Designed value ( $\Omega$ )	68	190	228	120
Measured value <sup>a</sup> ( $\Omega$ )	72.1	185.4	216.7	116.3

<sup>a</sup>The average value of tests on 10 random dies

**Fig. 20** Microwave test results of RF MEMS switch

with an error less than 6%. The errors due to the processing conditions were derived theoretically in our previous work (Xin et al. 2013). Generally, these results demonstrate the validity of the prediction approach based on a combination of curve fitting and grain size calculation.

## RF Measurement

To verify the function of the device, microwave performance of RF MEMS switch, three attenuation units, and the overall attenuator were measured. All the RF measurements were performed using an Agilent N5247A PNA-X vector network analyzer and on-wafer thru-reflect-line (TRL) calibrations. The analyzer provides a microwave testing signal with power of 20 mW and a sweep frequency range of DC ~ 20GHz (400 points). The microwave probes with GSG structure at each port are customized for RF tests.

## SPDT Switch

The insertion loss at “down” state and isolation at “up” state of RF MEMS switch were tested from DC to 40 GHz. As revealed in Fig. 20, the insertion loss is less than  $-0.5$  dB@ DC ~ 40 GHz, the isolation is  $-36$  dB@5 GHz,  $-30$  dB@10 GHz, and



-24 dB @20 GHz. The excellent performance of the switch plays a crucial role in realizing high performance reconfigurable attenuator.

### 20 dB Polysilicon Resistors

20 dB network is taken as the example to characterize and analyze the performance of polysilicon resistive networks. The attenuators were measured at room temperature (25 °C/289 K). The simulation and test results of S parameters are compared in Fig. 21. It indicates that up to 20GHz, the attenuation S21 is  $18.62 \pm 0.15$  dB and the return loss (RL) S11 is better than -17.77 dB (VSWR >1.29). The resistive network presents promising performance in operation frequency range, attenuation accuracy, and practical VSWR.

The deviation between measurements and predictions can be attributed to the resistance variations, which is critical to both the matching and attenuation accuracy performances. The difference of measured and simulated value of S11 is up to 10 dB (Fig. 21a), even though it is still good enough with a value around -20 dB at all frequency band interested. The difference of attenuation amplitude (S21) is relatively small in order of 2 dB (Fig. 21b). This may be caused by the resistance change due to the process fluctuations. In fact, there are many factors which affect the resistance of network resistors: the doping concentration, the lateral dimensions, the polysilicon film thickness and the contact between polysilicon and CPW. To analyze the cause of the derivation, careful measurement of the resistor size is performed again by using SEM. It is found that the lateral structural dimension deviations of attenuator are up to 5 μm, which results from the photolithography fluctuation and consequent etching process. The designed and measured parameters are listed in Table 8.

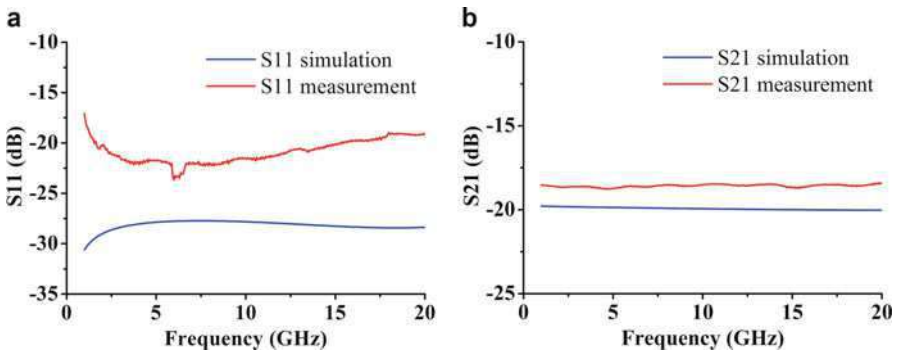


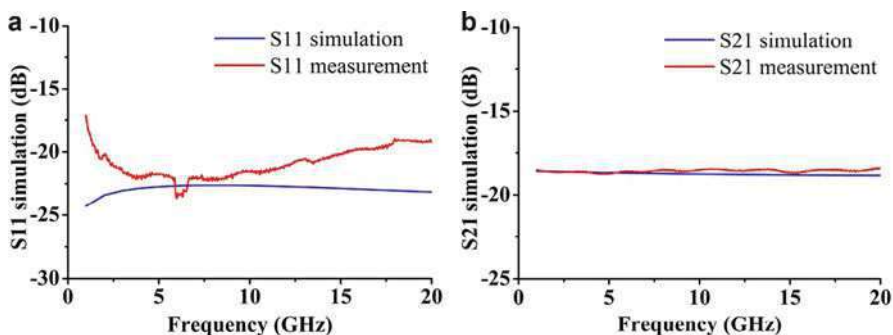
Fig. 21 Simulation and measurement results of the 20 dB resistive network. (a) S11. (b) S21

Table 8 Comparison of the designed and measured 20 dB resistor dimensions

	$R_{se}$	$R_{sh}$	CPW (center)	CPW (port)
Designed value (μm)	105/60	75/80	75/120 /200	52/80/243
Measured value (μm)	106/62	74/82	76/122/205	51/81/247

**Table 9** Comparison of the designed and measured resistor layer thickness

	Silicon dioxide	Doped-polysilicon	Silicon nitride
<b>Designed value (nm)</b>	240	300	300
<b>Measured value (nm)</b>	189	260	242

**Fig. 22** Performance verification with measured structural parameters. (a) S11 parameter (b) S21 parameter

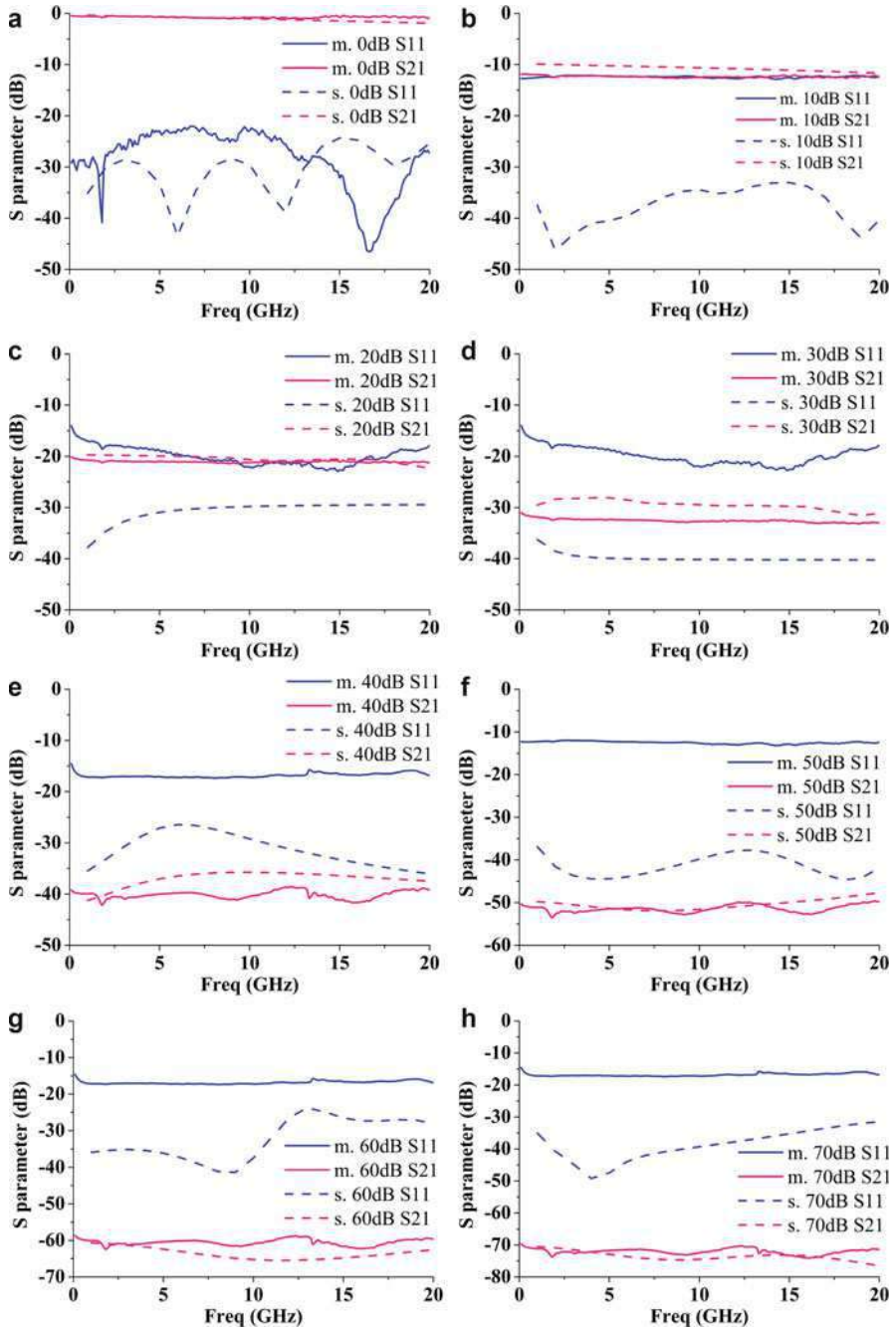
Another factor to influence the resistance is the polysilicon film thickness. Theoretically the final film thickness is determined precisely by growth rate and the time. But the rate fluctuation exists from batch to batch and the film difference in different position of one wafer is also observed. The layer thickness of the film was measured with Focused Ion Beam-SEM (LYRA3 TESCAN), as listed in Table 9.

Based on measured dimensions and thickness given in Tables 8 and 9, the resistive network is re-simulated. As shown in Fig. 22, the measurement performances show better consistency with the re-modeled simulation results, which means a precise process control is very important for manufacturing of the 3-bit attenuator.

### 3-Bit MEMS Attenuator

Twelve switches need to be actuated to characterize the 3-bit MEMS attenuator and it is not convenient to apply 6 DC probes in the given test bench. We designed a testing system including two PCB boards: DC-driven board and test board. The DC-driven board functions as the controlling part of the system and target attenuation value is determined by the manual switches. Driven voltage of the RF MEMS switches is boosted to 20 ~ 80 V using chip LT3482 and output to the driven pads of RF MEMS switches. The attenuator chip is Au-wire-bonded onto the test board with no DC probes.

Compared with the simulation predictions, the experimental results of all eight states obtained with the system are depicted in Fig. 23a–h and summarized in Table 10. Attenuation and matching performance of the 3-bit MEMS-based attenuator are discussed and compared to other achievements on miniaturized microwave attenuators using active or MEMS switches, as listed in Table 11. From Tables 10 and 11, it is found that:



**Fig. 23** Comparison of simulation and measurement results of the 3-bit attenuator in all eight states. *m.* measured *s.:* simulated

**Table 10** Measured performance of the fabricated 3-bit attenuator

State	Target (dB)	Attenuation (dB)				Return loss (dB)	VSWR
		center	accuracy	error	deviation <sup>a</sup>		
00	0	0.73	± 0.30	0.73	0.41	22.15	1.17
001	10	12.22	± 0.42	2.22	1.43	12.10	1.69
010	20	20.72	± 0.66	0.72	0.28	14.04	1.51
011	30	32.11	± 1.09	2.11	2.26	13.94	1.52
100	40	40.37	± 1.82	0.37	1.88	14.53	1.47
101	50	51.78	± 1.82	1.78	2.01	11.95	1.71
110	60	60.61	± 1.88	0.61	2.40	14.53	1.47
111	70	71.99	± 1.66	1.99	0.57	13.52	1.55

<sup>a</sup>The deviation refers to the difference between simulated and measured attenuation center values

**Table 11** Performance comparison of the microwave attenuators

Ref.	Technology	Freq. (GHz)	Num. of bit	Attenuation (dB)	Return loss (dB)
Yong-Sheng et al. (2007)	GaAs MESFET	0.045 ~ 50	6	2 ~ 70	11
Hyunchul et al. (2008)	InP/ InGaAs PIN	6 ~ 20	3	Up to 28	10
Marcus et al. (2011)	SOI FET	0.05 ~ 4	1	Up to 34	12
Bae et al. (2017)	BiCMOS	22 ~ 29/ 57 ~ 64	4	1 ~ 16	10
Iannacci et al. (2017)	MEMS	0.01 ~ 110	8	10 ~ 60	N/A <sup>a</sup>
Junfeng et al. (2016)	MEMS	DC ~ 20	3	0 ~ 35	12
This chapter	MEMS	DC ~ 20	3	0 ~ 70	12

<sup>a</sup>The return loss of this work was shown in figures but no exact value was mentioned in the text

1. For all eight states, the experimental results indicate that the attenuator achieves 0 ~ 70 dB attenuation at 10 dB intervals with an accuracy of better than ±1.88 dB (at 60 dB) and error better than 2.22 dB (at 10 dB). The difference between simulated and tested attenuation center value is less than 2.40 dB (at 60 dB) and can be as small as 0.28 dB (at 20 dB). Excellent attenuation performance of a wide range is demonstrated by comparison.
2. The measured return loss of the attenuator was better than 11.95 dB (VSWR <1.69), which shows large disagreement with the simulation result 23.75 dB (VSWR <1.14). The deviation can be attributed to modelling settings and process fluctuations. Nevertheless, the fabricated attenuator show same level of matching performance compared with the results in other works. The matching performance of multibit MEMS-based attenuator in wide operating frequency band should be improved in future work.

---

## Reliability Test

Reliability of the device is significant for practical application. It is determined primarily by the RF MEMS switch due to the possible failure of its mechanical movable structure (Chih-Hsiang et al. 2013). We characterized the operating cycles of the MEMS switches by adding a 4-kHz square wave signal to a basic DC-driven voltage, and we carefully design the actuation square wave parameter for stable up/down state (Anna et al. 2012). The fabricated switch was able to switch up/down continuously without failure for more than 21 h, which is equivalent to a cycle (cold switched condition) of more than  $3 \times 10^8$ . Furthermore, after a 1.5-year restoration in  $N_2$  atmosphere without packaging, the reconfigurable attenuator was tested again and the measured RF performance was consistent with the results obtained immediately after fabrication.

---

## Conclusion

Reconfigurable microwave device such as attenuator is a significant application domain of RF MEMS technique. An overview of recent research on reconfigurable attenuator is presented and discussed. A miniaturized reconfigurable 3-bit attenuator based on RF MEMS switches and polysilicon resistors is proposed and successfully fabricated with IC-compatible surface micromachining process. As a microwave device, the characteristic of the attenuator is determined by various factors, such as materials, structure topology, process flow. Performance modeling based on building blocks is essential to obtain a desired device. Customized design and optimization of each block should be performed elaborately and systematically. The reconfigurable attenuator with 0- to 70-dB attenuation magnitude at an interval of 10 dB is successfully realized and characterized. The attenuator exhibits high precision within a broad band range up to 20 GHz. The attenuation accuracy was better than 3.13% ( $\pm 1.88$  dB at 60 dB) with an error less than 2.22 dB for all eight states and the return loss better than 11.95 dB. The total device size is of  $2.45 \times 4.34 \times 0.5$  mm<sup>3</sup>. The current achieved results prove that RF MEMS provide an effective strategy to miniaturize the microwave/RF device. The RF MEMS switch could play an important role in the RF systems where high performance tuning/reconfiguration is important. Many challenges are still to be addressed from view of commercial applications such as improvement of the reliability, the productivity, and the packaging.

---

## References

- Anna P et al (2012) Reliability enhancement by suitable actuation waveforms for capacitive RF MEMS switches in III-V technology. *J Microelectromech Syst* 21(2):414–419
- Avago datasheet of AMMC-6650 DC–40 GHz variable attenuator
- Bae J et al (2017) A novel concurrent 22–29/57–64-GHz dual-band CMOS step attenuator with low phase variations. *IEEE Trans Microw Theory Tech* 64(6):1867–1875

- Bansal D et al (2017) Low voltage driven RF MEMS capacitive switch using reinforcement for reduced buckling. *J Micromech Microeng* 27(2):024001
- Cano JL et al (2010) Ultra-wideband chip attenuator for precise noise measurements at cryogenic temperatures. *IEEE Trans Microw Theory Tech* 58(9):2504–2510
- Chih-Hsiang K et al (2013) An electronically-scanned 1.8–2.1GHz base-station antenna using packaged high-reliability RF MEMS phase shifters. *IEEE Trans Microw Theory Tech* 61(2):979–985
- David AC et al (2013) RF MEMS switches with RuO<sub>2</sub>-Au contacts cycled to 10 billion cycles. *J Microelectromech Syst* 22(3):655–661
- David M et al (2016) Dielectric charging characterization of Triboelectric effects in MEMS switches. *J MEMS* 25(4):737–743
- Diego Correias S et al (2014) Graphene-based plasmonic tunable low-pass filters in the terahertz band. *IEEE Trans Nanotechnol* 13(6):1145–1153
- Hakan D et al (2008) Analysis and design of RF CMOS attenuators. *IEEE J Solid-St Circ* 43(10):2269–2283
- Hiroimi S et al (2004) Proposal and experimental study of a high-precision polycrystalline-silicon film resistor with a quasi-double-layer structure. *Electron Comm Jpn* 87(7):643–650
- Hyunchul E et al (2008) A 6–20GHz compact multi-bit digital attenuator using InP/InGaAs PIN Diodes IPRM 2008. In: 20th international conference on. Versailles, France. pp 1–8
- Iannacci J et al (2009) General purpose reconfigurable MEMS-based attenuator for radio frequency and microwave applications. In: EUROCON '09. IEEE. St Petersburg, Russia. pp 1197–1205
- Iannacci J et al (2016a) RF-MEMS technology for 5G: series and shunt attenuator modules demonstrated up to 110 GHz. *IEEE Electron Device Lett* 37(10):1336–1339. <https://doi.org/10.1109/LED.2016.2604426>
- Iannacci J et al (2016b) RF-MEMS for 5G mobile communications: a basic attenuator module demonstrated up to 50 GHz. *Proc IEEE Sens* 2016b:448–450. <https://doi.org/10.1109/ICSENS.2016.7808547>
- Iannacci J et al (2016c) RF-MEMS technology for future (5G) mobile and high-frequency applications: reconfigurable 8-Bit power attenuator tested up to 110 GHz. *IEEE Electron Device Lett* 37(12):1646–1649. <https://doi.org/10.1109/LED.2016.2623328>
- Iannacci J et al (2017) RF-MEMS for future mobile applications: experimental verification of a reconfigurable 8-bit power attenuator up to 110 GHz. *J Micromech Microeng* 27(4):1–11. <https://doi.org/10.1088/1361-6439/aa5f2c>
- Jiang HC et al (2010) Microwave power thin film resistors for high frequency and high power load applications. *Appl Phys Lett* 97(17):173504
- Jing W et al (2004) 1.156-GHz self-aligned vibrating micromechanical disk resonator. *IEEE Trans Ultrason Ferroelectr Freq Control* 51(12):1607
- Junfeng S et al (2016) A broadband DC to 20 GHz 3-bit MEMS digital attenuator. *J Micromech Microeng* 26(5):1–7
- Kim M et al (2001a) A DC-to-40GHz Four-Bit RF MEMS True-time delay network. *IEEE Microw Wireless Compon Lett* 11(2):56–58
- Kim M et al (2001b) A DC-to-40 GHz four-bit RF MEMS true-time delay network. *IEEE Microw Wirel Co Lett* 11(2):56–58
- Lee S et al (2004) Low-loss analog and digital reflection-type MEMS phase shifters with 1: 3 bandwidth. *IEEE Trans Microw Theory Tech* 52(1):211–219
- Ling L et al (2013) DC-contact radio-frequency microelectromechanical system symmetric toggle switch on a borofloat substrate. *Micro Nano Lett* 8(5):221–224
- Lucyszyn S (2010) *Advanced RF MEMS*. Cambridge University Press, New York
- Malhaire C et al (2003) Design of a polysilicon-on-insulator pressure sensor with original polysilicon layout for harsh environment. *Thin Solid Films* 427:362–366
- Marcus GJ et al (2011) A Broadband High Dynamic Range Voltage Controlled Attenuator MMIC with IIP3 > +47dBm over Entire 30dB Analog Control Range. *IEEE MTT-S Int Microw Symp Dig*, Baltimore, MD, USA 1–4

- Mingchun H (2010) Technology of phase-arrayed radar T/R modules. National Defence Industry Press, Beijing
- Ohzone T et al (1985) Ion-implanted Thin Polycrystalline-Silicon High-Value Resistors for High-Density Poly-Load Static RAM Applications. *IEEE Trans Electron Devices* 32(9):1749–1755
- Osseiran A et al (2014) Scenarios for 5G mobile and wireless communications: the vision of the METIS project. *IEEE Commun Mag* 52(5):26–35
- Petersen KE (1979) Micro-mechanical membrane switches on silicon. *IBM J Res Dev* 23(4): 376–385
- Petersen KE (1982) Silicon as a mechanical material. *Proc IEEE* 70(5):420–457
- Rebeiz GM et al (2001) RF MEMS switches and switch circuits. *IEEE Microw Mag* 1, 2(4):59–71
- Rebeiz GM et al (2003) Intermodulation distortion and power handling in RF MEMS switches, varactors, and tunable filters. *IEEE Trans Microw Theory Tech* 51(4):1247–1256
- Rebeiz GM et al (2012) A high-reliability high-linearity high-power RF MEMS metal-contact switch for DC-40-GHz applications. *IEEE Trans Microw Theory Tech* 60(10):3096–3112
- Rebeiz GM et al (2015) Symmetric and compact single-pole multiple-throw (SP7T, SP11T) RF MEMS switches. *J MEMS* 24(3):685–695
- Sanghyo L et al (2004) Low-loss analog and digital reflection-type MEMS phase shifters with 1: 3 bandwidth. *IEEE Trans Microw Theory Tech* 52(1):211–219
- Sekar V et al (2011) 1.2-1.6-GHz substrate-integrated-waveguide RF MEMS tunable filter. *IEEE Trans Microw Theory Tech* 59(4):866–876
- Simon O et al (2009) A distributed attenuator for K-band using standard SMD thin-film chip resistors. In: *Microwave conference, Asia Pacific*, pp 2148–2151
- Stukach OV (2006) Variable attenuator with low phase shift. In: *The 9th European conference on wireless technology*. Manchester, UK. pp 241–244
- Sukomal D et al (2015) Reliability analysis of Ku-Band 5-bit phase shifters using MEMS SP4T and SPDT switches. *IEEE Trans Microw Theory Tech* 63(12):3997–4012
- Van Caekenberghe K (2009) RF MEMS technology for radar sensors. In: *Radar conference – surveillance for a safer world*. IEEE, pp 1–6
- Vikram S et al (2011) 1.2-1.6-GHz substrate-integrated-waveguide RF MEMS tunable filter. *IEEE Trans Microw Theory Tech* 59(4):866–876
- Weller TM et al (1998) Optimization of MM-wave distribution networks using silicon-based CPW. *IEEE MTT-S Int Microw Symp Dig* 2:537–540
- Weston PW et al (1997) EM-ANN modeling and optimal chamfering of 90 Deg CPW bends with air-bridges. In: *IEEE MTT-S Digest*. pp 1603–1606
- Xin G et al (2013) Study on polysilicon control with nanoscale grain size Nanotech. In: *13th IEEE conference on*. Beijing, China. pp 950–953
- Xin G et al (2016) A miniaturized reconfigurable broadband attenuator based on RF MEMS switches. *J Micromech Microeng* 26:1–8
- Yasir M et al (2017) Enhanced tunable microstrip attenuator based on few layer graphene flakes. *IEEE Micro Wire Comp Lett* 27(4):332
- Yeh J-H et al (2016) Microwave attenuators for use with quantum devices below 100 mK. *J Appl Phys* 121:224501
- Yong-Sheng D et al (2007) A Novel UWB (0.045–50GHz) digital/analog compatible MMIC variable attenuator with low insertion phase shift and large dynamic range. *IEEE Microw Wirel Co Lett* 17(1):61–63
- Zavracky PM et al (1997) Micromechanical switches fabricated using nickel surface micro-machining. *J Microelectromech Syst* 6(1):3–9
- Zheng L et al (2016) A novel microwave attenuator on multilayered substrate integrated waveguide. *IEEE Trans Compon Packag Manuf Technol* 6(7):1106–1112
- Zhihao H et al (2009) Design and fabrication of DC to 30 GHz DC-contact shunt RF MEMS switch. *Opt Precis Eng* 17(8):1922–1927
- Zhong Q, Liu Z (2016) Wideband attenuators using distributed resistors for attenuation up to 30 dB. *IEICE Electron Exp* 13(10):1–9



# High-Performance Acoustic Devices for Wireless Communication and Sensor Applications

Changjian Zhou, Xiangguang Tian, and Tian-Ling Ren

## Contents

Overview .....	1150
High-Performance Surface Acoustic Wave Devices .....	1151
Fundamental Concepts of SAW Devices .....	1151
Multilayered Structures for High-Frequency SAW Devices .....	1153
Temperature-Compensated SAW Devices .....	1159
Super High Frequency SAW Devices .....	1162
Flexible Pseudo-SAW Devices .....	1166
Sensors Based on SAW Devices .....	1171
DNA Biosensor Based on SAW Device .....	1176
High-Performance Film Bulk Acoustic Wave Resonators .....	1177
Basic Structures and Fabrication Process .....	1178
Integration Strategy and System-Level Simulation .....	1183
Flexible FBAR .....	1186
Summary and Outlook .....	1191
References .....	1192

## Abstract

Surface acoustic wave (SAW) resonators and film bulk acoustic resonators (FBAR) are two of the most important passive resonators. After several decades of research and development, they are widely applied for wireless communications such as filters and duplexers, and broad sensing applications like pressure sensors, temperature sensors, gas sensors, biosensors, photoelectric sensors, and

---

C. Zhou

School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China

e-mail: [zhoucj@scut.edu.cn](mailto:zhoucj@scut.edu.cn)

X. Tian · T.-L. Ren (✉)

Institute of Microelectronics, Tsinghua University, Beijing, China

e-mail: [RenTL@tsinghua.edu.cn](mailto:RenTL@tsinghua.edu.cn)



many others. In this chapter, we first introduce basic concepts and mainstream topics of SAW and FBAR devices. Then we give a detailed presentation of several methods to fabricate high-performance resonators including using multilayered structures and improved process. Flexible SAW and FBAR devices are also introduced. Several sensing applications are given afterwards. Finally, a summary and outlook is presented regarding the further improvement of device performance and new research fields.

---

**Keywords**

Acoustic wave device · SAW · FBAR · Micro-nano machining · Sensors

---

**Overview**

SAW and FBAR devices are important electronic devices. They have received intense research and development efforts since several decades ago. Among the numerous applications in our daily life, radio-frequency (RF) filters in mobile phone and television, and various sensing applications like pressure sensors in automotive tire pressure monitoring, temperature sensors, and biosensors are the most known examples (Dubois and Muller 2013; Zhou et al. 2013a; Morgan 2010; Yeo and Friend 2014). The basic working principles of SAW and FBAR devices are similar, namely the piezoelectric and inverse piezoelectric effect. Coupling effect between mechanical and electric fields occurs, in other words, electric and mechanical energy can be transformed to each other. External electric signal at input port excites acoustic wave, then the wave propagates in the piezoelectric material and finally reaches output port to transform back to electric signal. The difference between SAW and FBAR device is the type of the wave mode. In the case of SAW, the wave mainly spreads near the surface while for FBAR the wave propagates in the whole device structure.

Commonly used piezoelectric material for acoustic wave device fabrication includes four types: single crystal, ceramics, polymer, and thin film (Setter et al. 2006; Ramadan et al. 2014; Shung et al. 2007). Typical examples, strength and weakness, together with application fields of the piezoelectric materials are listed in Table 1. Different materials are chosen according to their potential applications. At present, single crystal such as  $\text{LiNbO}_3$  and  $\text{LiTaO}_3$ , together with thin films like  $\text{AlN}$  and  $\text{ZnO}$  are the most commonly used piezoelectric materials for acoustic wave devices. Also, as these materials all have their own pros and cons, multilayered structures are also adopted to improve the overall performances of the acoustic wave devices.

Frequency characteristic is essential for acoustic wave resonators. For telecommunication applications, frequency characteristics of filters and duplexers are determined by the original resonators. For sensing applications, most of the acoustic wave sensors rely on the frequency shift versus other parameters to be detected. Resonance frequency  $f_r$ , quality factor  $Q$ , and coupling coefficient  $k^2$  are three key parameters for these resonators. Besides them, temperature coefficient of frequency (TCF) is also an important indicator. Basically, there are two kinds of methods to enhance

**Table 1** Typical piezoelectric materials for acoustic wave device fabrication

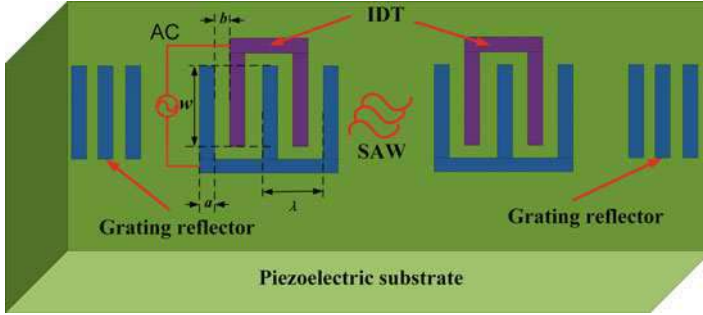
	Single crystal	Piezo ceramics	Piezo polymer	Piezoelectric thin film
Typical materials	Quartz, LiNbO <sub>3</sub> , LiTaO <sub>3</sub>	BaTiO <sub>3</sub> , PZT, PbTiO <sub>3</sub>	PVDF, PVF, PVC	AlN, ZnO, PZT, LiNbO <sub>3</sub>
Main strength	Low loss, high quality factor	High dielectric constants, high coupling coefficients	Soft, low acoustic impedance	Easy for integration
Main weakness	High manufacturing cost, hard in micromachining	Large loss, lack stability	Large loss	Hard to find optimal orientation
Application fields	Oscillator, filter	Ultrasonic imaging, high-k capacitors, high power transducer, wideband filter	Hydrophone, pressure sensor, ultrasonic measurement	High frequency device, FBAR, memory

device performance, including improving the piezoelectric material quality and the fabrication process. Many novel multilayered structures are fabricated to increase wave velocity, enhance the coupling coefficient, and improve temperature stability. Advanced fabrication techniques such as electron beam lithography are also adopted to achieve fine pattern for realizing ultra-high frequency acoustic wave devices.

## High-Performance Surface Acoustic Wave Devices

### Fundamental Concepts of SAW Devices

Interdigital transducers (IDT) consisting of multiple pairs of interdigital electrodes are fabricated on the piezoelectric substrate to excite and detect SAW. After the introduction of IDTs by White and Voltmer (1965), the potential of SAW devices for various applications becomes reality. A typical SAW device consists of one or more sets of comb-like IDTs to realize complex functions like filtering, phase delay, signal convolution, etc. To improve the quality factor, a SAW resonator usually consists of grating reflectors at both sides of IDTs to trap energy in the device area. Figure 1 shows the schematic structure of a two-port SAW resonator, which is consisting of two IDTs and two grating reflectors. The main parameters of an IDT includes the width of the electrodes  $a$ , pitch between electrodes  $b$ , periodic length  $\lambda = 2(a + b)$ , and the aperture  $w$ . If  $a$  equals to  $b$ , then the IDT is named as uniform transducer. As indicated in Fig. 1, when an Ac signal is applied to the input IDT, a periodic electrical potential induces periodic mechanical displacements in the piezoelectric substrate because of the inverse piezoelectric effect. The mechanical wave propagates along the piezoelectric surface and induces electrical charges in the output IDT because of the piezoelectric effect.



**Fig. 1** Schematic structure of a two-port SAW resonator

As each IDT consists of multiple pairs of electrodes, the excited acoustic wave will superimpose according to the wave interference principles at the receiving IDT. Assume that the IDT consists of  $N$  pairs of electrodes, the angular frequency of the applied voltage is  $\omega$ , and the SAW velocity is  $V_p$ , the phase difference between two adjacent pair of electrodes is given by

$$\Delta\theta = \frac{\omega}{V_p} \frac{\lambda}{2}. \tag{1}$$

And the output signal due to the  $N$ -pair IDT superimposition is expressed as

$$E_t = E_0 e^{i\omega t} \left[ 1 - e^{i\Delta\theta} + e^{i2\Delta\theta} - \dots + e^{i2(N-1)\Delta\theta} - e^{i(2N-1)\Delta\theta} \right], \tag{2}$$

where  $E_0$  is the amplitude of the excited wave for a single pair of electrode. It is found that  $E_t$  reaches maximum when  $\Delta\theta = \pi$ , which is

$$E_{t, \max} = 2NE_0 e^{i\omega t}. \tag{3}$$

The corresponding signal frequency is defined as the synchronism frequency or resonance frequency, and is expressed as

$$\omega = \omega_0 = 2\pi V_p / \lambda_0. \tag{4}$$

The wavelength  $\lambda_0$  is determined by the fabricated periodic length and has an inverse relationship with the resonance frequency.

For characterizing the SAW devices, network analyzer is usually used to measure the scattering parameters (S parameters).  $S_{11}$  and  $S_{21}$  are reflection and transmission coefficients respectively for a two-port SAW device. S parameters are also easily to be transformed to Y or Z parameter, and therefore circuit model of acoustic wave device is able to be derived from the measurement result.

## Multilayered Structures for High-Frequency SAW Devices

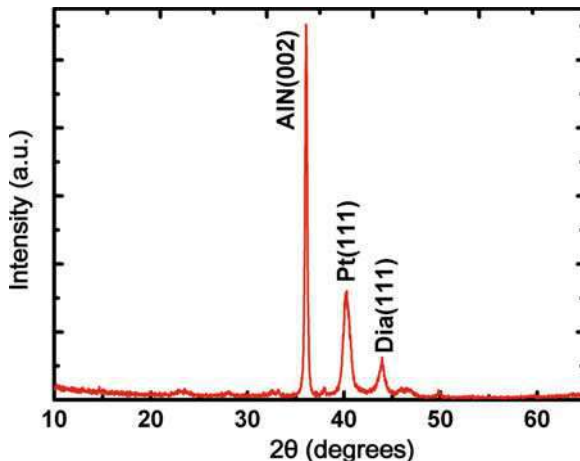
Conventional single-crystal piezoelectric substrates such as LiTaO<sub>3</sub>, LiNbO<sub>3</sub>, and Quartz possess limited SAW velocities, and it requires high-cost sub-micron lithography and ultrathin IDTs to fabricate high-frequency SAW devices for future wireless communication systems. The working frequency of a SAW device is given by  $f_r = V_p/\lambda$ , where  $V_p$  is the SAW velocity and  $\lambda$  is the SAW wavelength and determined by the IDT linewidth. To satisfy the high-frequency requirement, while at the same time ensuring the reliability of the SAW device under the working load of high-power RF signals, it is optimal to use relatively low-resistance, i.e., wider and thicker IDTs. In this regard, one has to choose piezoelectric substrate with higher SAW velocity which cannot be provided by single-crystal piezoelectric substrates. Fortunately, there are many materials like diamond, sapphire which have much higher acoustic velocities than conventional piezoelectric substrates, and the combination of these high-velocity materials with piezoelectric materials can potentially provide a multilayered structure with an overall much higher SAW velocity (Nakahata et al. 2003; Hashimoto et al. 2013). In the following, we demonstrate our work in integrating diamond film with two commonly used piezoelectric thin films, i.e., AlN and LiNbO<sub>3</sub>, for high-frequency SAW devices (Zhou et al. 2011, 2013b).

### SAW Resonators Based on AlN/Diamond/Si Multilayered Structures

#### Material Preparation and Device Fabrication

The diamond film used in the study is provided by Hebei Institute of Laser. The diamond film is about 10  $\mu\text{m}$  in thickness and directly deposited on a silicon wafer for following processes. The upper surface of the diamond film is polished to obtain an optical-grade surface, which is beneficial for the device fabrication. Prior to the deposition of AlN thin film, a seed layer of 50-nm Pt thin film is deposited on the diamond film. Then the 1.2  $\mu\text{m}$ -thick AlN thin film is prepared on the Pt thin film using DC magnetron sputtering method. The optimized parameters for obtaining (002) preferred orientation AlN are as follows: power 280 W, pressure 0.4 Pa, gas flow N<sub>2</sub> 65 sccm/Ar 50 sccm, target to substrate distance 70 mm, substrate temperature 450 °C, and a deposition speed of 20 nm/min. Figure 2 shows the x-ray diffraction (XRD) spectrum of the AlN/diamond/Si multilayered structure. The sputtered AlN film exhibits well (002) preferred crystal orientation, which is desirable for SAW excitation. It is also noticed that there are another two peaks corresponding to the (111) orientation of diamond and Pt. The root mean square of the surface roughness of AlN film is about 11 nm within a measured range of 5  $\mu\text{m} \times 5 \mu\text{m}$ . It is known that the attenuation and scattering of the acoustic wave are closely related to the surface properties of the substrate. Future advanced polishing process is required to reduce the surface roughness of the diamond film and improve the surface properties of the upper layers. To form the IDTs, a 100 nm aluminum thin film is evaporated on the patterned substrate followed by a lift-off process. Three kinds of two-port SAW resonators are fabricated in this study to investigate the dispersion relationship of the SAW velocity propagated in the

**Fig. 2** X-ray diffraction (XRD) spectrum of the AlN/diamond/silicon layered structure (Reprinted with permission of Elsevier)



multilayered structure. The linewidths of the SAW devices are 1.5  $\mu\text{m}$ , 1.78  $\mu\text{m}$ , and 3  $\mu\text{m}$ , respectively, and the SAW resonators corresponding to these linewidths are denoted as u1, u2, and u3 in the following discussions.

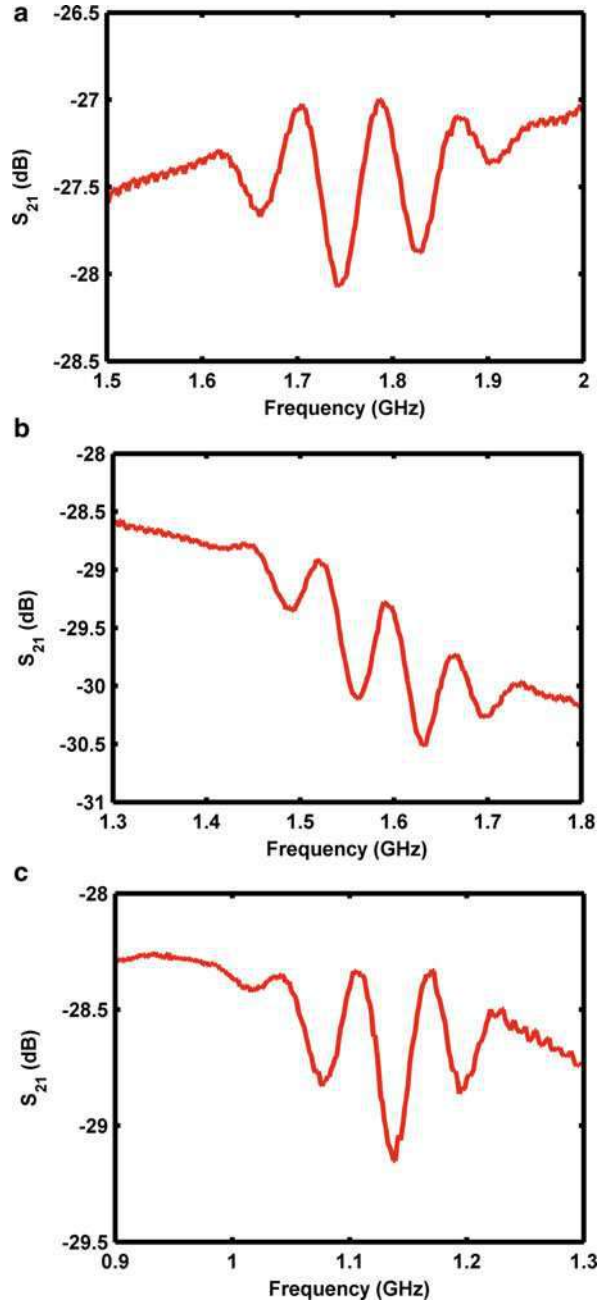
### Characterization of the SAW Resonators

Network analyzer HP8721s is used to measure the S-parameters of the SAW resonators. Figure 3 shows the measured  $S_{21}$  of three SAW resonators, i.e., u1, u2, and u3. All of the SAW resonators have five pairs of interdigital electrodes in the IDTs. The typical sinc function is clearly observed in all of the three SAW resonators, which is a representative of the frequency response due to the superimposition of multiple interdigital electrode (Campbell 1989). The resonance frequency and SAW velocity are listed in Table 2. Notably, a high frequency of 1.834 GHz is obtained with a 1.5  $\mu\text{m}$  linewidth SAW resonator, which gives an equivalent SAW velocity of 11,004 m/s. It is also shown that the SAW velocity increases with increasing linewidths. This can be understood as more energy is penetrated into the diamond with increasing wavelengths while keeping a constant thickness of AlN thin film. Figure 4 shows the dispersion relationship between the SAW velocity and the normalized AlN thickness  $KH_{\text{AlN}}$ , where  $K = 2\pi/\lambda$  and  $H_{\text{AlN}}$  is 1.2  $\mu\text{m}$ . Comparing with the theoretical calculation, one can notice that all the wave modes excited in the fabricated SAW devices are first-order probably due to the larger coupling coefficients of the first-order mode than other modes. It is also shown that the experimental SAW velocities are well matched with the theoretical calculations. One disadvantage of the AlN/diamond/Si multilayered structure lies in its low coupling coefficient, which limits its application to probably narrowband filters or sensors.

### SAW Resonators Based on LiNbO<sub>3</sub>/Diamond/Si Multilayered Structures

LiNbO<sub>3</sub> single-crystal substrate is widely used to fabricate SAW devices for its large coupling coefficient and low propagation loss. Thus it is promising to integrate LiNbO<sub>3</sub> with diamond for high-frequency large-bandwidth RF filters. Unlike AlN, it seems much more difficult to prepare high-quality LiNbO<sub>3</sub> thin films. Although

**Fig. 3** Measured  $S_{21}$  parameters of the SAW resonators with a linewidth of (a)  $1.55 \mu\text{m}$ , (b)  $1.78 \mu\text{m}$ , and (c)  $3 \mu\text{m}$  (Reprinted with permission of Elsevier)

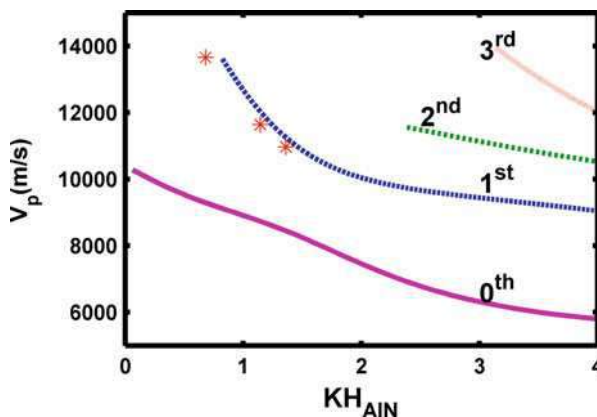


various methods such as sputtering, metalorganic decomposition, and pulsed laser deposition (PLD) have been reported to deposit  $\text{LiNbO}_3$  thin films for SAW devices, the reported coupling coefficients are much smaller than theoretically expected, and the propagation loss is very large (Dogheche et al. 2005; Lee et al. 2003). We have

**Table 2** Properties of the SAW resonators based on AlN/diamond/Si multilayered structure

	u1	u2	u3
$\lambda$ ( $\mu\text{m}$ )	6	7.128	12
$kh_{\text{AlN}}$	1.26	1.06	0.63
$f_0$ (GHz)	1.834	1.632	1.138
$V_p$ (m/s)	11,004	11,633	13,656

**Fig. 4** Dispersion relationship between the SAW velocity and the normalized AlN thickness  $KH_{\text{AlN}}$  (Reprinted with permission of Elsevier)

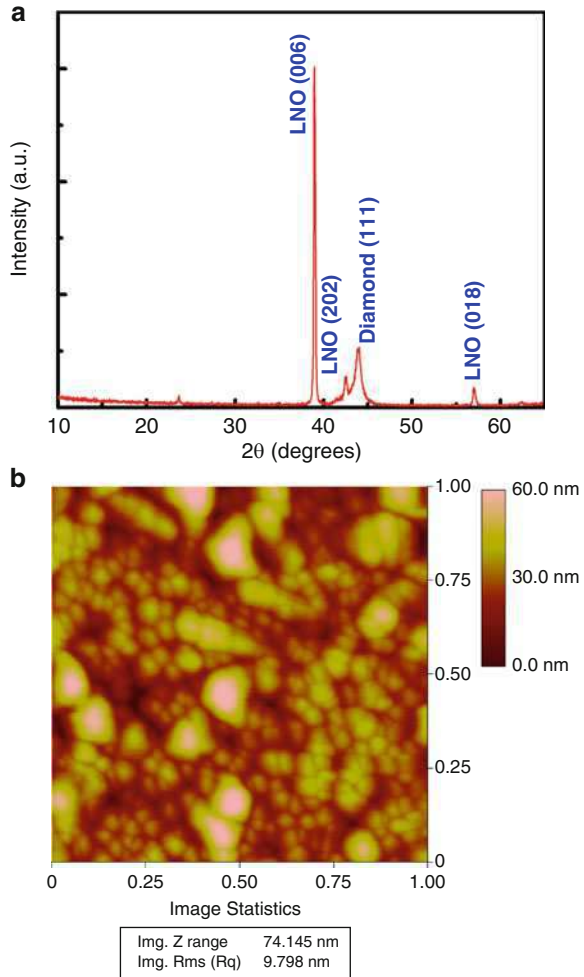


optimized the PLD process to obtain c-axis oriented  $\text{LiNbO}_3$  thin film and fabricated high-frequency SAW resonators based on the  $\text{LiNbO}_3$ /diamond/Si multilayered structure.

### Material Preparation and Device Fabrication

The PLD method is used to deposit a  $0.6 \mu\text{m}$   $\text{LiNbO}_3$  thin film on the diamond film using the following parameters: KrF laser pulse width 25 ns, repetition rate 3 Hz, substrate temperature  $650^\circ\text{C}$ ,  $\text{O}_2$  partial pressure 40 Pa, and target to substrate distance 40 mm. Figure 5a shows the XRD spectrum of the multilayered structure. It is shown that the  $\text{LiNbO}_3$  exhibits very good c-axis (006) preferred crystal orientation, and the (111) crystal orientation of the diamond is also clearly observed. There is also another peak corresponding to the (202) crystal orientation of  $\text{LiNbO}_3$  with a much smaller intensity than that of (006). This suggests that further optimization is still needed to get a perfect c-axis oriented  $\text{LiNbO}_3$  thin film. Figure 5b shows the atomic force microscopy (AFM) image of the  $\text{LiNbO}_3$  thin film measured in a range of  $1 \mu\text{m} \times 1 \mu\text{m}$ . The root mean square of the surface roughness is 9.8 nm. It is noting that there are some large crystal grains due to the aggregation of small grains during the PLD process, which results in a large propagation loss and a low coupling coefficient. In general, the surface is still smooth enough for the following lithography and lift-off processes. The fabrication processes to obtain the IDTs are the same as described above for AlN/diamond/silicon structure. Figure 6 shows a photograph of a two-port SAW resonator with five pairs of interdigital electrodes for each IDT.

**Fig. 5** (a) XRD spectrum of the LiNbO<sub>3</sub>/diamond/silicon multilayered structure, (b) AFM image of the LiNbO<sub>3</sub> thin film (Reprinted with permission of AIP)

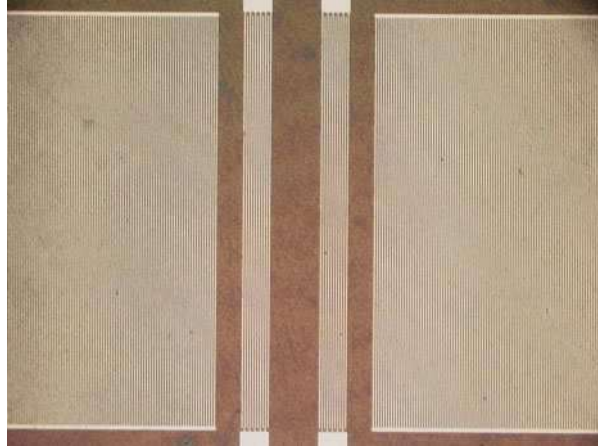


**Characterization of the SAW Resonators**

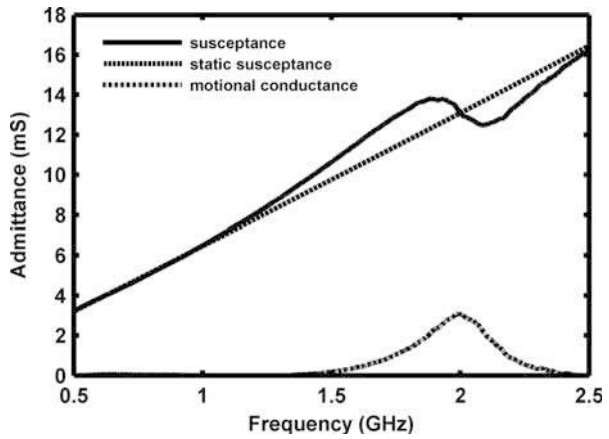
According to the equivalent circuit of the IDTs, the coupling coefficient is defined as  $K^2 = \pi G_m(f_0)/4NB_s(f_0)$ , where  $N$  is the number of pairs of electrodes in an IDT, and  $G_m(f_0)$  and  $B_s(f_0)$  are the motional conductance and static susceptance at the center frequency  $f_0$ . Figure 7 shows the admittance spectrum of a SAW resonator with a linewidth of 1.78 μm. The center frequency is 1.75 GHz, and  $G_m(f_0)$  and  $B_s(f_0)$  are 3.77 mS and 12.2 mS. The SAW velocity and coupling coefficient are 12,474 m/s and 4.85%, respectively. The SAW resonators with different wavelengths are measured and the performance parameters are calculated and summarized in Fig. 8. The experimental SAW velocities are quite consistent with the theoretical ones. Of all the four kinds of samples with different linewidths, both the zeroth and first order SAW mode can be excited with the SAW resonator of 1.78 μm linewidth.



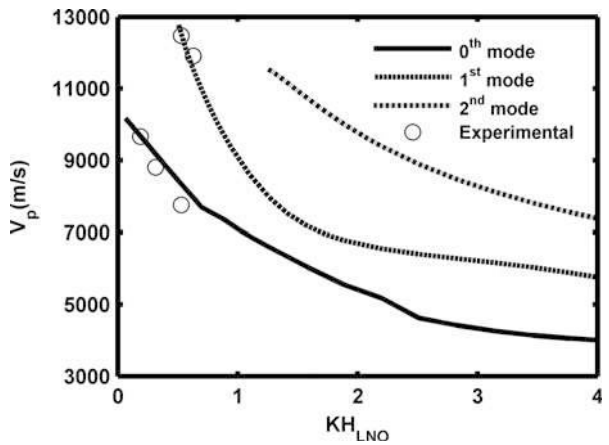
**Fig. 6** Optical photograph of a SAW resonator with linewidth of 1.5  $\mu\text{m}$  (Reprinted with permission of AIP)



**Fig. 7** Measured admittance spectrum of a SAW resonator with linewidth of 1.78  $\mu\text{m}$  (Reprinted with permission of AIP)



**Fig. 8** Dispersion relationship between the SAW velocity and the normalized  $\text{LiNbO}_3$  thickness  $KH_{\text{LNO}}$  (Reprinted with permission of AIP)



Comparing these two multilayered structure, we notice that both structures can provide very high SAW velocities larger than 11,000 m/s, which are very promising for future high-frequency SAW devices. In terms of coupling coefficients, the LiNbO<sub>3</sub>/diamond/silicon multilayered structure can provide much higher coupling coefficients, while experimentally future optimization is needed to obtain perfect c-axis oriented LiNbO<sub>3</sub> thin film with lower propagation losses.

## Temperature-Compensated SAW Devices

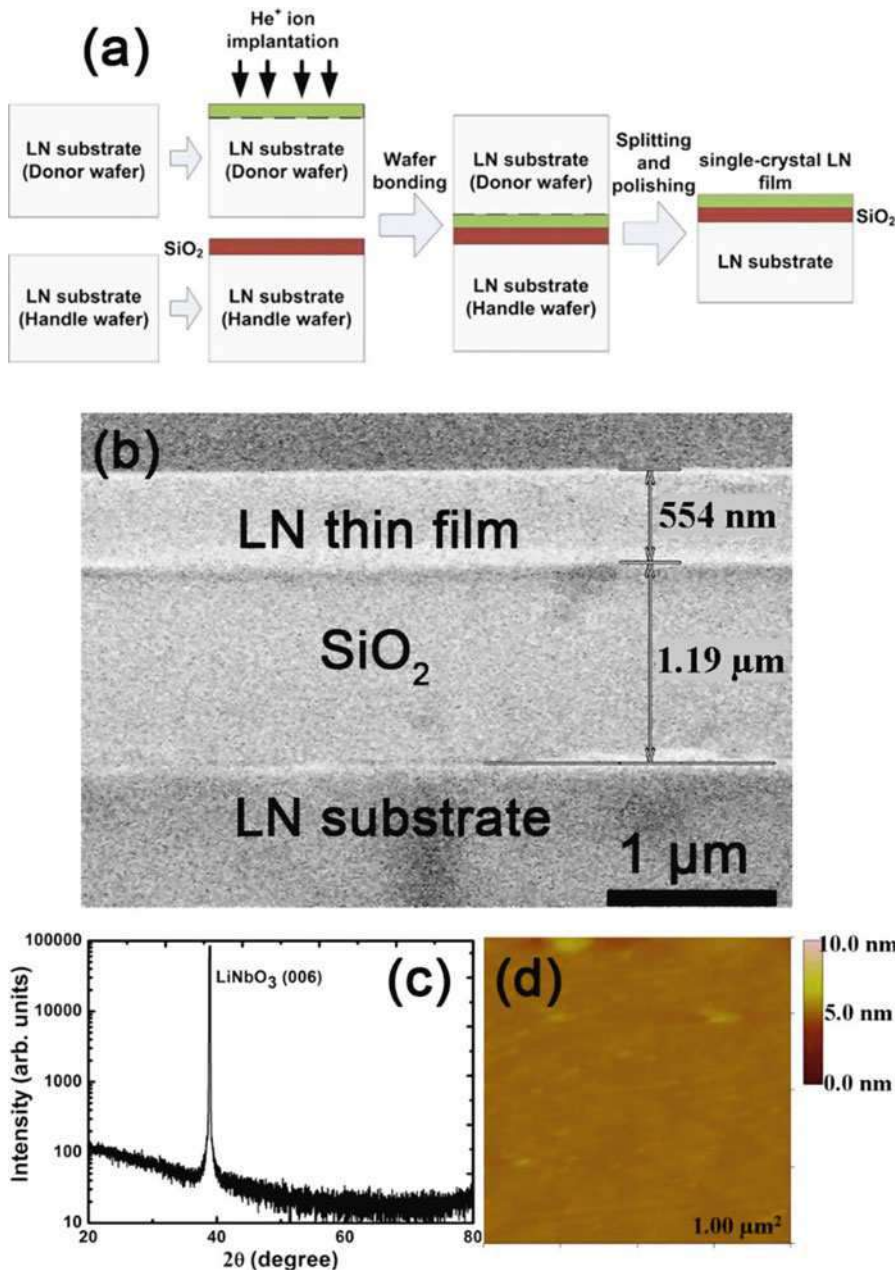
Temperature stability is of great importance for SAW devices especially in telecommunication applications. It is difficult to use conventional single crystal piezoelectric substrate to obtain high-performance SAW devices with required temperature stability. Quartz has extremely high temperature stability, but its coupling coefficient  $K^2$  is very small which precludes its use in many fields. Single crystal piezoelectric substrates LiNbO<sub>3</sub> (LN) and LiTaO<sub>3</sub> (LT) have the opposite characteristics, that is, a very high  $K^2$  but very poor temperature stability. Temperature coefficient of frequency (TCF) is used to represent the frequency shift versus temperature change. LN and LT both have a negative TCF, and therefore, one strategy to enhance temperature stability is integrating materials with a positive TCF with these piezoelectric substrates. With special design, the TCF of the combined structure can even reach zero. SiO<sub>2</sub> is the most commonly used compensation material because of its easy accessibility and high quality. The TCF of the SiO<sub>2</sub>/LN-based SAW device can be improved from  $-80$  to  $-30$  ppm/°C (Nakamura et al. 2008). Several other novel structures for temperature compensation are presented here to show not only high temperature stability but also other advantages (Tian et al. 2016a; Zhou et al. 2013c).

### Temperature-Compensated Structure Z-Cut LN/SiO<sub>2</sub>/LN

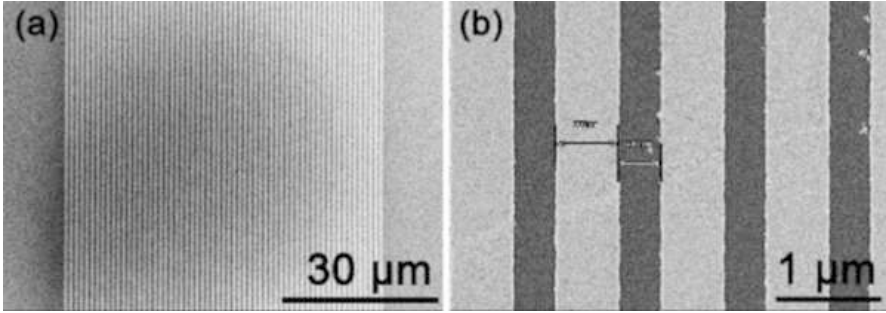
A novel single-crystal Z-cut LN thin film/SiO<sub>2</sub>/LN layered structure for temperature-compensated SAW devices is demonstrated. In comparison with the conventional SiO<sub>2</sub>/LN structure, SiO<sub>2</sub> is used as an interlayer here. The metal IDTs are directly formed on the top of the thin LN film thus precluding deterioration caused by uneven SiO<sub>2</sub> profile. Moreover, because SiO<sub>2</sub> has a near-zero coefficient of thermal expansion (CTE), the thermal expansion of the top LN layer is also suppressed by the SiO<sub>2</sub> interlayer, which results in a smaller TCF.

### Device Fabrication

Figure 9a shows the process flow to achieve the LN/SiO<sub>2</sub>/LN layered structure. Two LN substrates named donor and handle wafer are used. The single-crystal LN thin film is transferred from the donor wafer to the handle wafer through ion implantation, wafer bonding, and splitting. By adjusting the depth of the ion implantation, different thicknesses of the top LN film can be obtained. Final splitting is adopted to obtain a smooth surface which is necessary for the following SAW device fabrication. The top LN film and the SiO<sub>2</sub> interlayer are 0.55 and 1.2  $\mu\text{m}$ , respectively. Figure 9c presents the XRD measurements of the LN layer. The result shows a single



**Fig. 9** (a) Fabrication process flow. (b) Cross-sectional SEM image of LN/SiO<sub>2</sub>/LN layered structure. (c) XRD patterns of LN/SiO<sub>2</sub>/LN structure, where only a strong (006) peak is observed. (d) AFM image of the top LN thin-film surface, with root mean square surface roughness of 0.28 nm over measured area of 1 μm<sup>2</sup> (Reprinted with permission of IEEE)



**Fig. 10** SEM images of the SAW device showing (a) 30 pairs of interdigital electrodes. (b) Close-up view of two electrode pairs (Reprinted with permission of IEEE)

strong (006) orientation with a full-width at half-maximum of  $0.0468^\circ$  which confirms its single-crystal property. Figure 9d shows that the root square surface roughness of the top LN is about 0.28 nm. SAW device is fabricated on the obtained structure using e-beam lithography, Au/Ti (30/5 nm) evaporation, and lift-off process. The nominal line width and pitch of the IDTs are both set as  $0.5 \mu\text{m}$ , which gives a wavelength of  $2 \mu\text{m}$ . Figure 10 shows the SEM image of a fabricated SAW device.

### Results and Discussion

The fabricated SAW device is characterized utilizing a network analyzer with a temperature-controlled microwave probe station. Figure 11 shows the  $S_{11}$  curve with a temperature range from  $20^\circ\text{C}$  to  $80^\circ\text{C}$ . Table 3 summarizes the parameters of the SAW device, including TCF, wave velocity  $V_p$ , and  $K^2$ . Those parameters can be derived from  $S_{11}$  results easily. The proposed structure exhibits much more improved temperature stability than single LN. For three wave modes, the TCF are  $-6.79$ ,  $-16.22$ , and  $-2.34 \text{ ppm}/^\circ\text{C}$ , respectively.

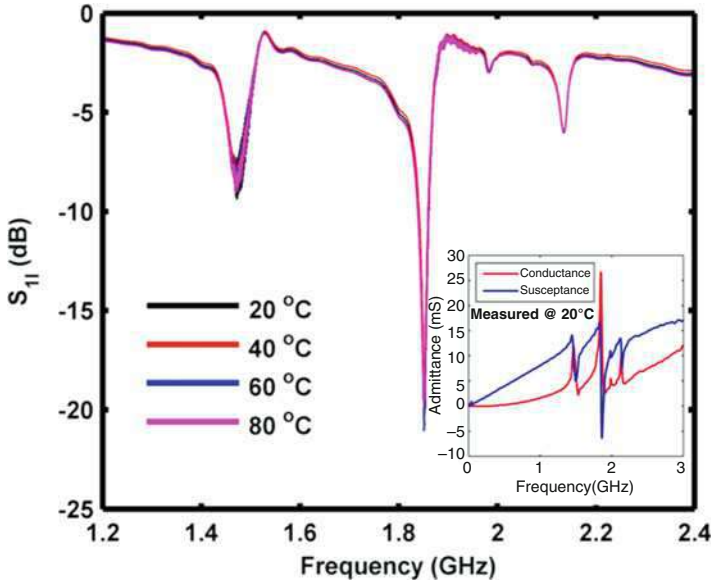
### Temperature-Compensated Structure LN/SiO<sub>2</sub>/Si

#### Device Fabrication

A novel X-cut LN/SiO<sub>2</sub>/Si for temperature stable SAW device fabrication is prepared at first. SiO<sub>2</sub> functions as compensation layer and Si functions as support substrate with low CTE. Similar with LN/SiO<sub>2</sub>/LN structure layer in the former part, the process flow is also mainly based on ion implantation and wafer bonding. Figure 12 gives a brief chart of the process. Thickness of LN, SiO<sub>2</sub>, and Si are 718 nm, 2054 nm, and  $400 \mu\text{m}$ , respectively. Electron beam lithography, metal evaporation, and lift-off processes are adopted to fabricate SAW device on the new structure. Figure 13 shows the optical images of a fabricated SAW device.

#### Results and Discussion

Characterization of the SAW devices made on these two kinds of substrates is performed using a network analyzer (Agilent N5247A) with a temperature



**Fig. 11**  $S_{11}$  parameters of the SAW device measured from 20 °C to 80 °C. Inset: the admittance (conductance and susceptance) spectrum measured at 20 °C for  $K^2$  extraction (Reprinted with permission of IEEE)

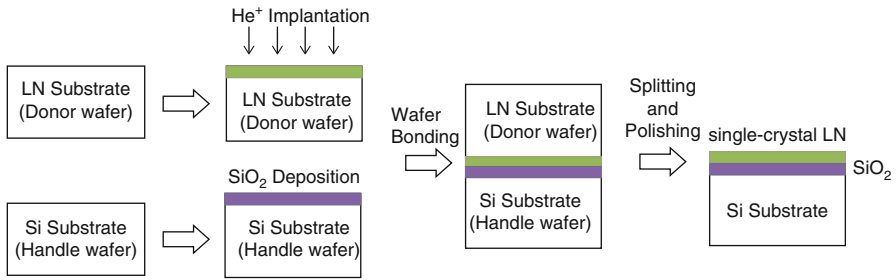
**Table 3** Experimental results of the SAW device on LN/SiO<sub>2</sub>/LN structure

$f_c$ (GHz)	TCF(ppm/°C)	$V_p$ (m/s)	$K^2$ (%)
1.4720	-6.79	2944.0	2.43
1.8527	-16.22	3705.4	4.89
2.1356	-2.34	4271.2	1.08

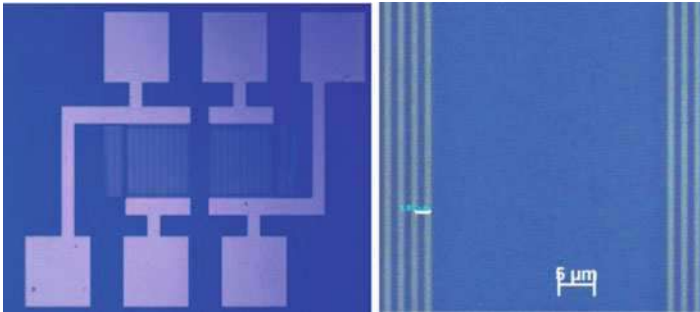
adjustable probe station. The  $S_{11}$  curve at several different temperature points of the fabricated device is shown in Fig. 14. It is presented in figure that the  $f$ - $T$  relation is almost linear for this structure. The calculated TCF value is near  $-18$  ppm/°C, which is largely improved compared with  $-80$  ppm/°C of single LN. The TCF of the LN/SiO<sub>2</sub>/Si structure is not only at a rather low value to show excellent temperature stability, but also provides a perfect constant performance which offers advantage for both communication use and sensing applications.

### Super High Frequency SAW Devices

For SAW devices, both for communication and sensing applications, frequency characteristic is the most significant property. High frequency offers improved sensitivity for sensors and fast data transfer rates for signal transmission. Resonance



**Fig. 12** Fabrication process flow of LN/SiO<sub>2</sub>/Si (Reprinted with permission of IEEE)



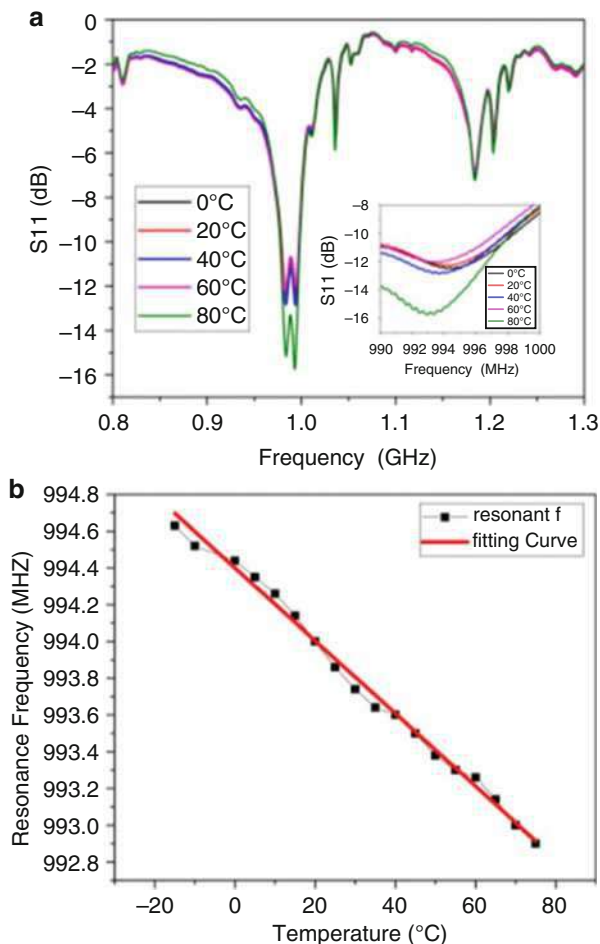
**Fig. 13** Optical image of SAW device. (a) Overall SAW device pattern. (b) Close-up view of IDT pairs (Reprinted with permission of IEEE)

frequency of SAW device is related to the substrate and the IDT period  $\lambda$ . Once the choice of the substrate material is made by using whether a single crystal substrate or a multilayered structure, the most effective method to further increase working frequency is to decrease the  $\lambda$  by reducing the IDT line width. A recent work has reported using LiNbO<sub>3</sub> to get 8.6 GHz frequency but the insertion loss (IL) is large to  $-60$  dB (Chen et al. 2011). A systematic study fabricating LiNbO<sub>3</sub> SAW device to get the highest frequency exceeding 14 GHz with significantly lower losses is reported here (Chen et al. 2015).

### Device Fabrication

The complete fabrication process overflow is shown in Fig. 15. 128 °YX LiNbO<sub>3</sub> substrate is diced into  $1 \times 1.5$  cm<sup>2</sup> chips. Piranha (H<sub>2</sub>SO<sub>4</sub>:H<sub>2</sub>O<sub>2</sub> 3:1) and deionized (DI) water are used to clean the surface. After baking at 175 °C for 5 min, PMMA 950 k A2 and A4 resist is spun on the substrate, yielding near 100 nm and 200 nm thick layers. Because LiNbO<sub>3</sub> is an insulator, to avoid charge accumulation, a 20 nm conductive layer of ESPACER is coated on top of the PMMA resist. Then the pre-designed SAW patterns are exposed at 10KeV using a Raith 150 electron beam lithography (EBL) tool. After exposure, the ESPACER is removed by immersion in DI water and development is conducted in MIBK:IPA. The chosen electrode

**Fig. 14** (a)  $S_{11}$  parameters of SAW device on LN/SiO<sub>2</sub>/Si structure measured from 0 °C to 80 °C. Inset:  $S_{11}$  curves near resonance frequency. (b) The extracted relation of resonance frequency and temperature for SAW device on LN/SiO<sub>2</sub>/Si structure (Reprinted with permission of IEEE)



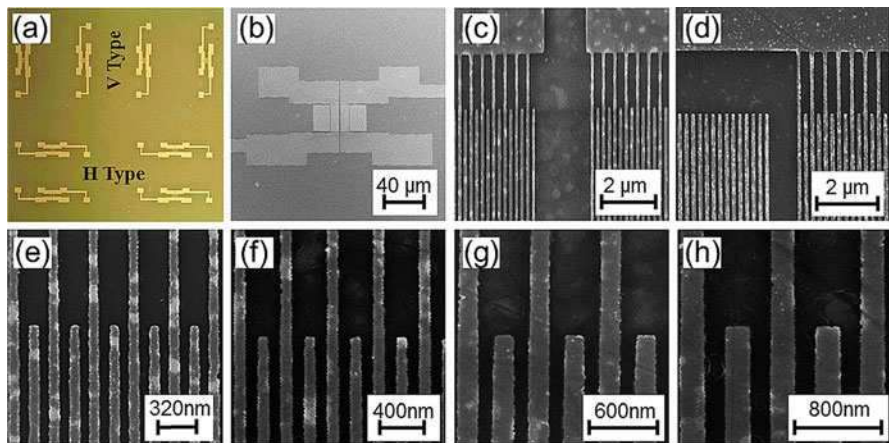
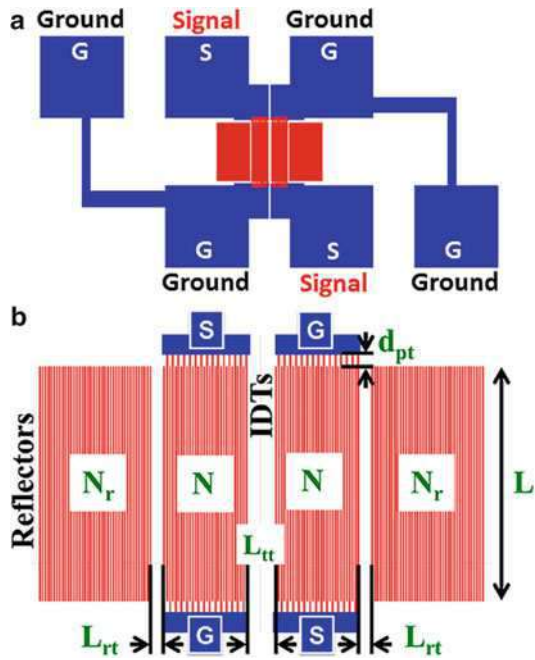
material is Au. The metallization is carried out by electron beam evaporation. Ti acting as a good seed layer is coated first followed by Au metallization. For different SAW periods, corresponding metal thicknesses are adopted. Finally, lift-off is conducted in ultrasonic acetone. Figure 15 shows the schematic layout of the super high frequency SAW devices.

## Results and Discussion

The SEM images of fabricated SAW devices are shown in Fig. 16. Then the SAW devices are characterized using Agilent N5230C network analyzer equipped with a probe station. Useful parameters such as resonance frequency, insertion loss, Q-factor, bandwidth, and coupling coefficient can be derived from the experimental data. And it can be seen from Fig. 17 that 320 nm device can achieve frequencies of

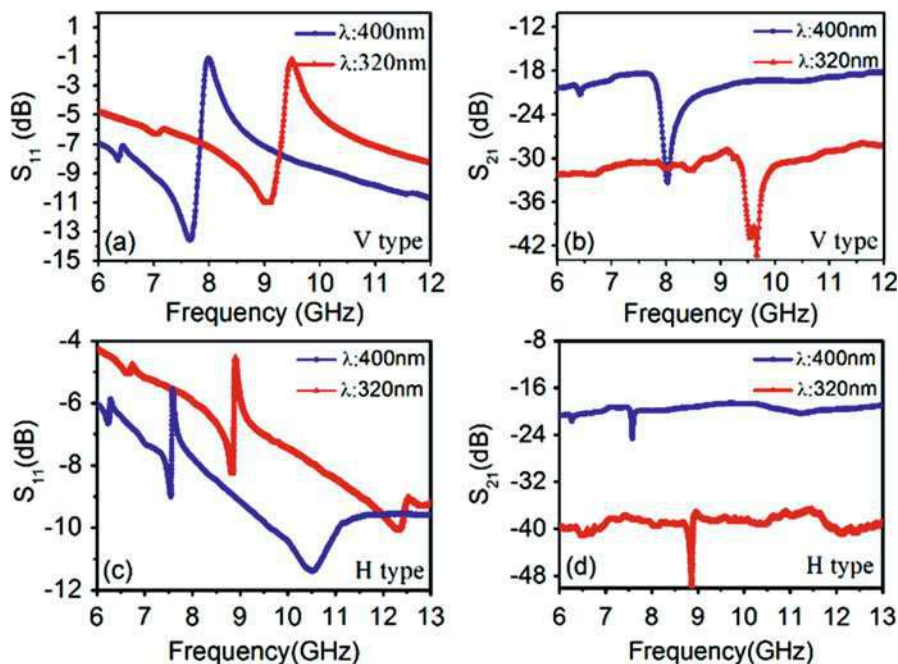


**Fig. 15** Two-port (a) SAW device design with GSG electrode configuration and (b) the magnified IDTs and reflectors structure with relevant labels (Reprinted with permission of AIP)



**Fig. 16** Optical microscope images of (a) fabricated H-type and V-type SAW devices. SEM micrographs of 400 nm period SAW devices showing (b) central region, (c) magnified region between two IDTs, and (d) magnified region between an IDT and a reflector. High-resolution SEM micrographs of (e) 320 nm, (f) 400 nm, (g) 600 nm, and (h) 800 nm period IDTs. The SEM voltage and working distance are 10 keV and 5 mm, respectively. The bright patterns in panels (c–f) are possibly charging artifacts caused due to defects (notching) in the thin EPACER layer spun-cast before SEM imaging (Reprinted with permission of AIP)





**Fig. 17** Comparing frequency characteristics of 320 versus 400 nm SAW devices. Reflection ( $S_{11}$ ) and transmission ( $S_{21}$ ) parameters, respectively, for (a and b) V-type, and (c and d) H-type devices. Each panel contains spectra for both 320 nm (triangles) and 400 nm (circles) period devices (Reprinted with permission of AIP)

12.5. Compared with previous works, this work shows the highest frequency and best insertion loss for LiNbO<sub>3</sub> SAW devices.

## Flexible Pseudo-SAW Devices

Flexible electronics has many unique qualities such as bendable, lightweight, and low-cost fabrication. Since the advent of flexible silicon solar cells about 40 years ago, a wealth of flexible electronic devices has been demonstrated by both industry community and academic institutions in recent years (Khang et al. 2006; Oh et al. 2016). Various methods have been used to implement active components like diodes and transistors by using organic materials and transfer printed semiconductor thin films (Kim et al. 2016). Meanwhile, there are many reports on flexible passive devices like conductors, inductors, and antennas, which are mainly constructed by integrating conductive materials with flexible elastomers (Jung et al. 2016; Ahmed et al. 2015). However, it is still to be found an effective way to realize a flexible RF resonator which is a must for the implementation of flexible wireless electronics. We explore the possibilities by directly integrating piezoelectric ZnO thin film with

flexible spin-coated PI substrate to implement the flexible acoustic wave resonators (Zhou et al. 2012).

PI has good biocompatibility and high chemical resistance, and it can withstand a high temperature up to 250 °C. These properties are important for succeeding fabrication processes like sputtering and chemical etching, etc. ZnO thin films can be readily deposited by using reactive sputtering method at room temperature, ensuring the feasibility of the PI-based fabrication process.

### Device Structure and Fabrication

The fabrication process flow and the resultant device structure are illustrated in Fig. 18. The PI film with about 10 μm thickness was spin-coated on the handle wafer with PECVD SiO<sub>2</sub> sacrificial layer (300 nm) and the piezoelectric ZnO film was subsequently deposited by sputtering method. The sputtering conditions have been optimized to obtain (002) preferred oriented piezoelectric thin film. Two-port resonators with the IDTs in the middle and grating reflectors on the two sides were formed by using lift-off process. The distances between the two IDTs and between the IDT and the grating reflectors were designed to ensure the formation of standing waves. After electrodes patterning, the handle wafer was coated with a thin layer of photoresist and immersed into HF solution to etch off the SiO<sub>2</sub> and release the flexible structure.

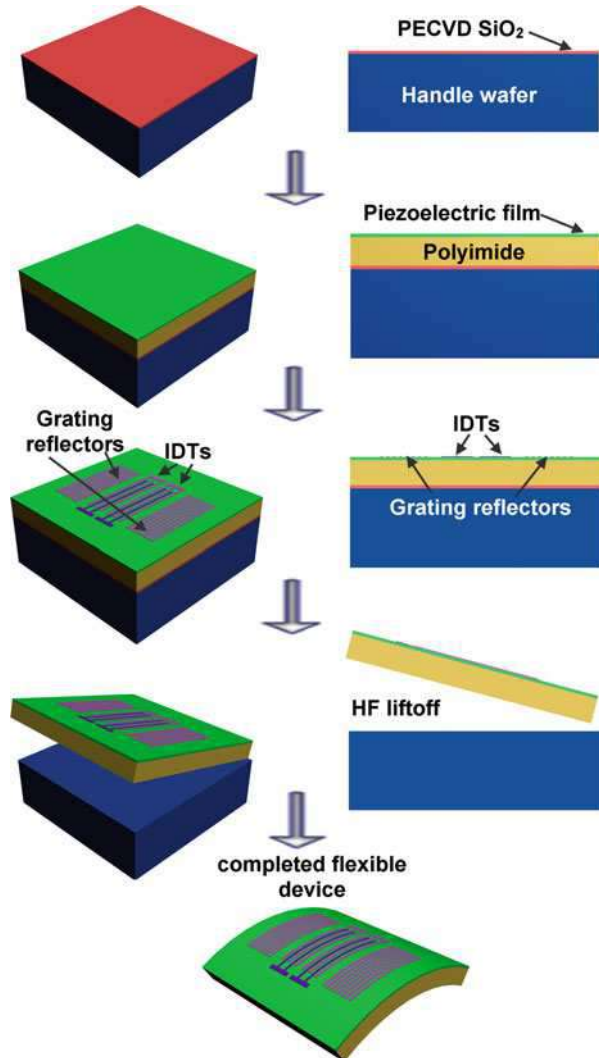
The fabricated device is very flexible and mechanically robust (Fig. 19a). The column crystal structure of the deposited piezoelectric film can be clearly observed from the cross-sectional SEM of the structure (Fig. 19b), suggesting the well aligned crystal structure of the sputtered ZnO film. It is also shown that the PI film exhibits charging effect when imaged with the SEM, which is the typical phenomenon for high resistance samples. Figure 19c shows an optical picture of a typical device on PI substrate before the final peeling off process with 6 μm wavelength. The inset highlights the interdigital electrodes at one end of the transducer, which clearly shows the electrodes with smooth edges.

It is noted that the spin-coated PI thin film has a very smooth surface, with a root mean square of roughness of 0.26 nm. This is very beneficial for obtaining high-quality sputtering piezoelectric films. After the ZnO sputtering, the surface roughness of the ZnO film is 4.90 nm, which is smooth enough for the photolithography and lift-off processes. Also the XRD spectrum of the ZnO/PI structure shows highly (002) preferred orientation, which is beneficial for acoustic wave stimulation and propagation.

### Device Working Principle

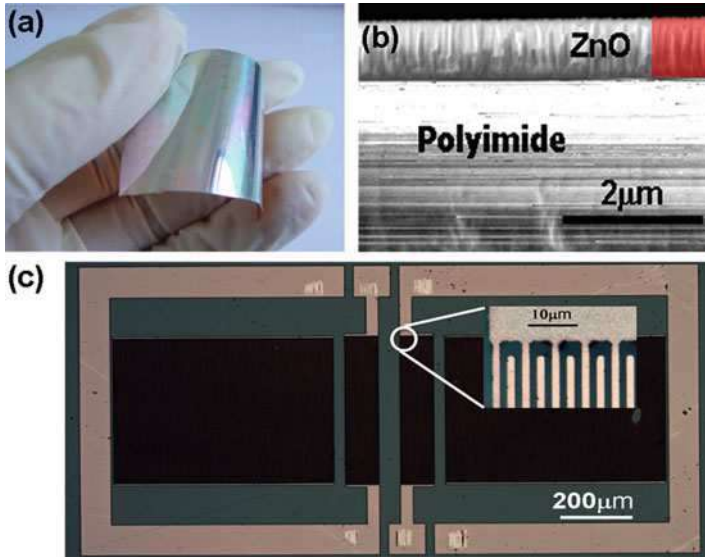
To study the resonant mode shape of the proposed flexible structure, two-dimensional finite element model has been constructed and both harmonic and modal analyses have been performed by using ANSYS software. The model was based on the flexible resonator with 6 μm wavelength fabricated on ZnO/PI structure. The thickness of the ZnO film is 1.6 μm and the thickness of the PI film is 10 μm, which is the same as used in the experiments. One pair of electrodes was modeled and periodic boundary conditions were applied to simulate multiple pairs of electrodes as designed in the IDTs (Fig. 20a). The fixed displacement constraints at the bottom

**Fig. 18** Schematic illustration of the device structure and fabrication process of the resonators. The two-port resonators with two IDTs in the middle and grating reflectors are formed by photolithography and lift-off processes. The handle wafer can be reused after the releasing process, ensuring a low manufacturing cost for the proposed strategy (Reprinted with permission of IEEE)

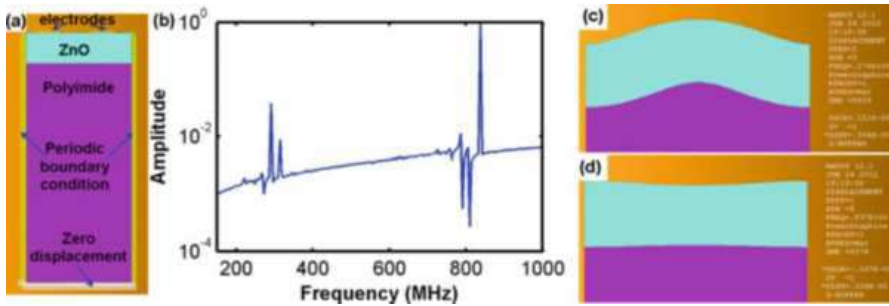


side of the PI is applied considering the very small part of acoustic wave energy reaching to the substrate. It is expected that the acoustic wave is excited by the surface interdigital electrodes, and transported along the surface and also into the PI substrate. Due to the large acoustic impedance mismatch between piezoelectric material and the flexible PI material, most part of the acoustic wave will be reflected at the boundary between ZnO and PI. The effect of adding the flexible PI film to the free-standing ZnO film has thus to be quantitatively investigated with the proposed numerical model.

Firstly, we performed harmonic analysis to get the frequency spectrum of the effective admittance of the resonator. One volt as voltage is applied to one electrode



**Fig. 19** (a) Photograph of the flexible acoustic resonators; Cross-sectional SEM of released (b) ZnO/PI and (c) optical photograph of the fabricated resonator with 6 μm wavelength (Reprinted with permission of IEEE)



**Fig. 20** (a) Two-dimensional finite element model of the ZnO/PI (1.6 μm/10 μm) structure; (b) Harmonic analysis results: effective admittance (normalized to the maximum value) as a function of the applied Ac signal frequency, of which two wave modes can be clearly observed; Modal analysis results: mode shapes of the first wave mode (c) and the second wave mode (d) (Reprinted with permission of IEEE)

with another electrode set to the ground. Figure 20b shows the frequency spectrum of the resonator, in which two wave modes are presented. Their center frequencies are about 274 MHz and 837 MHz, respectively. We then performed modal analysis to study the mechanical vibration characteristics of the two wave modes. The first mode shape is the same as the known Lamb asymmetrical mode  $A_0$ , and the second shape is the same as the symmetrical  $S_0$  mode (Fig. 20c, d). Thus we will denote these modes as pseudo-SAW wave modes hereafter. The results indicate that the

adding of a flexible PI film actually has a little effect in terms of wave mode excitation due to its much lower acoustic impedance. This phenomenon is further explained by theoretical calculation as detailed below.

The theoretical dispersive characteristics of the ZnO/PI are calculated using transfer matrix method (Wang and Rokhlin 2001). The traditional Lamb wave characteristics based on suspended ZnO film is also computed for comparison. The core principle in constructing the transfer matrix is to choose a state vector to include basic physical properties. In this study, the state vector is expressed as:

$$\tau = [T_{13} \quad T_{23} \quad T_{33} \quad D_3 \quad v_1 \quad v_2 \quad v_3 \quad i\omega\phi], \quad (5)$$

where  $T_{i3}$  is the stress component along the  $i$ th direction and the 3rd direction stands for the thickness direction,  $D_3$  is the electric displacement,  $v_i$  is the  $i$ th velocity component,  $\phi$  is the electrical potential, and  $\omega$  is the angular frequency. The state vectors at each surface are connect by a linear differential equation, which is constructed by combining the Newton's second law of motion and the constitutive relations of piezoelectric material. With this method, the continuous conditions are automatically satisfied and only boundary conditions are to be considered to compute the dispersive characteristics for a certain wave mode.

In the implementation of the calculation, both the open-circuit acoustic velocity  $v_o$  and short-circuit acoustic velocity  $v_s$  are calculated with a given  $kh_{\text{ZnO}}$ , where  $k = 2\pi/\lambda$  is the wavenumber, and  $h_{\text{ZnO}}$  is the thickness of the ZnO film. The coupling coefficient is thus given by

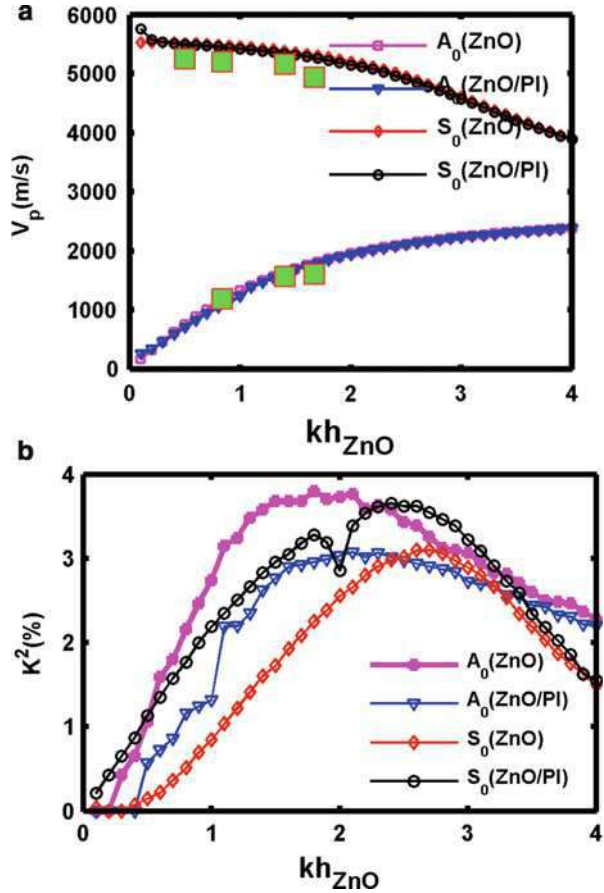
$$K^2 = 2 \frac{v_o - v_s}{v_o}. \quad (6)$$

Figure 21a shows the calculated results of the dispersive characteristics of ZnO/PI and ZnO structures. Two wave modes are considered in the calculations, namely,  $A_0$  and  $S_0$  wave modes. In general, the structures with PI bounded to the lower surface of the piezoelectric film exhibit a similar phase velocity dispersive characteristics. The effect of adding a PI layer to the ZnO film only results in a slightly lower phase velocities. In terms of coupling coefficients,  $A_0$  mode exhibit lower  $K^2$  for the structures with PI, while  $S_0$  mode exhibit the opposite effect. Although the Lamb wave is usually excited in the suspended piezoelectric film, it is found that the PI substrate bounded piezoelectric film can be also used to excite Lamb wave modes. This finding is in accordance with the previous FEM simulation.

### Electrical Characterization

Electrical characteristics of the flexible resonators have been measured by using vector network analyzer HP8722ES and microprobes. Figure 22 shows the measured results of the resonator with 6  $\mu\text{m}$  wavelength based on ZnO/PI structure. Two wave modes are clearly observed, and their center frequencies are about 266.6 MHz and 822.2 MHz, respectively. The results are in good agreement with the FEM results. The measurement results of resonators with 7.128  $\mu\text{m}$  and 12  $\mu\text{m}$  wavelength also present

**Fig. 21** Dispersion curves of (a) acoustic wave phase velocities and (b) coupling coefficients based on ZnO suspended film and ZnO/PI structure for Lamb wave  $A_0$  and  $S_0$  modes. The experimental results are also marked in this figure (Reprinted with permission of IEEE)



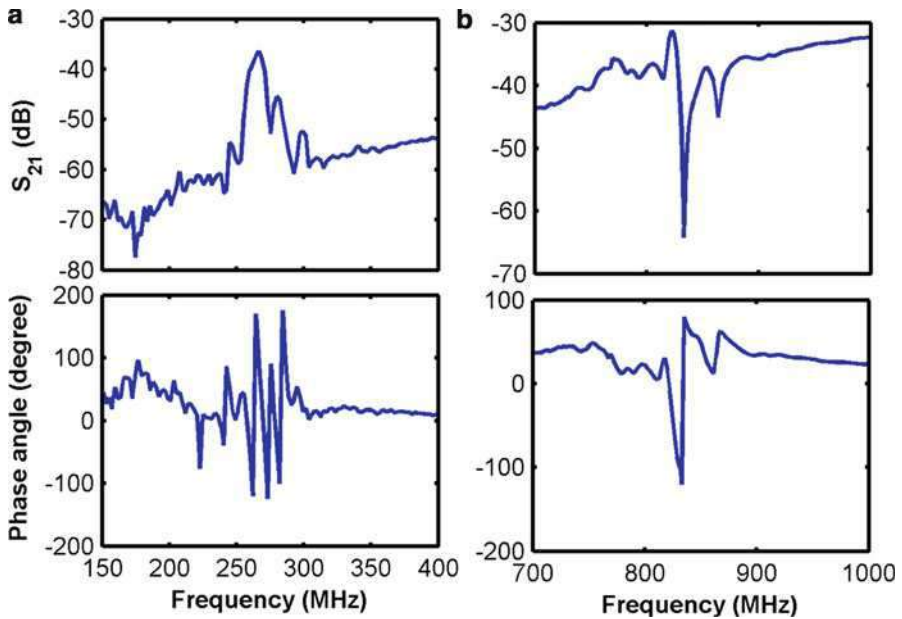
two wave modes, while only one wave mode, i.e.,  $S_0$  mode, is obtained for the resonator with  $20 \mu\text{m}$  wavelength. This is mainly due to the near-zero coupling coefficient for a very small  $kh_{ZnO}$  value. The experimental results fit well with the theoretical ones. It is also noticed that a higher  $Q$  factor can be obtained under the  $S_0$  mode, of which a  $Q$  factor up to 1000 has been achieved. To examine the mechanical robustness of the flexible resonator, measurement after the bending operation is conducted for the resonator with  $6 \mu\text{m}$  wavelength, and it is found that the resonator exhibits almost the same characteristics as the initial status except for a slight frequency shift.

## Sensors Based on SAW Devices

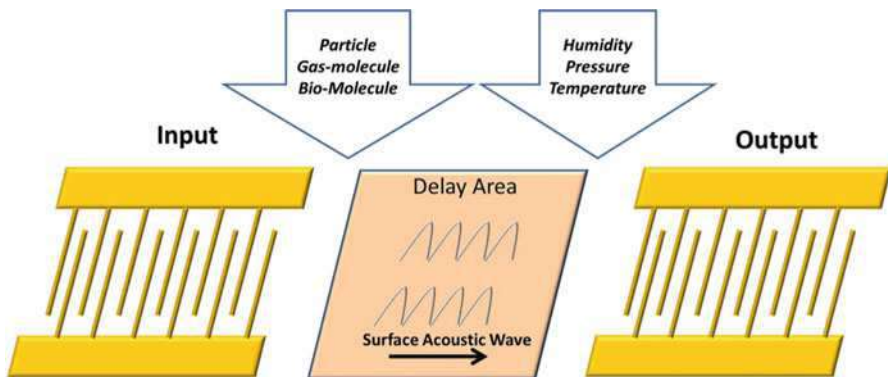
### Temperature Sensor Based on SAW Device

SAW devices have been widely used in sensing field because of their relatively cheap price, easy fabrication process, tiny size, and compatibility with standard





**Fig. 22** Transmission responses of the resonator with  $6\mu\text{m}$  wavelength based on ZnO/PI structure: (a) and (b) present results of  $A_0$  and  $S_0$  modes. The center frequencies are about 266.6 MHz and 822.2 MHz, respectively (Reprinted with permission of IEEE)



**Fig. 23** Diagram of SAW physical quantity sensor (Reprinted with permission of AIP)

CMOS process. The basic working principle of SAW sensors is ambient conditions such as temperature, pressure, gas concentration, and mass of matter could cause resonance frequency  $f_r$  shift (Bo et al. 2016; Tian et al. 2016b). Figure 23 is a diagram which shows SAW device works as sensor for multiple physical quantities.

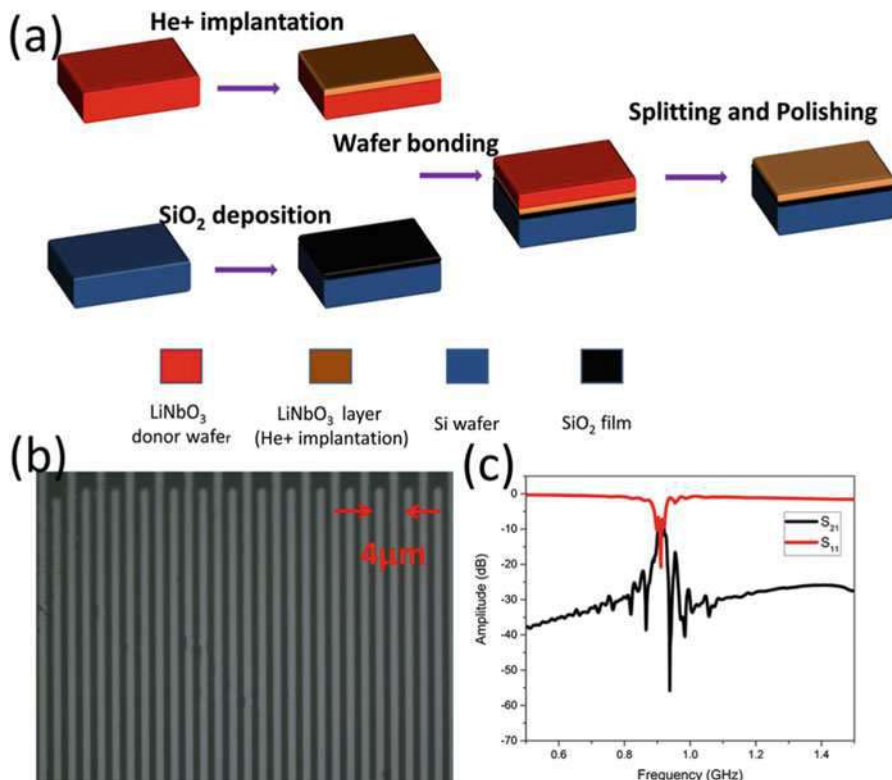
For temperature detection, temperature coefficient of frequency (TCF) is the core parameter to represent the shift of  $f_r$  with temperature  $T$ . Some most commonly used piezoelectric materials like  $\text{LiNbO}_3$  and  $\text{LiTaO}_3$  have a negative TCF which means  $f_r$  of SAW devices made on them decreases with temperature increment. Usually, TCF is not a constant for a given material. It also varies when temperature changes due to the higher order effect. One widely used method for temperature compensation is to integrate a positive TCF material with the piezoelectric substrate, where  $\text{SiO}_2$  is the most common choice.

For sensing applications, frequency readout is one of the most important steps. Traditional method utilizing network analyzer is not suitable as the bulky size and operation difficulty. Specialized  $f_r$  readout chip is a promising way to achieve portable sensing. A high-resolution and high-linearity SAW temperature sensor is achieved, which consists of a SAW resonator device fabricated on  $\text{LiNbO}_3/\text{SiO}_2/\text{Si}$  piezoelectric substrate and a  $f_r$  readout chip using standard 180 nm CMOS technology (Tian et al. 2016b). The obtained relation between frequency and temperature has a pretty good linearity. The frequency resolution of the sensor is 733 Hz and the corresponding temperature accuracy of 0.016 °C is derived.

Figure 24a shows the fabrication process of the  $\text{LiNbO}_3/\text{SiO}_2/\text{Si}$  substrate. The first step is implanting He<sup>+</sup> ions into a  $\text{LiNbO}_3$  wafer to get a thin film layer. Then a  $\text{SiO}_2$  layer is deposited on a Si wafer. The two prepared wafers are bonded together followed by bonded pair heating. The thin  $\text{LiNbO}_3$  layer is split off from the original  $\text{LiNbO}_3$  wafer but stays on the  $\text{SiO}_2/\text{Si}$  surface. An annealing step is adopted to increase the bonding strength. Finally, a chemical-mechanical-polishing (CMP) process is conducted to smooth the surface of the new obtained structure. SAW device is designed with a resonator structure. The linewidth of IDTs is set to 1  $\mu\text{m}$  corresponding to a period of 4  $\mu\text{m}$ . A cleaning-EBL patterning-metal evaporation-lift off process is performed to make resonator structure. Figure 24b shows the optical image of the fabricated IDT structure. As in the traditional way, network analyzer (Agilent N5247A) is used to characterize the device. Figure 24c presents the  $S_{11}$  and  $S_{21}$  curves. Resonance frequency is about 902 MHz.

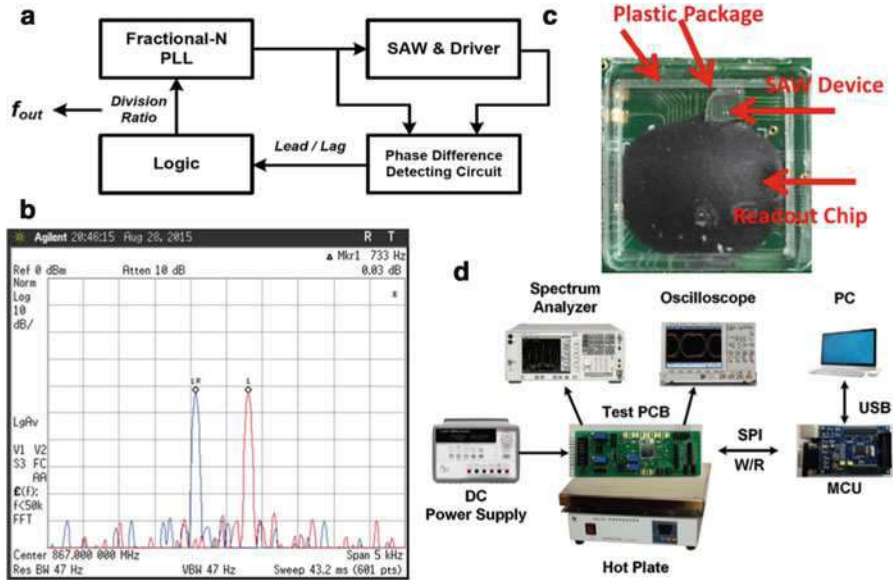
As mentioned before, as network analyzer has bulky size and operation difficulty, it is near impossible to use this method in sensing applications. A frequency readout chip is designed and fabricated using 180 nm CMOS technology to match the SAW device. The chip consists of a fractional-N phase-locked loop (PLL) as the stimulus source and a frequency detecting circuit using successive-approximation (SAR). Figure 25a shows the diagram of the proposed circuit. Figure 25b shows the measured result of the PLL, from which we can obtain 733 Hz frequency resolution. After the SAW device and the chip are both prepared and tested, we combine them together to build a portable sensor system. Those two parts are bonded to a PCB board which functions as connection and assistant discrete electronic components supplier. Figure 25c shows the picture where plastic package and black adhesive are used for protecting the SAW device and the chip. The system in Fig. 25d shows equal test results compared to the traditional method. The much scaled size and simplified operation offer great potential to the use of SAW device in real-time sensing applications.



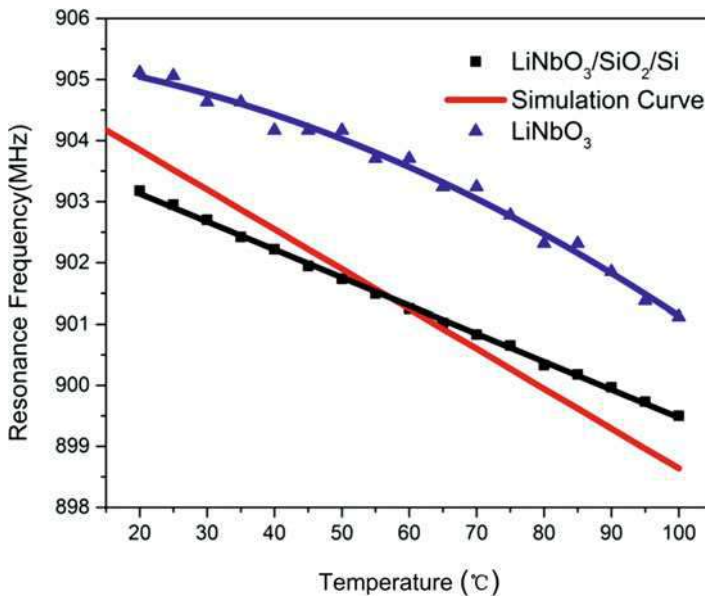


**Fig. 24** (a) Fabrication process overflow of  $\text{LiNbO}_3/\text{SiO}_2/\text{Si}$ ; (b) Optical image of a fabricated SAW resonator device; (c)  $S_{11}$  and  $S_{21}$  of a fabricated SAW device (Reprinted with permission of AIP)

To verify the ability of the sensor to detect temperature, the PCB holding the SAW device and the readout chip is placed on a hot plate to achieve temperature variation. The test range of temperature is set from  $20^\circ\text{C}$  to  $100^\circ\text{C}$  with a gradient of  $5^\circ\text{C}$ . Figure 26 shows the obtained relation of resonance frequency and temperature and simulation result. In the given temperature range, almost linear relation can be achieved. When linear fitting is applied to the test data, a 0.99842 R-square can be achieved. The test is repeated for several cycles and the result was always stable. A comparison test is set using the same arrangement but the substrate is changed to normal  $\text{LiNbO}_3$  substrate. The result is also shown in Fig. 26. It is easily inferred that linearity of the temperature performance of SAW device based on  $\text{LiNbO}_3/\text{SiO}_2/\text{Si}$  substrate is much optimized than the device based on  $\text{LiNbO}_3$  substrate. Although the frequency variation is smaller due to the compensation of  $\text{SiO}_2$  for  $\text{LiNbO}_3/\text{SiO}_2/\text{Si}$  structure, the linearity is much optimized which means it is more suitable in practical sensing field.



**Fig. 25** (a) Diagram of the proposed circuit; (b) Measured frequency resolution of the fractional-N PLL; (c) Picture of the packaged SAW device and readout chip; (d) Test set-up (Reprinted with permission of AIP)



**Fig. 26** Experimental and simulation resonance frequency-temperature relation of SAW device on  $\text{LiNbO}_3/\text{SiO}_2/\text{Si}$  substrate, experimental result of SAW device on  $\text{LiNbO}_3$  substrate as a comparison (Reprinted with permission of AIP)

## DNA Biosensor Based on SAW Device

In this section, we introduce our recent work in using SAW devices for DNA detection based on the mass loading effect of the SAW devices (Cai et al. 2014, 2015).

### Fabrication of SAW Sensors

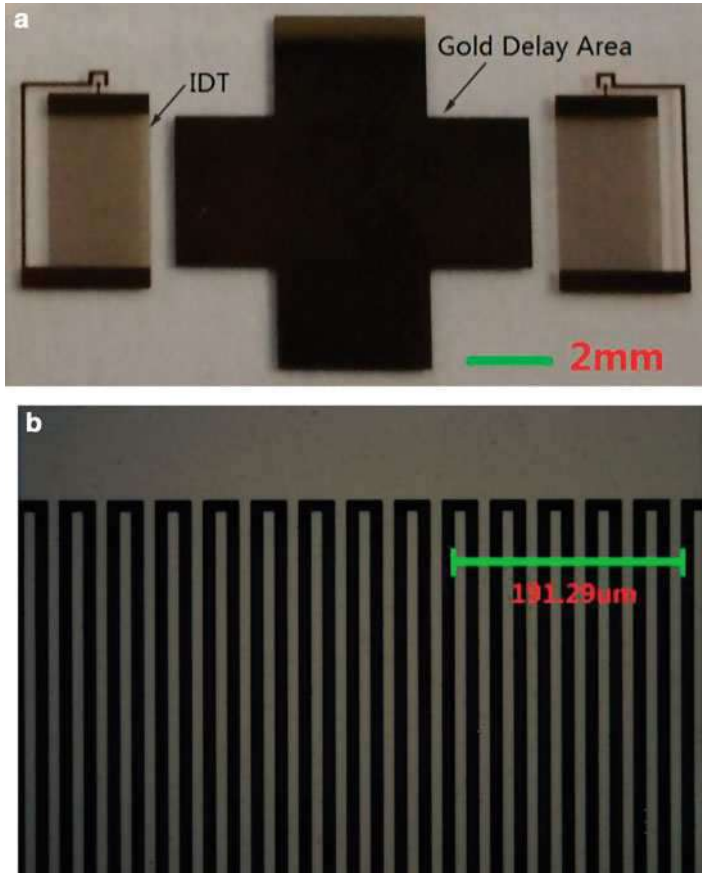
The SAW devices used in this study has a two-port delay line structure, and the distance between the input IDT and output IDT is about 1 mm. The linewidth of the IDT is 0.5  $\mu\text{m}$ , which gives an equivalent wavelength of 2  $\mu\text{m}$ . Conventional single crystal  $\text{LiNbO}_3$  substrate is used as the piezoelectric substrate, and the electrodes are formed by electron-beam lithography, gold sputtering, and lift-off processes. It is noted that the space between the two IDTs is also coated with a thin layer of gold at the same time. This gold-coating layer between the two IDTs serves as the perturbation area when the DNA to be measured is affiliated. Figure 27 shows the photographs of a fabricated device.

To immobilize the DNA on the gold delay area, the probe-target mechanism is used following the procedure as described below:

1. The device is cleaned in an ultrasonic bath using ethanol and deionized water for 5 min, and dried with  $\text{N}_2$  gas flow.
2. The probe DNA is first bonded to the gold delay area. Firstly, the gold surface is activated and the probe DNA with the sequence of 3'-SH-AAAAAAAAA CCCCCTAAAA-GGTCCTAAAAACCTAAAAACCT-5' is coated onto the surface of gold delay area. The probe DNA solution is mixed with 0.1 M  $\text{KNO}_3$  to enhance the hybridization. Then the SAW device is kept in a chamber at 4  $^\circ\text{C}$  for 24 h to promote the immobilization of the probe DNA on the delay area. After that, the device is cleaned using Milli-Q water and dried with  $\text{N}_2$ .
3. The field area without probe DNA is blocked with bull serum albumin solution. Further hybridization in a solution of mixed phosphate buffered saline and target DNA is conducted for 4 h, followed by cleaning process using PBS to remove any unhybridized DNA.

### Results and Discussions

The SAW device is mounted on a printed circuit board and connected to a network analyzer to measure the real-time frequency response. Figure 28 shows the measured  $S_{11}$  parameters of the original SAW devices and that after probe and target DNA hybridizations. It is well known that mass-loading effect results in the downshift of the resonance frequency of SAW devices. This effect is clearly revealed by this experiment. Along with the downshifts of the resonance frequency, the quality factors of the SAW devices after DNA coating also decrease a lot, which is probably caused by the enhanced acoustic wave absorption and diffraction due to the DNA coating on the gold delay area. The dynamic range of the biosensor based on SAW device is investigated by studying the frequency response under various concentrations of the target DNA. Figure 29 shows the SAW device exhibits a linear relationship between the frequency downshift and the concentration of the target



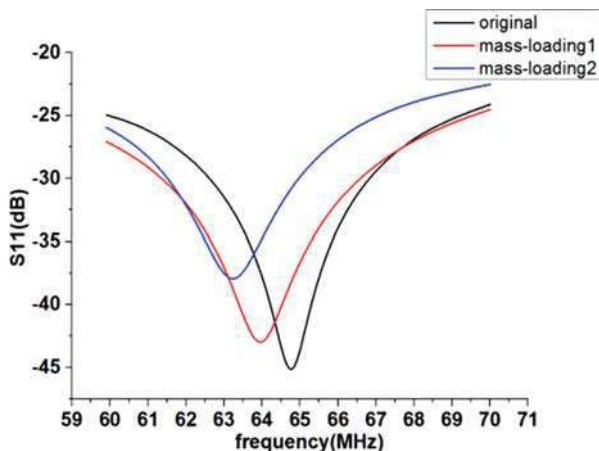
**Fig. 27** Optical image of a SAW device. (a) A complete SAW sensor showing the two IDTs and the gold delay area (Reprinted with permission of Elsevier)

DNA in a range of 1.3 nM ~ 130 nM. The resonance frequency of the SAW device is affected by many ambient parameters like temperature and humidity, which can be clumped into a single term as noise-equivalent frequency shift. When the concentration is too low, the mass-loading effect cannot be distinguished from the noise term, which also sets a lower detection limit of the SAW biosensor.

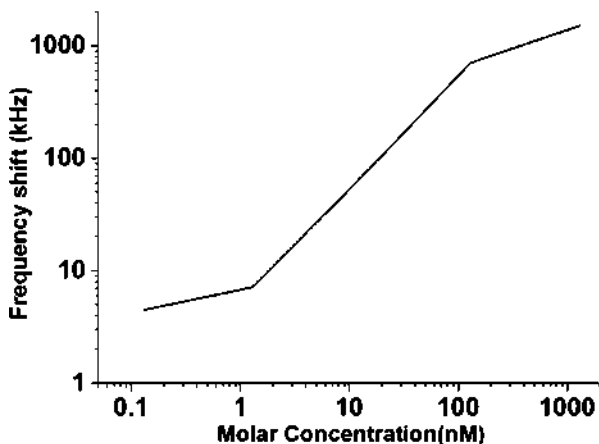
## High-Performance Film Bulk Acoustic Wave Resonators

Film bulk acoustic wave resonator (FBAR) was successfully commercialized by Ruby in Agilent (now with Avago) (Ruby et al. 1999). Compared with the competing SAW technology, FBAR provides better performance in terms of power handling capability and quality factors, and was first implemented in the duplexers at

**Fig. 28** Measured  $S_{11}$  parameters of the original SAW devices, the SAW devices after probe DNA hybridization (mass-loading1), and target DNA hybridization (mass-loading2) (Reprinted with permission of Elsevier)



**Fig. 29** Relationship between the frequency downshift and the concentration of the target DNA (Reprinted with permission of Elsevier)



1900 MHz frequency band. Until now, the FBAR is mainly used as a standalone device although the fabrication process of AlN-based FBAR is potentially CMOS compatible. Further advancement towards integrative RF front-end module is expected to build a more compact portable wireless handset. On the other hand, it is promising to exploit the high-frequency FBAR for highly sensitive sensing applications.

## Basic Structures and Fabrication Process

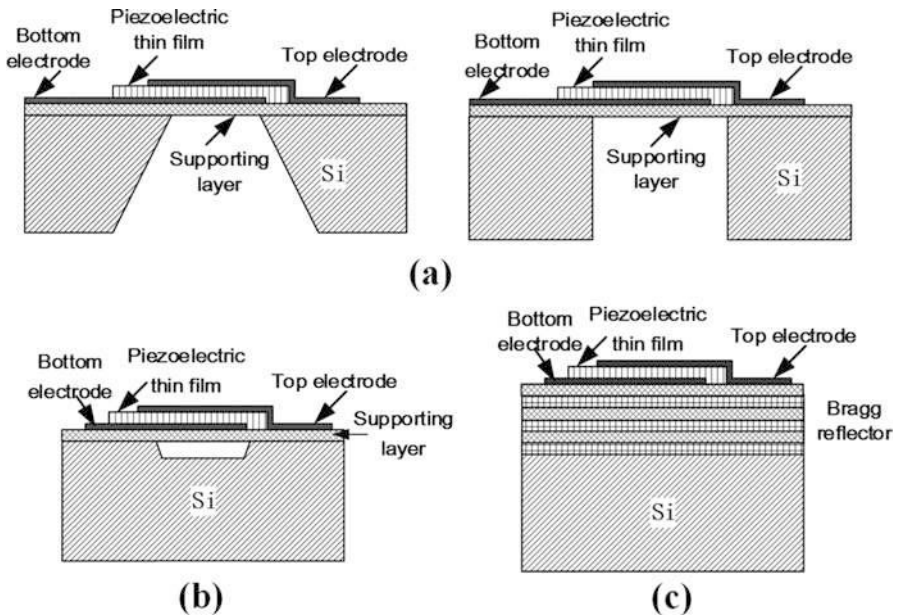
The resonance frequency of a FBAR is given by

$$f_c = V_p/\lambda, \quad (7)$$

where  $\lambda = 2d$  and  $d$  is the thickness of the piezoelectric thin film sandwiched between top and bottom electrodes. As the frequency is determined by thin film thickness, it is relatively easier to achieve high-frequency resonator using FBAR technology than SAW technology. To obtain a high quality factor, the acoustic wave energy should be confined in the piezoelectric thin film. And there are several kinds of device structures that can satisfy this requirement.

**Basic Structures**

Figure 30a shows the FBAR structure obtained with bulk micromachining method. The structure is suspended by removing the silicon at the backside. The left figure of Fig. 30a shows the structure formed by wet anisotropic etching. With this structure, the acoustic wave can be reflected totally at the upper and lower boundary and the wave energy is confined in the FBAR ensuring a high quality factor. As a mature process in fabricating micro-electro-mechanical system (MEMS), this structure is adopted in the early development of FBARs. One can notice that the sidewall forms an angle of  $54.7^\circ$  with the silicon planar surface, which results in a larger area occupied than the effective FBAR area. Take a FBAR fabricated on a  $400\ \mu\text{m}$  silicon wafer for example: the resultant device area is  $480\ \mu\text{m} \times 480\ \mu\text{m}$  if an effective device area of  $200\ \mu\text{m} \times 200\ \mu\text{m}$  is required. Comparatively, the device area is comparable to the effective area if a deep dry etching method is used to get a steep



**Fig. 30** Basic structures of FBAR: (a) Structure released by bulk micromachining; (b) Air-gap type FBAR formed by surface micromachining; (c) Solid mounted resonator (SMR) with Bragg reflector

sidewall, as shown in the right figure of Fig. 30a. However, both of the two structures have the mechanical reliability issues, and usually a supporting layer is used to enhance the mechanical stability.

Figure 30b shows the FBAR structure obtained with surface micromachining method. The air gap is formed on the top layer of the silicon wafer, and the overall mechanical stability is improved. Also the surface micromachining method has the potential to be integrated with standard CMOS processes. To release the FBAR structure, the sacrificial material filling the air gap should be carefully chosen to sustain the following processes like sputtering, lithography, plasma etching, etc.

The distinct feature of the third type FBAR, i.e., solidly mounted resonator (SMR) from the other two structures, lies in its different principle to realize the acoustic wave reflection at the lower boundary. The Bragg reflector consists of multiple stacking layers. Each layer has a thickness of a quarter of wavelength, and the adjacent layer has quite different acoustic impedance to ensure the total reflection of the acoustic wave. This structure processes the best mechanical reliability. It usually requires multiple layers of up to 8–10 for deposition, and the thickness of each layer should be controlled very precisely, which poses great challenges to the fabrication processes. Also the quality factor realized with this structure is still inferior to the device structure bounded by air.

### Fabrication Process

Because of the CMOS compatibility and the relative higher mechanical stability, the air-gap type FBAR is chosen in our experiments. The detailed fabrication processes based on the surface micromachining process is illustrated in Table 4. The process flow consists of four lithography steps, three thin film sputtering and etching processes. Dry etching methods are used for the electrode etching, AlN etching and FBAR releasing, which satisfies the CMOS compatibility requirements for further integration processes. Figure 31 shows the optical images of the same FBAR during the final step of air-gap formation. As indicated by the color changing around the center area, the lateral etching area becomes wider with an etching time increase from 20 min to 130 min. The center effective area is fully suspended after 130 min dry etching. In the experiment, we actually increased the etching time to remove all the possible residues at the bottom of the released structure.

### FBAR Characterizations



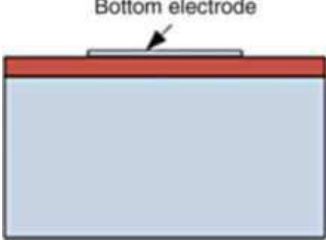
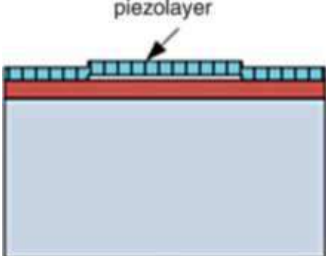

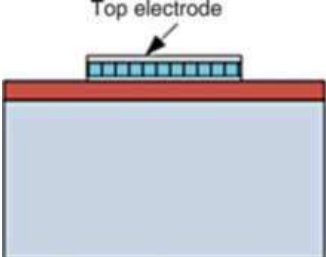
S parameters of the FBAR are measured with network analyzer HP8722S and microwave probe station. Standard calibration is performed prior to the test using the following parameters: frequency range 50 MHz ~ 3500 MHz, data points 1601, IFBW 100 Hz. The equivalent coupling coefficient can be obtained by

$$K_{eff}^2 = \frac{\pi^2}{4} \frac{f_p - f_s}{f_p}, \quad (8)$$

where  $f_p$  and  $f_s$  are the parallel and serial resonance frequency, respectively. Figure 32 shows the  $S_{11}$  and  $S_{21}$  parameters of a FBAR with a square center area. The  $f_p$  and  $f_s$



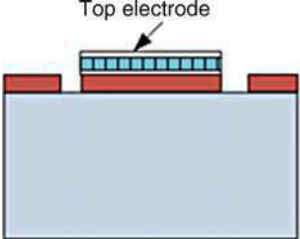
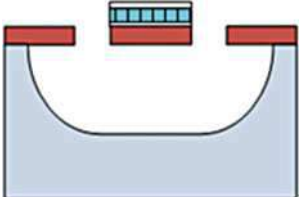
**Table 4** Fabrication processes of the air-gap type FBAR

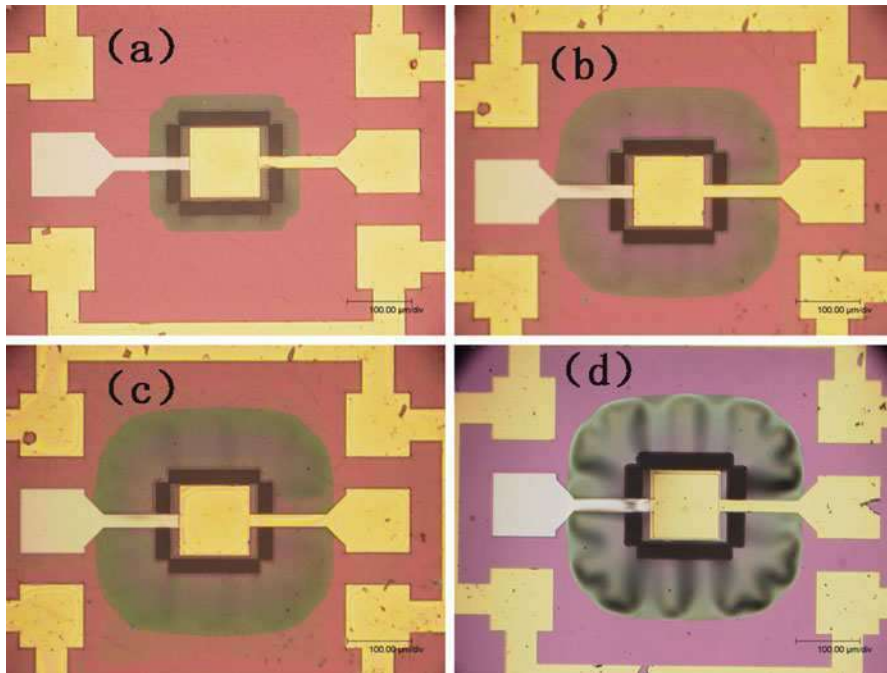
Fabrication process	Schematic illustration
<p>1. Silicon wafer preparation: High-resistivity 4-inch silicon wafer is used to suppress the loss, and standard RCA cleaning process is used to clean the silicon wafer</p>	
<p>2. SiO<sub>2</sub> deposition: 5000 Å SiO<sub>2</sub> is formed by thermal oxidation. The SiO<sub>2</sub> is used as the supporting layer and insulation layer between the FBAR and substrate</p>	
<p>3. Bottom electrode deposition and patterning: Pt/Ti thin film is sputtered on the SiO<sub>2</sub>, followed by photolithography and ion beam etching to define the bottom electrode</p>	
<p>4. Piezoelectric thin film deposition: A 1 μm AlN is sputtered on the bottom electrode and SiO<sub>2</sub>. DC magnetron sputtering method using Al as the substrate is used in an atmosphere of Ar/N<sub>2</sub> to obtain the (002) preferred orientation thin film</p>	
<p>5. Piezoelectric thin film patterning: A second photolithography is used to define the AlN pattern, followed by reactive ion etching using Cl<sub>2</sub>/BCl<sub>3</sub> gas</p>	
<p>6. Top electrode deposition and patterning: Top electrode is formed by using photolithography, electrode material Au/Ti (1500 Å/100 Å) evaporation, and lift-off processes</p>	

(continued)



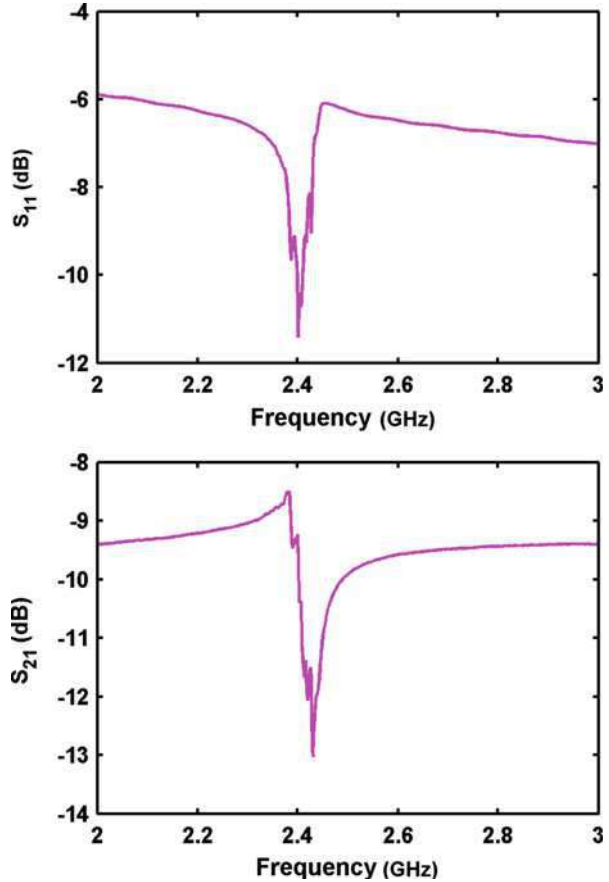
**Table 4** (continued)

Fabrication process	Schematic illustration
<p>7. Fourth photolithography: This step is used to form the etching window for the following etching process. After the photolithography, the <math>\text{SiO}_2</math> exposed is etched out by wet etching or dry etching method, and the silicon underneath is exposed</p>	
<p>8. Air-gap formation: By using the anisotropic dry etching method, the FBAR structure is released from the silicon substrate. <math>\text{SF}_6</math> is used as the etchant</p>	



**Fig. 31** Optical images of a FBAR after dry etching for (a) 20 min, (b) 100 min, (c) 130 min, and (d) photoresist stripping out

**Fig. 32** Measured S parameters of a FBAR with a square center area



are 2.382 GHz and 2.430 GHz, respectively, which gives an equivalent coupling coefficient of 4.87%.

## Integration Strategy and System-Level Simulation

To the date, the FBAR and the CMOS RF integrated circuits are separately manufactured and assembled on the same printed board or packaged by using wire bonding, and thus have the limitation in reduction of the system size. Furthermore, parasitic effects will degrade the system performance. To solve the above problems, various integration strategies including monolithic and heterogeneous integration methods were proposed (Nam et al. 2008; Carpentier et al. 2005). Based on 0.25  $\mu\text{m}$  SiGe BiCMOS processes, the European Martina project implemented the integration of the FBAR and a RF front-end module by using polymer as the sacrificial material (Kerhervé et al. 2006). By utilizing a similar structure, researchers in Taiwan University and Columbia University implemented the integration of air-gap type

and Bragg reflector type FBAR based on 0.18  $\mu\text{m}$  CMOS process (Sung et al. 2007; Johnston et al. 2010). While achieving good results, the reported methods utilized complicated processes like chemical mechanical polishing (CMP) and some special sacrificial materials, and their high integration cost greatly limited their applications.

In this section, a novel strategy for post-CMOS integration of AlN-based FBAR has been proposed. The AlN can be deposited by reactive sputtering method at moderate temperature and the whole integration process is compatible with the CMOS process. The feasibility of the integration strategy like the sacrificial layer deposition and the FBAR structure releasing has been experimentally verified. The design of the 2.4 GHz oscillator based on the UMC 0.18  $\mu\text{m}$  CMOS process and the FBAR has been demonstrated and the simulation results of the oscillator in terms of the phase noise clearly illustrates the merits by introducing the FBAR into the oscillator.

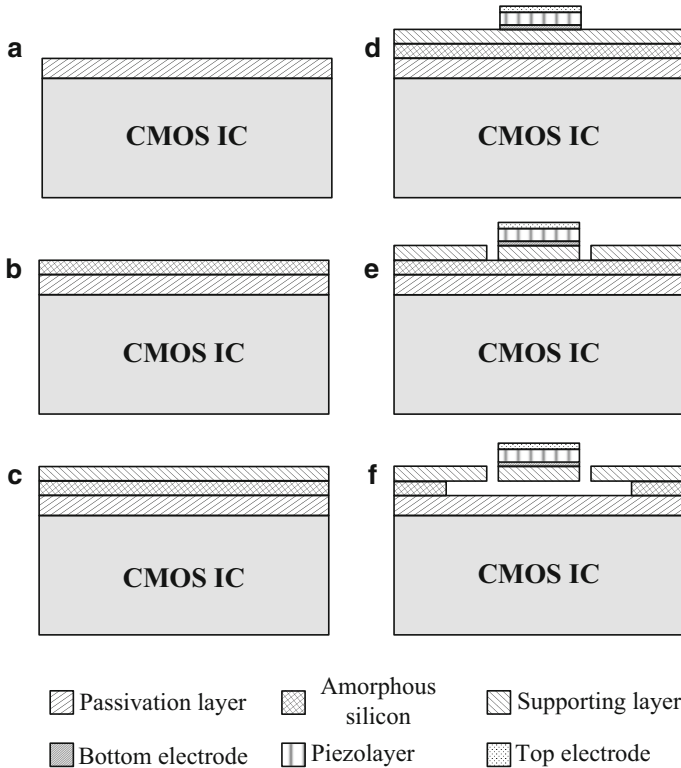
### Integration Strategy

It is known that the two sides of the FBAR should be acoustically isolated to ensure a high-quality factor FBAR. This is usually implemented by using air-gap below the lower electrode or Bragg reflectors. We propose to integrate of the FBAR onto the CMOS IC by using the air-gap structure. Shown in Fig. 33 is the proposed integration process flow for FBAR post-CMOS integration. The supporting layer could be  $\text{SiO}_2$  deposited by PECVD method because of the low temperature characteristic of the PECVD process and high etching selectivity of silicon over  $\text{SiO}_2$ . Both of the electrode and the piezoelectric materials can be deposited by sputtering method at room temperature. It is designed that the highest temperature through the proposed back end process is no more than 350  $^\circ\text{C}$ . To verify the feasibility of the proposed strategy, we have fabricated the FBAR using Si thin film as the sacrificial material.

### FBAR Fabrication and Characterization

Mostly, the fabrication processes in the integration strategy is similar to that described in Table 4 except for the sacrificial layer material. Figure 34a shows the cross-sectional SEM picture of the FBAR structure before the final releasing process (Xiao et al. 2014). The silicon layer with a thickness of 1.8  $\mu\text{m}$  is used as the sacrificial layer. The AlN thickness used in this study is about 640 nm and a 450 nm  $\text{SiO}_2$  is used as the supporting layer. Figure 34b shows the optical image of a FBAR device after etching the Si sacrificial layer. It is noticed that the effective center area has an eclipse shape, which is designed to eliminate the superimposition of the reflected wave from the boundaries. This design is also helpful to achieve a high quality factor.

The measured S parameters of the FBARs are shown in Fig. 35, from which multiple wave modes can be identified. Both the thickness shear modes (TSM) and thickness extension modes can be excited with the fabricated FBARs, which is due to the nonperfect (002) aligned AlN thin film. Comparatively, the shear wave modes have lower wave velocities than extension modes, which can be also verified by



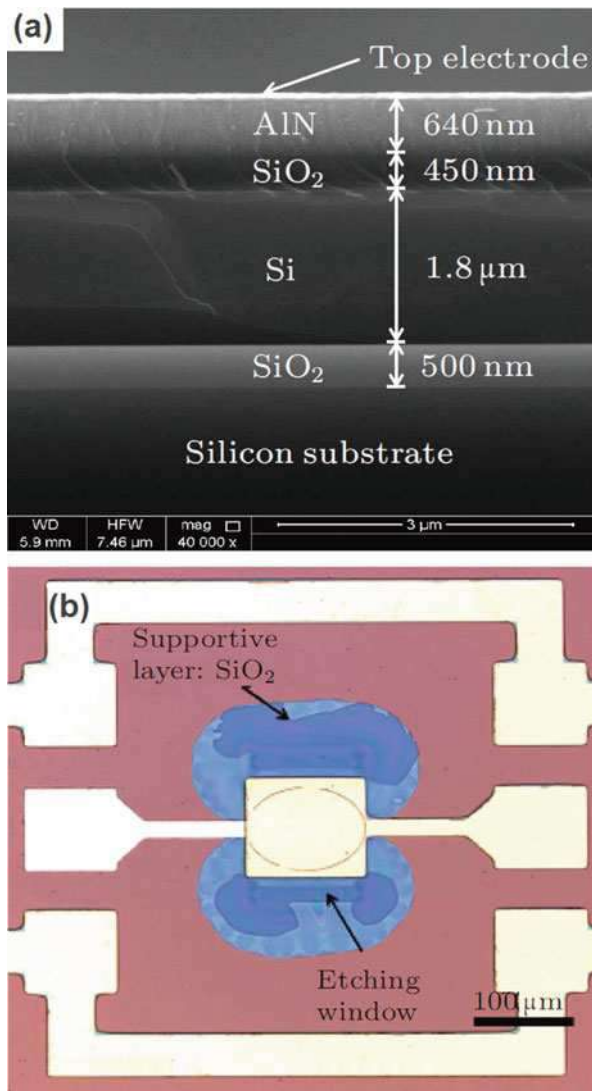
**Fig. 33** FBAR post-CMOS integration process flow: (a) Fabricated CMOS IC with a passivation layer; (b) PECVD amorphous silicon as the sacrificial material; (c) Supporting layer deposition; (d) Top electrode/piezoelectric layer/bottom electrode sandwich structure formation; (e) Etching hole formation beside the sandwich structure to expose the underneath sacrificial layer; (f) Isotropic plasma dry etching to form the air gap, thus releasing the FBAR structure

comparing the resonance frequency of the two wave modes. For the first TEM mode, a high frequency of 4.10 GHz along with a high coupling coefficient of 6.47% is obtained. Thus it is conclusive that the proposed integration strategy is feasible for high-performance FBAR fabrication.

**System-Level Simulation**

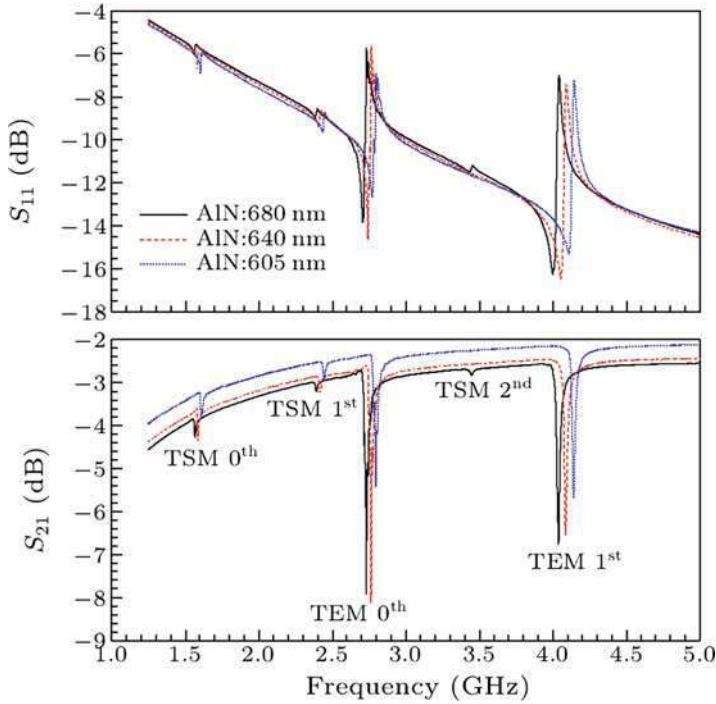
The UMC 0.18 μm CMOS process-based oscillator using the fabricated FBAR has been designed and its performance has been simulated. The parameters of the FBAR extracted from the measured results are used in the simulation. Figure 36 presents the phase noise of the FBAR-based oscillator. It is noticed that the phase noise is as low as -158dBc/Hz@1 MHz, which is not achievable by using traditional lumped elements or wire-bonded FBARs.

**Fig. 34** (a) Cross-sectional SEM image of the FBAR structure before the final releasing process; (b) optical image of a FBAR device after the final releasing process



### Flexible FBAR

Flexible FBAR has potential applications in flexible wireless communication systems and biosensors. As the FBAR resonant structure should be acoustically isolated from the environment by using air-gap or Bragg reflector, a flexible substrate cannot be directly used with the conventional FBAR design. In this regard, a novel mechanism is proposed to utilize the surface electrode-only structure for the flexible FBAR, whose working mechanism, fabrication process, and characterization are detailed below (Zhou et al. 2015).



**Fig. 35** Measured S parameters of the FBARs

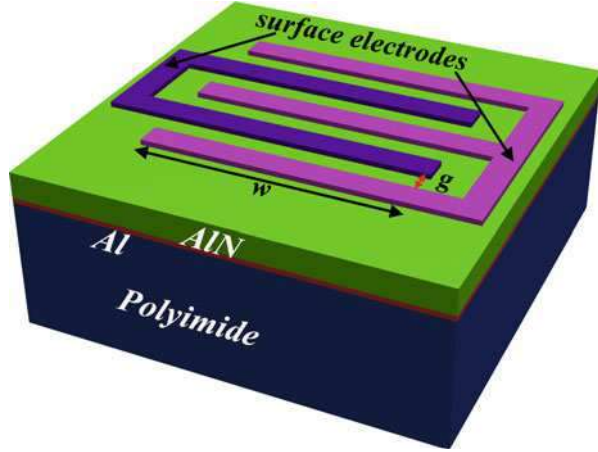


**Fig. 36** Simulated phase noise of the oscillator with integrated FBAR

**Device Structure and Working Mechanism**

Figure 37 shows the schematic of the flexible FBAR. Here  $w$  stands for the overlap length of the electrodes, and  $g$  denotes the distance of the electrodes. Compared to the traditional electrode/piezoelectric layer/electrode sandwich structure, this

**Fig. 37** Schematic of the flexible FBAR (Reprinted with permission of IOP)



structure only consists of a layer of electrode, just as in the case of a SAW device. In this structure, polyimide is used as the flexible substrate and AlN is preferred as the piezoelectric material.

As the lower boundary of the resonant structure is not bounded by air or a perfect Bragg reflector, a part of the acoustic wave will be transported into the underlying PI substrate. Considering that the sound pressure of the acoustic wave excited in AlN is  $p_0$ , and the sound pressure of the reflected wave at the AlN/PI boundary is  $p_r$ , the reflecting coefficient of the sound pressure is defined as

$$r = \frac{p_r}{p_0} = \frac{z_2 - z_1}{z_2 + z_1} \tag{9}$$

The transmission coefficient of the sound pressure is then defined as the ratio of the sound pressure of the transmitted wave to that of incident wave:

$$t = \frac{p_t}{p_0} = \frac{2z_2}{z_2 + z_1} \tag{10}$$

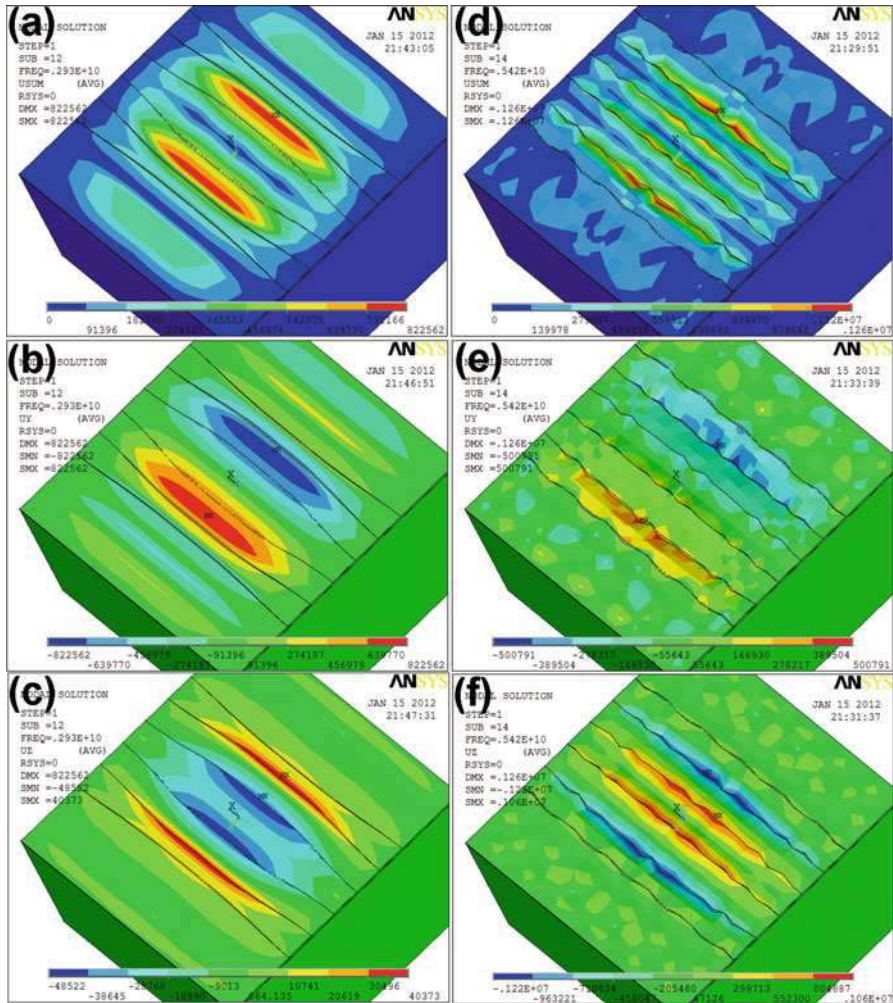
where  $z_1$  and  $z_2$  are the acoustic impedance of AlN and PI, and  $z_1 = 36 \times 10^6 \text{ kg/m}\cdot\text{s}$ ,  $z_2 = 1.7 \times 10^6 \text{ kg/m}\cdot\text{s}$ , which result in

$$r = \frac{1.7 - 36}{1.7 + 36} = -0.91 \tag{11}$$

$$t = \frac{2 \times 1.7}{1.7 + 36} = 0.09. \tag{12}$$

It is noticed that most of the acoustic wave energy will be reflected into the upper layer because of the large acoustic impedance mismatch at the boundary.

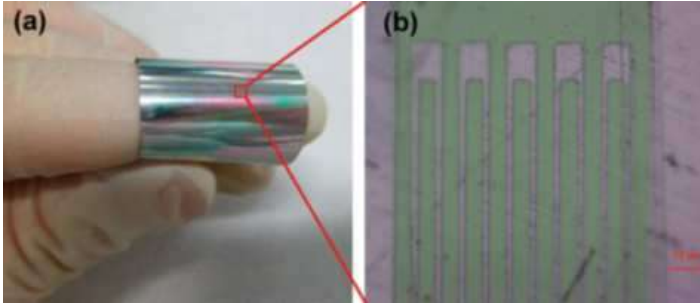




**Fig. 38** Mechanical displacements characteristics of the two wave modes. (a, b, and c) depict the total displacement (Usum), y-component displacement (Uy), and z-component displacement (Uz) of the first mode, respectively; (d, e, and f) depict Usum, Uy, and Uz of the second mode, respectively (Reprinted with permission of IOP)

The working mechanism of the flexible FBAR is elaborated by performing finite element analysis. As shown in Fig. 38, only one pair of surface electrode is modeled to reduce the computation time while still keeping the accuracy. First, a harmonic analysis is performed to obtain the frequency spectrum of the impedance, from which we can identify two wave modes, i.e., TSM and TEM. And the resonance frequencies of the two wave modes are 2.93 GHz and 5.42 GHz. The ratio of the acoustic wave velocity of the TSM to that of TEM is 0.54. The modal analysis is then performed to confirm the dynamic vibration mode shape of the two wave modes.





**Fig. 39** (a) Photograph of the fabricated LFE FBAR; (b) Microscope image (1000 $\times$ ) of one sample with five pairs of interdigital electrodes: the width of the electrode and the gap distance between the positive and negative electrode are 5  $\mu\text{m}$  and 1.5  $\mu\text{m}$ , respectively (Reprinted with permission of IOP)

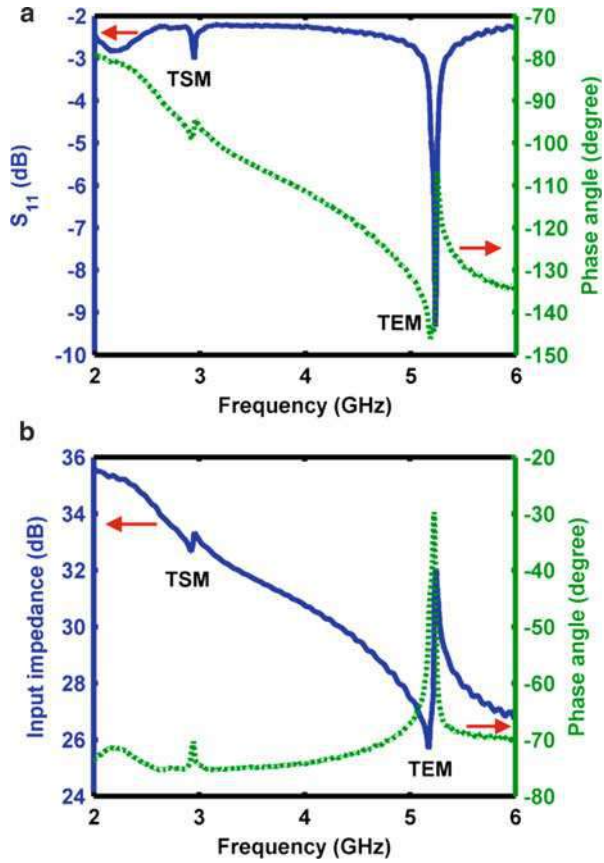
Figure 37a, b, and c present the total displacement ( $U_{\text{sum}}$ ), the lateral y-component of displacement ( $U_y$ ) and thickness directional z-component of displacement ( $U_z$ ) of the first wave mode, and figure (d), (e), and (f) shows the displacement of the second wave mode with a higher resonance frequency. For the first wave mode, the lateral displacement  $U_y$  is about one order of magnitude larger than  $U_z$ , confirming its characteristic as TSM wave. Similarly, the thickness directional displacement of the second wave mode is much larger than that of the lateral displacement, which corresponds to the characteristic of a TEM wave. Another difference lies in the position of the maximum displacement, which is at the center of the electrode gap for TSM, and at the boundary of the electrode for TEM.

### FBAR Fabrication and Characterization

Figure 39a shows the photograph of the flexible FBAR based on PI substrate. It is indicated that the FBAR can be attached to a curved surface like the finger. The photo micrograph image of a sample is shown in Fig. 39b. The surface electrodes consist of five pairs of interdigital electrodes. The width of the electrode and the gap distance  $g$  between the positive electrode and negative electrode are 5  $\mu\text{m}$  and 1.5  $\mu\text{m}$ , respectively. The length of the electrode  $w$  is 500  $\mu\text{m}$  and the total area of a fabricated device is about 0.5mm<sup>2</sup>.

The measured frequency responses of the flexible LFE FBAR are presented in Fig. 40. Two resonant peaks can be observed from the  $S_{11}$  and input impedance frequency spectrum. According to the finite element analysis, the first peak is due to the TSM wave, which has a resonance frequency of 2.945 GHz and a coupling coefficient of 2.72%. In case of the TEM wave, the resonance frequency of 5.233 GHz and a coupling coefficient of 3.14% are obtained. The resonance frequencies of the two modes are inconsistent with the finite element analysis. In terms of quality factor, the TEM wave shows a much higher quality factor of 176 than that of the TSM wave, suggesting its more suitability for wireless communication and sensing applications.

**Fig. 40** Measured (a)  $S_{11}$  and (b) input impedance spectrum of the flexible FBAR (Reprinted with permission of IOP)



## Summary and Outlook

This chapter presents an overview of the basic principles of acoustic wave devices and some recent work in this area. SAW and FBAR devices are two of the most important passive devices. They are widely used both in telecommunication and sensing fields. Resonance frequency  $f_r$ , quality factor  $Q$ , and coupling coefficient  $k^2$  are three key parameters for these resonators. As one of the most efficient methods to improve the device performance, several novel multilayered structures are presented and discussed in this chapter. For wave velocity enhancement, there are structures including AlN/diamond/Si and LN/diamond/Si. For temperature compensation, LN/SiO<sub>2</sub>/LN and LN/SiO<sub>2</sub>/Si are two typical examples. Flexible SAW and FBAR devices are also introduced to show their potential on many other novel substrates. As for practical applications, a SAW temperature sensor and a SAW biosensor for DNA detecting are presented as well.

SAW and FBAR devices have shown excellent performance in telecommunication fields especially as filters and duplexers. They are already mass produced and the technology is well-developed. However, for sensing fields, situation is quite different. Acoustic wave devices are promising candidates for sensing function because of their low price, small size, simple process, and high sensitivity. Although many kinds of acoustic wave sensors are realized in laboratory, few commercial sensor products are achieved. We believe there are still works to do for not only performance improvement but also system integration. Recent research trends also include combination with many other new materials such as two-dimensional materials.

---

## References

- Ahmed S, Tahir FA, Shamim A (2015) A compact Kapton-based inkjet-printed multiband antenna for flexible wireless devices. *IEEE Antennas Wirel Propag Lett* 14:1802–1805
- Bo L, Xiao C, Hualin C (2016) Surface acoustic wave devices for sensor applications. *J Semicond* 37(2):021001
- Cai HL, Yang Y, Zhang YH (2014) A high sensitivity wireless mass-loading surface acoustic wave DNA biosensor. *Mod Phys Lett B* 28(07):1450056
- Cai HL, Yang Y, Chen X (2015) A third-order mode high frequency biosensor with atomic resolution. *Biosens Bioelectron* 71:261–268
- Campbell C (1989) Surface acoustic wave devices and their signal processing applications. Academic, Boston
- Carpentier JF, Cathelin A, Tilhac C (2005) A SiGe: C BiCMOS WCDMA zero-IF RF front-end using an above-IC BAW filter, Solid-state circuits conference, 2005. Digest of technical papers. ISSCC. 2005 I.E. International. IEEE, San Francisco, CA, USA
- Chen NH, Huang JC, Wang CY (2011) Fabrication of a GHz band surface acoustic wave filter by UV-nanoimprint with an HSQ stamp. *J Micromech Microeng* 21(4):045021
- Chen X, Mohammad MA, Conway J (2015) High performance lithium niobate surface acoustic wave transducers in the 4–12 GHz super high frequency range. *J Vac Sci Technol B Nanotechnol Microelectron: Mater Process Meas Phenom* 33(6):06F401
- Dogheche E, Remiens D, Shikata S (2005) High-frequency surface acoustic wave devices based on LiNbO<sub>3</sub>/diamond multilayered structure. *Appl Phys Lett* 87(21):213503
- Dubois MA, Muller C (2013) Thin-film bulk acoustic wave resonators. In: MEMS-based circuits and systems for wireless communication. Springer, New York
- Hashimoto K, Sato S, Teshigahara A (2013) High-performance surface acoustic wave resonators in the 1 to 3 GHz range using a ScAlN/6H-SiC structure. *IEEE Trans Ultrason Ferroelectr Freq Control* 60(3):637–642
- Johnston ML, Kymissis I, Shepard KL (2010) FBAR-CMOS oscillator array for mass-sensing applications. *IEEE Sensors J* 10(6):1042–1047
- Jung YH, Qiu Y, Lee S (2016) A compact parylene-coated WLAN flexible antenna for implantable electronics. *IEEE Antennas Wirel Propag Lett* 15:1382–1385
- Kerhervé E, Ancy P, Aid M (2006) 4D-5 BAW technologies: development and applications within MARTINA, MIMOSA and MOBILIS IST European projects, Ultrasonics symposium, 2006. IEEE
- Khang DY, Jiang H, Huang Y (2006) A stretchable form of single-crystal silicon for high-performance electronics on rubber substrates. *Science* 311(5758):208–212
- Kim J, Son D, Lee M (2016) A wearable multiplexed silicon nonvolatile memory array using nanocrystal charge confinement. *Sci Adv* 2(1):e1501101
- Lee TC, Lee JT, Robert MA (2003) Surface acoustic wave applications of lithium niobate thin films. *Appl Phys Lett* 82(2):191–193

- Morgan D (2010) Surface acoustic wave filters: with applications to electronic communications and signal processing. Academic, Amsterdam
- Nakahata H, Fujii S, Higaki K (2003) Diamond-based surface acoustic wave devices. *Semicond Sci Technol* 18(3):S96
- Nakamura H, Nakanishi H, Tsurunari T (2008) Miniature surface acoustic wave duplexer using  $\text{SiO}_2/\text{Al}/\text{LiNbO}_3$  structure for wide-band code-division multiple-access system. *Jpn J Appl Phys* 47(5):4052–4055
- Nam K, Park Y, Ha B (2008) Monolithic 1-chip FBAR duplexer for W-CDMA handsets. *Sensors Actuators A Phys* 143(1):162–168
- Oh JY, Rondeau-Gagné S, Chiu YC (2016) Intrinsically stretchable and healable semiconducting polymer for organic transistors. *Nature* 539(7629):411–415
- Ramadan KS, Sameoto D, Evoy S (2014) A review of piezoelectric polymers as functional materials for electromechanical transducers. *Smart Mater Struct* 23(3):033001
- Ruby R, Bradley P, Larson JD (1999) PCS 1900 MHz duplexer using thin film bulk acoustic resonators (FBARs). *Electron Lett* 35(10):794–795
- Setter N, Damjanovic D, Eng L (2006) Ferroelectric thin films: review of materials, properties, and applications. *J Appl Phys* 100(5):051606
- Shung KK, Cannata JM, Zhou QF (2007) Piezoelectric materials for high frequency medical imaging applications: a review. *J Electroceram* 19(1):141–147
- Sung PH, Chen PY, Chin YC (2007) Method of forming film bulk acoustic wave filter assembly. US Patent 7,214,564
- Tian XG, Tao LQ, Liu B (2016a) Surface acoustic wave devices based on high quality temperature-compensated substrates. *IEEE Electron Device Lett* 37(8):1063–1066
- Tian XG, Liu H, Tao LQ (2016b) High-resolution, high-linearity temperature sensor using surface acoustic wave device based on  $\text{LiNbO}_3/\text{SiO}_2/\text{Si}$  substrate. *AIP Adv* 6(9):095317
- Wang L, Rokhlin SI (2001) Stable reformulation of transfer matrix method for wave propagation in layered anisotropic media. *Ultrasonics* 39(6):413–424
- White RM, Voltmer FW (1965) Direct piezoelectric coupling to surface elastic waves. *Appl Phys Lett* 7(12):314–316
- Xiao C, Yi Y, Hua-Lin C (2014) A multiple resonant mode film bulk acoustic resonator based on silicon-on-insulator structures. *Chin Phys Lett* 31(12):124302
- Yeo LY, Friend JR (2014) Surface acoustic wave microfluidics. *Annu Rev Fluid Mech* 46:379–406
- Zhou C, Yang Y, Zhan J (2011) Surface acoustic wave characteristics based on c-axis (006)  $\text{LiNbO}_3$ /diamond/silicon layered structure. *Appl Phys Lett* 99(2):022109
- Zhou CJ, Yang Y, Shu Y (2012) Ultra flexible pseudo-lamb wave RF resonators based on  $\text{ZnO}/\text{PI}$  and  $\text{AlN}/\text{PI}$  structures, *Electron Devices Meeting (IEDM), 2012 IEEE International* (pp. 5–4)
- Zhou CJ, Yang Y, Shu Y (2013a) Visible-light photoresponse of  $\text{AlN}$ -based film bulk acoustic wave resonator. *Appl Phys Lett* 102(19):191914
- Zhou C, Yang Y, Jin H (2013b) Surface acoustic wave resonators based on (002)  $\text{AlN}/\text{Pt}/\text{diamond}/\text{silicon}$  layered structure. *Thin Solid Films* 548:425–428
- Zhou C, Yang Y, Cai H (2013c) Temperature-compensated high-frequency surface acoustic wave device. *IEEE Electron Device Lett* 34(12):1572–1574
- Zhou C, Shu Y, Yang Y (2015) Flexible structured high-frequency film bulk acoustic resonator for flexible wireless electronics. *J Micromech Microeng* 25(5):055003

---

**Part VII**

**Optical MEMS**



# Silicon-Based Optical Sensor: Uncooled Infrared Focal Plane Array Based on Bi-Materials Cantilever Microstructures

Tianchun Ye, Dapeng Chen, Shali Shi, and Binbin Jiao

## Contents

Development of MEMS-Based Optical-Readable Thermal Imaging Technology .....	1198
Modeling and Analysis of MEMS-Based Optical-Readable Thermal Imaging Technology .....	1201
Operating Principle of MEMS-Based Optical-Readable Thermal Imaging System .....	1204
Establishment of Equivalent Circuit Model .....	1210
Noise Analysis of the System .....	1213
Experimental Measuring of the Effects of System Noise .....	1218
Section Summary .....	1221
The Examination of and the Improvement on the FPA Unit Model .....	1221
Structure of the Single-Layer FPA .....	1221
Early Model of the Substrate-Free Single-Layer FPA .....	1222
The Model Validation .....	1229
Correction of the Model .....	1232
Simulation Design of Full Hollow Double-Layer Infrared Focal Plane Array Devices Based on the Sacrificial Layer .....	1240
Summary .....	1244
The Optimization and Improvement of Monolayer FPA .....	1244
The Shortage of the Previous Monolayer FPA .....	1244
Higher Sensitivity Gained in the Situation of Pixels' Size Scaling Down .....	1245
Improvement of Thermal Interference Among Pixels .....	1254
Designing of Optimized Single-Layer FPA .....	1255
Test of Optimized Single-Layer FPA .....	1256
Summary .....	1261
References .....	1262

---

T. Ye · D. Chen (✉) · S. Shi (✉) · B. Jiao  
Institute of Microelectronics of Chinese Academy of Sciences (IMECAS), Beijing, China  
e-mail: [dpchen@ime.ac.cn](mailto:dpchen@ime.ac.cn); [shishali@ime.ac.cn](mailto:shishali@ime.ac.cn)

---

**Abstract**

This chapter introduces the research work of MEMS-based optical-readable thermal imaging technology. An uncooled infrared focal plane array based on bi-materials cantilever microstructures is presented.

---

**Keywords**

MEMS · Optical-readable · IR image system · FPA

---

---

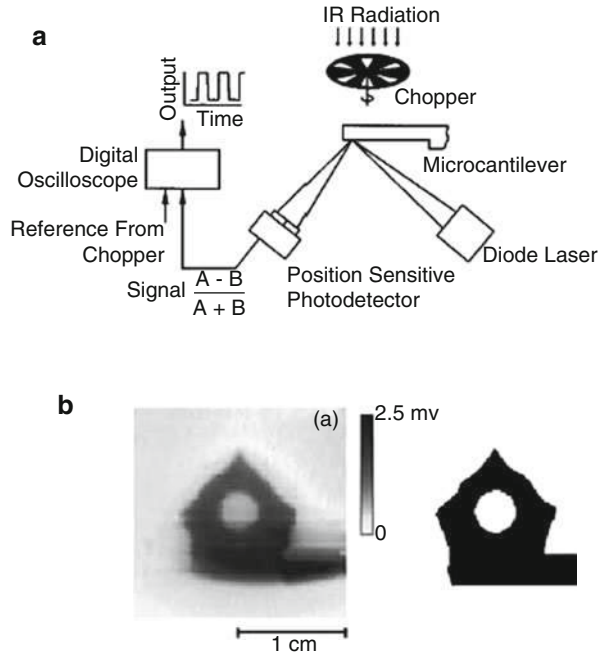
## Development of MEMS-Based Optical-Readable Thermal Imaging Technology

In recent years, MOMC (micro opto-mechanical cantilever) structure (Barnes et al. 1994; Lai et al. 1997; Varesi et al. 1997) and other optical readout method have been gradually introduced into the field of infrared detection and extensively researched as the technology of AFM (atomic force microcantilever) (Wachter et al. 1996) getting matured. The detecting unit of this optical readout is bi-material micro mechanical cantilever which is to be bent by shear-stress between the two levels of it because of the thermal expansion coefficient difference when absorbing infrared radiation and getting temperature rise. The temperature difference between the test object and the background can be obtained through detecting the thermal-induced deformation via optical method. Lai J (Lai et al. 1997) and other researchers provide that the theatrical detecting sensitivity of uncooled infrared focal plane array (IRFPA) can reach K magnitude.

It has been manifested practicable that bi-material micro-mechanical cantilever can be exploited as infrared detector in 1996 when a research group of Oak Ridge National Lab developed a sort of thermal detector which is made of civil piezoresistive microcantilever (Varesi et al. 1997; Oden et al. 1996; Datskos et al. 1998; Datskos 2000). The typical dimensions of bi-material cantilever in experiments were as follows: length in 100–200  $\mu\text{m}$ , thickness in 0.3–0.4  $\mu\text{m}$ , width in 10–50  $\mu\text{m}$ . The optical readout method was standard AFM optical imaging system as illustrated in Fig. 1a. The light generated by a He-Ne laser focused on the tip of the cantilever as an optical probe. The PSD would receive reflection beam and detect deflection displacement of the light spot as thermal deformation of the cantilever occurred. During the experiment, they observed the thermal image of an object of 250 °C for the first time. Thereafter, Varesi and his group optimized the cantilever and measured its thermal deformation with the same optical path, thus improving the energy resolution to picowatt ( $10^{-12}$  W) (Varesi et al. 1997).

In 1997, a research group of Stanford University developed a sort of two-dimensional cantilevers array with detecting units composed of 96  $\mu\text{m} \times 96 \mu\text{m}$  bi-material microcantilever structures similar to Fresnel zone plates (Manalis et al. 1997), whose focal distance alters as intervals among the spiral changing slightly when temperature rising occurs after absorbing incident infrared radiation. Thus, the experiment can detect the magnitude of the incident infrared radiation when visible light radiates upon the Fresnel zone plates by means of examining the intensity variation of its reflection through observation via a CCD installed at the original focal position.

**Fig. 1** Piezoresistive microcantilever used by Oak Ridge National Lab as infrared detector. (a) Detecting optical path. (b) Thermal imaging outcome (left: thermal imagery; right: corresponding object)



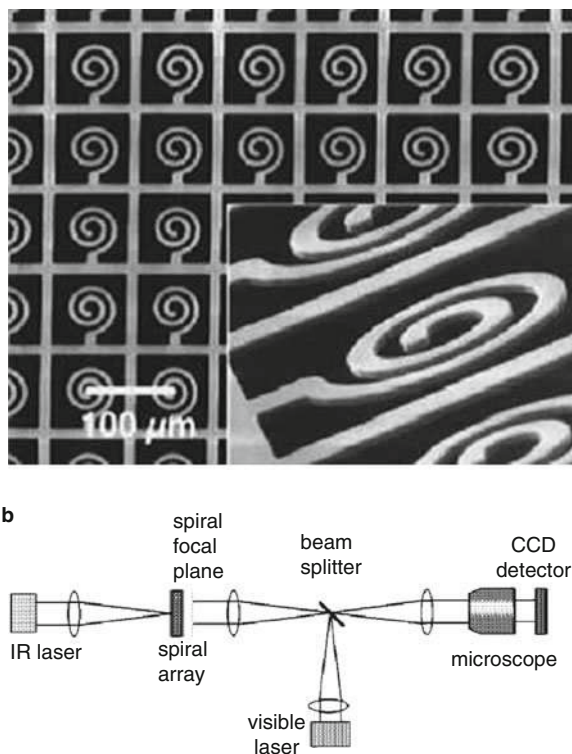
Majumdar's research group of University of California, Berkeley, had been designing and developing an IRFPA of bi-material microcantilevers based on optical interference readout method since 1979 (as illustrated in Fig. 2a) (Perazzo et al. 1999; Mao et al. 1999a, b; Norton et al. 2000; Zhao et al. 1999, 2002; Choi et al. 2003). In order to detect the deformation of each microcantilever unit, Majumdar's group improved the optical interference method (As illustrated in Fig. 2b). The light distribution of interference field read from the CCD directly displays the intensity of infrared thermal image as the optical path difference between FPA and comparison mirror plane being detected through interference method. With this interference method, they obtained a thermal image of an experimenter at a distance of 2 m. The Noise equivalent temperature difference of this system was 3–5 K (Fig. 3) (Zhao et al. 2002).

From the year 2000, Nikon Cooperation successively exposed an optical readout system of thermal-deformation inclination angle (as illustrated in Fig. 4) (Ishizuya et al. 2001, 2000, 2002; Ishizuya et al.; Suzuki et al.). Their system positioned on the spectral flatness of the optical  $4f$  processing system a pinhole through which the amount of light would differ when the inclination angle changes and thus the light intensity received by CCD will be different. With this system they also obtained thermal image of human beings and manifested the feasibility of the visibility of thermal image to naked eyes.

Zhang Liuqiang and other Chinese scholars designed and developed a sort of FPA structure (as illustrated in Fig. 5) and have the temperature field color coded with the optical imaging system provided by Veeco Cooperation (Zhang and Yang 2001). When the FPA element absorbs the infrared radiation temperature, the bi-material microcantilever will be deformed. Because of the screw symmetry of structure, the



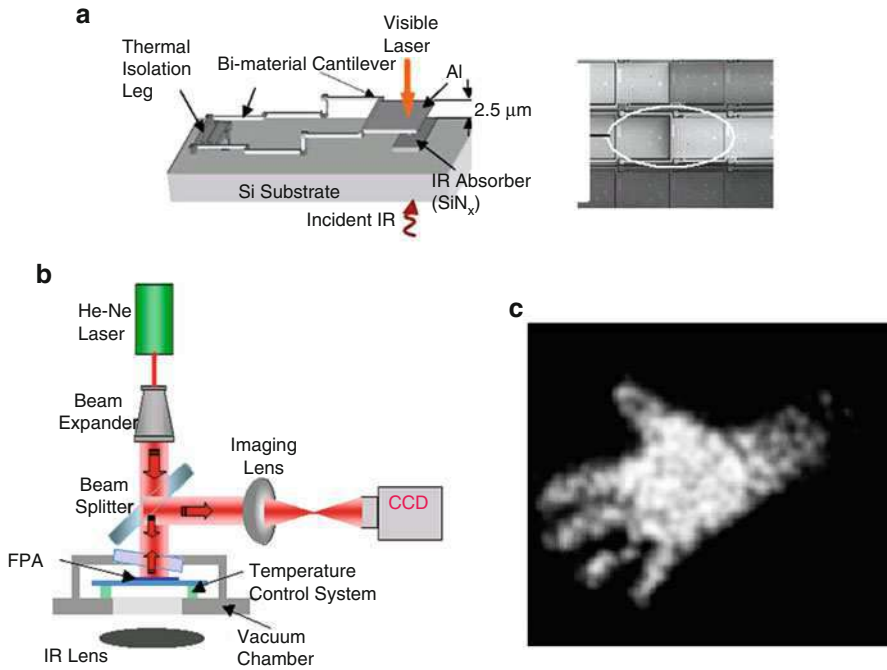
**Fig. 2** The Fresnel zone plate FPA and its optical readout system designed by Stanford University. **(a)** Detecting units. **(b)** A diagram of the optical readout



micro mirror at central merely occurs off-plane displacement. With the interference optical path, white illuminant can produce colorful image bearing temperature data as the distinct wavelength of lights in different color varies the intensity of interference thickness. Since the off-plane displacement occurs rather slightly, however, the sensitivity of this detecting method is quiet low.

All the productions of designed FPAs mentioned above are based on surface manufacturing process with sacrificial layer, which will develop a subsequently removable thin-film layer of substance (such as PSG and polycrystalline silicon) as sacrificial layer which is to be removed after the figure wanted being built upon to acquire a suspended component on some layer of thin-film cantilever element. The component's feature and manufacturing process are illustrated in Figs. 6 and 7. The absorbing efficiency of the detector unit, however, is considerably degraded as the essential silicon substrate of this structure will reflect and absorb radiation of the infrared object that a loss of approximately 40% of the radiation energy is inevitable for not being able to reach the absorbing layer.

Since 2001, Institute of Microelectronics of the Chinese Academy of Sciences and University of Science and Technology of China have cooperated on a research of optical readout uncooled infrared imaging technology and developed a sort of free-standing stencil bi-material microcantilever structure without silicon substrate as



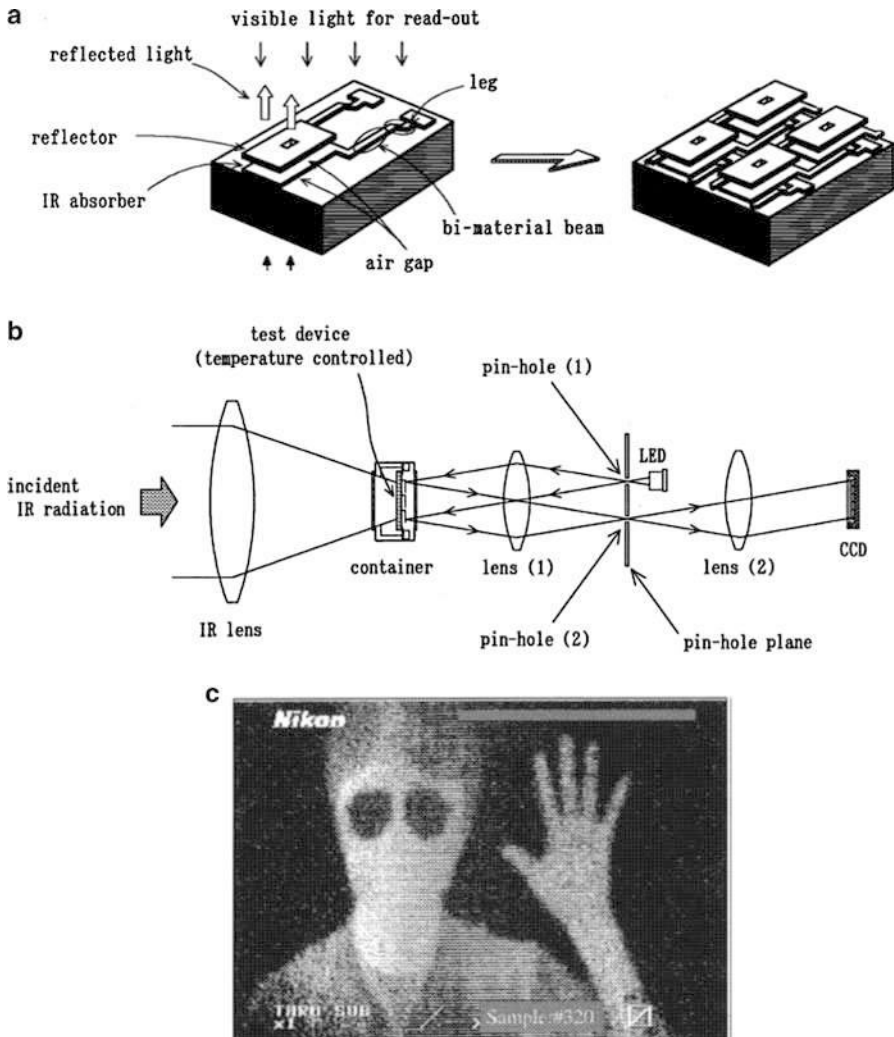
**Fig. 3** The FPA and its optical readout system designed by Majumdar group. (a) Bi-material microcantilever unit. (b) Fizeau interferometric optical readout system. (c) The thermal image of a hand

illustrated in Fig. 8a (Weibing et al. 2003, 2004; Li et al. 2006a, b; Dong et al. 2005, 2006; Miao et al. 2006), which magnificently improves infrared absorption efficiency of microcantilever units of FPA by allowing all the infrared radiation emitted by an object reaching its absorbing layer as well as noninterferometry light optical readout technology with higher detective sensibility as illustrated in Fig. 8b (Duan et al.; Duan 2003; Duan et al. 2003).

Figure 9 gives the outcomes of thermal imaging of human body at room temperature and Fig. 9a is an infrared thermal image of human hand whereas Fig. 9b is the one of head. The experimental result manifests that the system is able to detect IR image of objects at room temperature and its NETD is estimated at 200mK.

## Modeling and Analysis of MEMS-Based Optical-Readable Thermal Imaging Technology

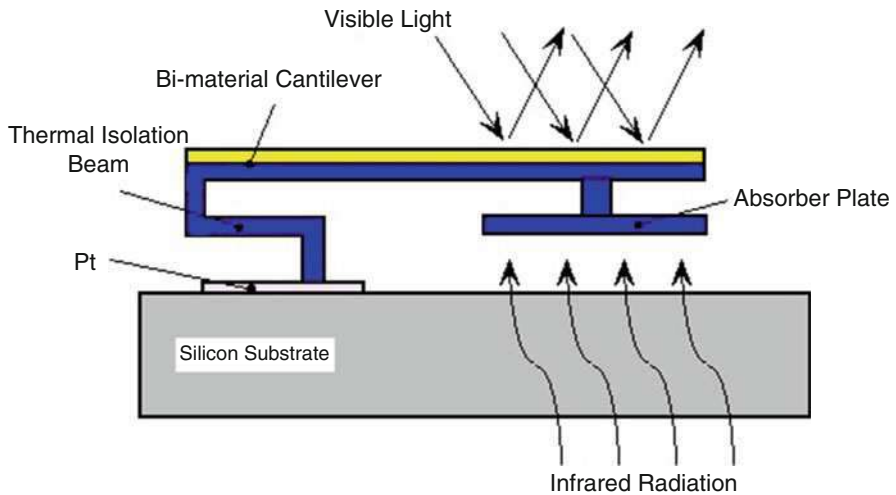
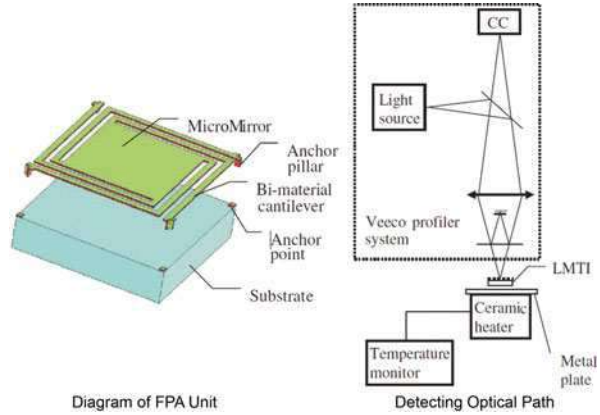
Though researches on MEMS-based optical-readable thermal imaging technology have been undertook for many years, their hotspots are mainly focused on modeling and analysis of core FPA whereas there is still a lack of a system-level description model of the system's comprehensive operating principle and performance. The



**Fig. 4** The bi-material microcantilevers array and pinhole filter system designed by Nikon Coop. (a) The microcantilevers array of Nikon Coop. (b) The optical readout system of the pinhole filter system. (c) The output thermal image of human beings

performance feature of optical readable infrared imaging system is not merely limited by the FPA itself but also considerably affected by optical detection modules at back-end (detecting optical path and CCD) and thus optical detecting module must be considered when representing the system's behavior. This section establishes a system-level model of equivalent circuit to illustrate a series of complex conversion process of light to heat, heat to deformation, deformation to light intensity, and light intensity to electricity with simplified circuit. The systematical performance will also

**Fig. 5** The FPA unit and its detecting optical path designed by Zhang’s group



**Fig. 6** A feature diagram of a component with sacrificial layer structure

be analyzed upon the base of the model to provide theoretical foundation of follow-up study. Infrared FPA characterized as thermal-mechanical sensor is a sort of uncooled infrared focal plate array (FPA) with structures of thin-film stencil bi-material microcantilever. It obtains images with photomechanical-effect-based optical readout system via temperature distribution of the infrared source, which can be gained through measuring the thermal-deformation distribution of cantilevers composed of two different materials with distinct thermal transmission under infrared radiation.

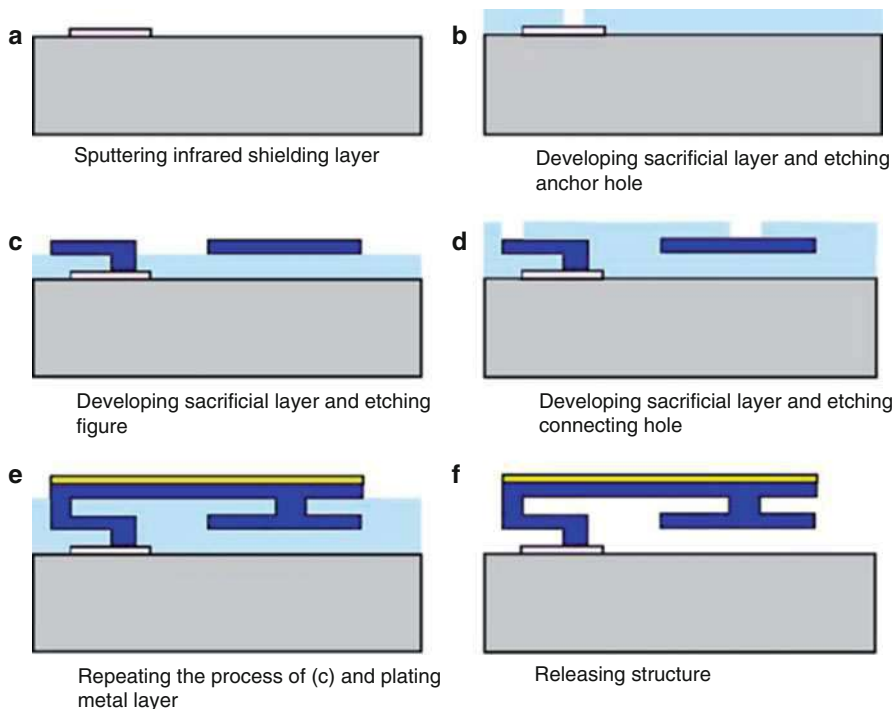


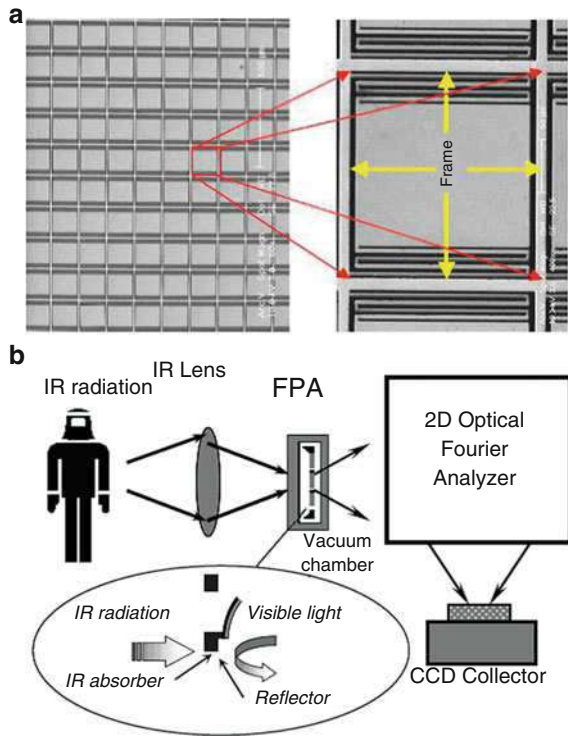
Fig. 7 A process sketch of component manufacturing

### Operating Principle of MEMS-Based Optical-Readable Thermal Imaging System

An MEMS-based optical-readable thermal imaging system will transform radiation energy into thermal energy of detecting units by IR absorptive structure, namely the MEMS technology processed microcantilever array, then transforming the thermal energy of units into microcantilever's inclination angle (displacement) through bi-material cantilever structure composed of metal thin film covered microcantilever before transforming the inclination angle of microcantilever into variation of light intensity via noninterference optical system of Knife-edge filter to obtain final outputs of grayscale after the light intensity being transformed into current by optical CCD (Weibing et al. 2004; Li et al. 2006a, b; Dong et al. 2005, 2006; Miao et al. 2006; Duan et al.; Duan 2003; Duan et al. 2003). Its operating principle is illustrated in Fig. 10.

The operating process will be introduced in two separate parts, the thermal-deformation process of MEMS microcantilever and the readout process of optical micro inclination angle, as below.

**Fig. 8** The FPA and its optical readout system designed by our group. (a) The stencil FPA. (b) Diagram of the optical readout system



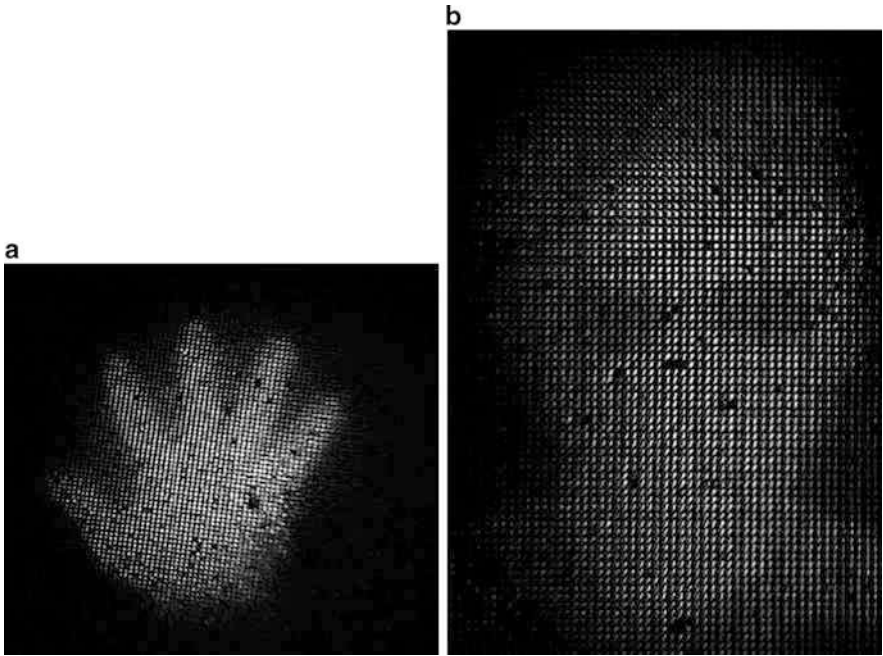
**Principle of Thermal-Deformation Process of MEMS Microcantilever**

The MEMS microcantilever referred in this article is a sort of double-layer-cantilever structure compounded by two materials with distinct thermal transmission (Weibing et al. 2004; Li et al. 2006a, b; Dong et al. 2005, 2006; Miao et al. 2006). The microcantilever will bend to deformation as the two bonded layers of material have different trends of expansion. The thermal deformation of double-layer-material cantilever is illustrated in Fig. 11.

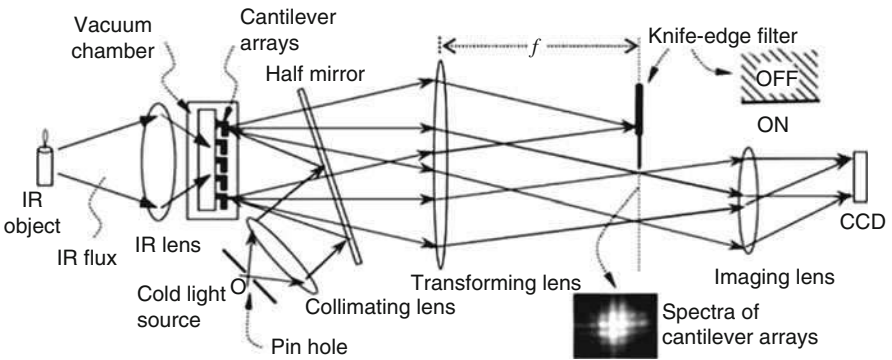
Supposing the thermal transmission of the two material being  $\alpha_1$  and  $\alpha_2$ , the Young modulus separately  $E_1$  and  $E_2$ , thicknesses  $t_1$  and  $t_2$ , lengths  $L$ , widths  $b$ , it can be deduced from relevant knowledge of material mechanics that the formula of displacement of microcantilever’s free end and temperature rise as below:

$$\delta = 3(\alpha_2 - \alpha_1) \left( \frac{p + 1}{K} \right) \left( \frac{L^2}{t_1} \right) \Delta T \tag{1}$$

Hereinto,  $K = 4 + 6p + 4p^2 + \varnothing p^3 + \frac{1}{\varnothing p}$  and  $p = \frac{t_2}{t_1}$  is the ratio of thicknesses of the two materials,  $\varnothing = \frac{E_1}{E_2}$  the ratio of the Young modulus of them.



**Fig. 9** The IR imaging outcomes of human body. (a) IR image of human hand. (b) IR image of human head

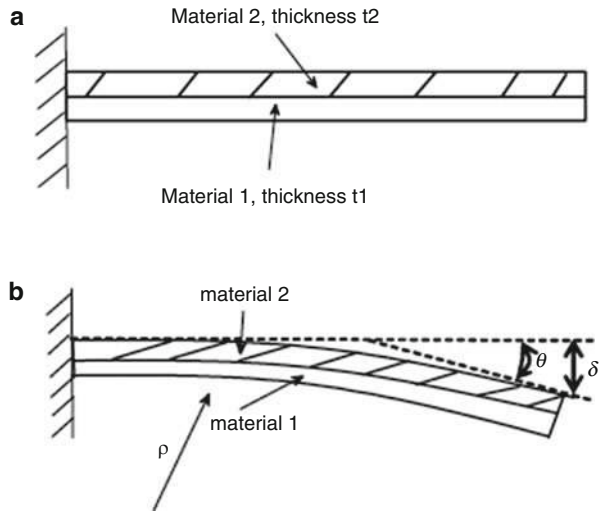


**Fig. 10** The schematic diagram of optical-readable uncooled infrared system

Thus, through the differential of the displacement of microcantilever's free end and inclination angle direction, the relation between the inclination angle  $\theta$  of cantilever and temperature variation is as below:

$$\theta = \left. \frac{d\delta}{dz} \right|_{x=L} = 6(\alpha_2 - \alpha_1) \left( \frac{t_1 + t_2}{t_1^2 K} \right) L \Delta T \quad (2)$$

**Fig. 11** Diagram of thermal deformation of bi-material oblique cantilever. (a) The bi-material oblique cantilever, when  $T = T_0$ . (b) The bi-material cantilever, when  $T = T_0 + \Delta T$

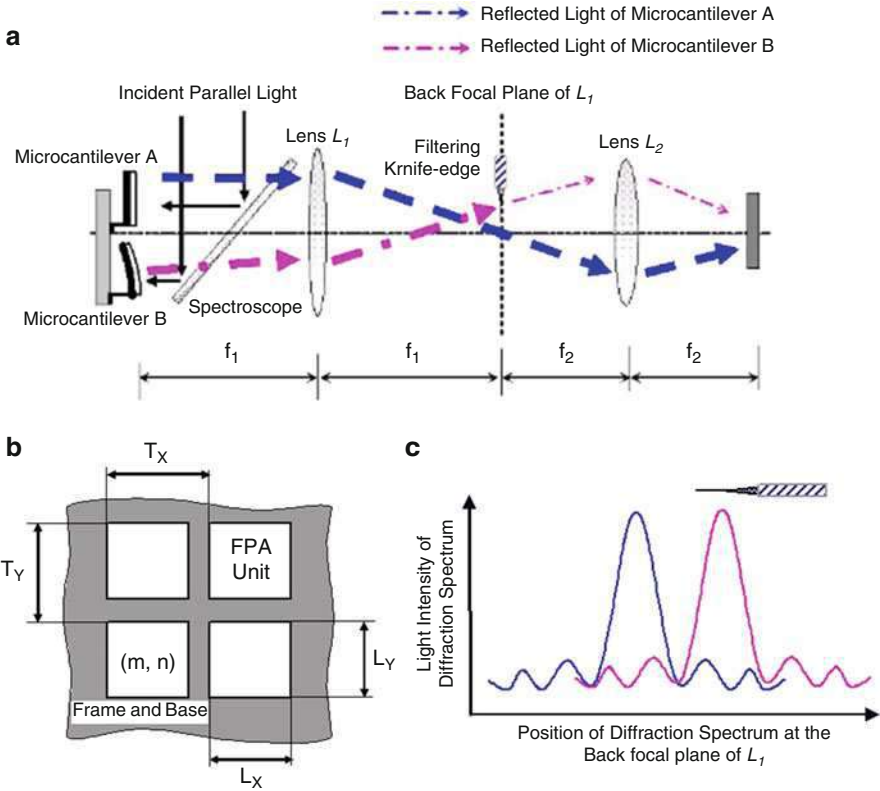


### Principle of Optical Readout System

The optical readout system referred below adopts spectral-flatness knife-edge filtering technology (Duan et al.; Duan 2003; Duan et al. 2003) which is a part of incoherent optical spatial filtering technology developed by University of Science and Technology of China and its operating principle is illustrated in Fig. 12.

Figure 12a illustrates the optical principle of  $4f$  optical system-based spectral-flatness knife-edge filtering technology. Through the half mirror, incident visible light reaches the surface of the microcantilevers array located at the front of the focal plane of the lens  $L_1$ . The light returned from a cantilever beam is transformed by Fourier to form a diffraction spectrum at the rear of the  $L_1$  focal plane where there placed an optical filtering knife-edge to control luminous flux of the diffraction spectrum. Having went through the filtering knife-edge, the light is received by the CCD positioned at the focal plane of the second lens  $L_2$  which should be passed and will produce a visible image. It is assumed that the optical reflection surface of each cantilever can be simplified to a regular rectangular plate. So a rectangular plate is arrayed to form a two-dimensional grating structure. The geometrical parameters of the grating are shown in Fig. 12b. the grating structure is under visible light, each rectangular plate of it gives the same diffraction spectrum upon spectral flatness whereas the diffraction spectrums of them are to displace from each other as the inclination angles of them are a little different when infrared radiation is being absorbed (as illustrated in Fig. 12c). The diffraction light of some cantilevers with larger inclination angles will mostly shift to opaque area of the filtering knife-edge thus the light passed through decreasing obviously, while one of those with smaller inclination angles mostly remains in the nonopaque area thus the decreasing of light being less (vice versa), so the light intensity received by different parts of an optical receiver will be distinct – in other words, the infrared thermal image is to be “seen.”





**Fig. 12** Diagram of the principle of spectral-flatness knife-edge filtering technology. (a) Diagram of the optical path of spectral-flatness knife-edge filtering technology. (b) Cantilevers array. (c) Diffraction spectrum pattern

Though light intensity passed through knife-edge filter and inclination angle of cantilevers are not linear dependent to each other (as illustrated in Fig. 13), it can be regarded as so when the inclination angle being tiny, especially as the approximate linear interval is located at the most sensitive part of the curve because optical energy concentrates at zero-order diffraction spectrum that the border of knife-edge just falls upon the center line of zero-order diffraction spectrum when the rate of change of optical energy transmittance maximized.

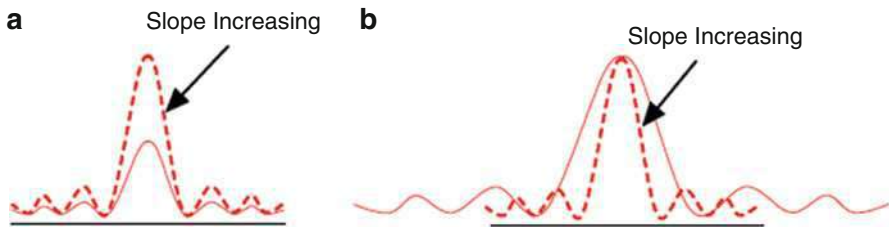
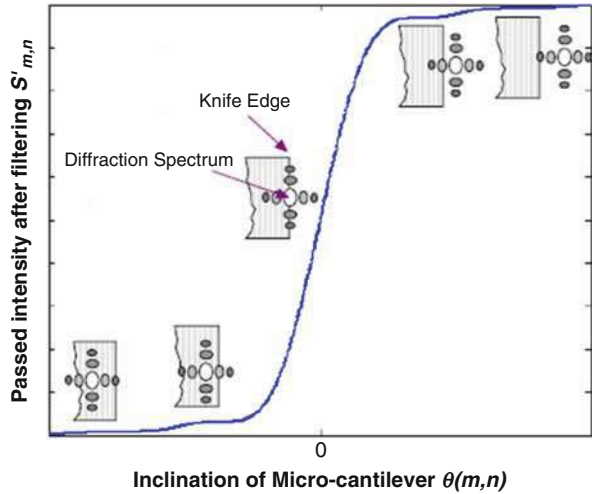
Detective sensitivity of spectral-flatness knife-edge filtering technology can be calculated by Fourier optics as:

$$D = -\frac{4\pi}{\lambda} * I_s * L_Y * R \tag{3}$$

$$\propto I_s * L_Y$$

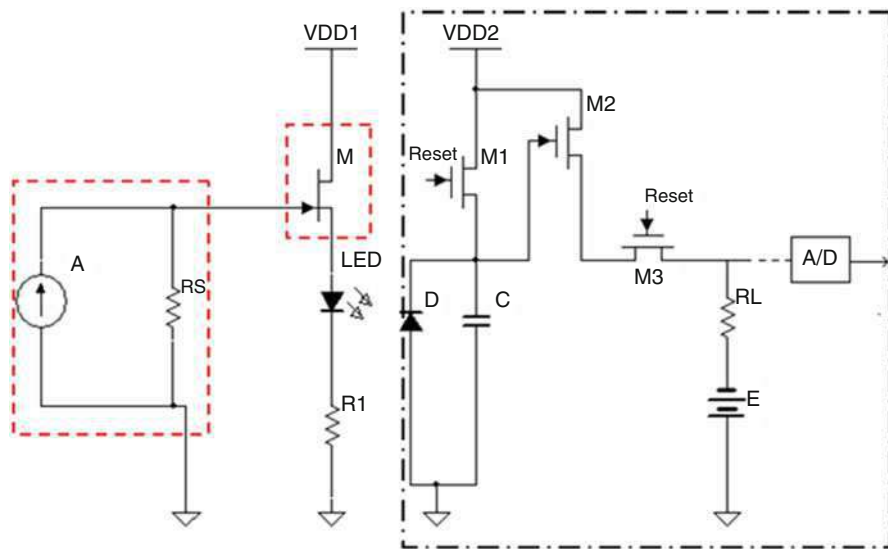
Hereinto,  $\lambda$  is the readout optical wavelength and  $I_s$  the total light intensity upon cantilevers, which is a constant as to the linear interval. The formula (3) manifests

**Fig. 13** Light intensity through knife edge filter



**Fig. 14** Influence factors of optical detective sensitivity. (a) Sensitivity improving as incident light intensity increasing, diffraction peak heightening, and change rate of light intensity per unit inclination-angle enlarging. (b) Sensitivity improving as geometric scale of cantilever at inclination direction increasing, width of diffraction peak decreasing, and change rate of light intensity per unit inclination-angle enlarging

that detective sensitivity are in direct proportion to the light intensity upon micro-cantilevers and the geometric scale of microcantilevers' inclination angle. As diffraction width of cantilever only depends on its geometric scale, increasing light intensity upon cantilever just raises the diffraction peak – namely the change rate of light intensity led by unit's inclination is more intense which is to improve detective sensitivity. Meanwhile, the larger geometric scale of cantilever at inclination direction is the narrower width of diffraction peak becomes (as illustrated in Fig. 14b), and the more intense change rate of light intensity is caused by per unit inclination-angle; thus, the sensitivity will improve with them. According to formula (3), detective sensitivity is irrelevant to focal distance of lens so it can be chosen at convenience which paves the way of miniaturization and practicalization of MEMS-based optical-readout infrared detection.



**Fig. 15** Photoelectric coupling circuit model of MEMS-based optical-readout uncooled infrared system

### Establishment of Equivalent Circuit Model

As operating principle of MEMS microcantilever-array based on optical infrared imaging system involves mechanical deformation process of cantilevers, optical modulation process of optical readout and the final electrical process, a photoelectric coupling model is established to illustrate its operating principle and noise analysis as Fig. 15 for the purpose of describing optical readout infrared imaging system clearly and explicitly (Jiao et al. 2008). The components marked with red dotted-line are equivalent ones which do not actually exist in a practical system whereas others are real indeed. The black dotted-line marked sector is a structure diagram of the CMOS CCD module of the experimental system. Physical significances and reasons of introduction of each virtual component will be presented as below.

### Equivalence of Microcantilevers

As has been stated above, microcantilevers in this imaging system transform absorbed IR into unit's temperature rising and inclination-angle changing. As to processes of IR absorbing and temperature rising, all thermal infrared imaging systems are of similar structure composed of radiation absorption structure and thermal isolation structure as illustrated in Fig. 5. Here the ability of microcantilevers' infrared absorption and temperature transmission represented by  $H$  is defined as temperature variation of microcantilevers caused by per unit temperature variation of the object (the  $H$  of FPA microcantilever pixel will be elaborated in section "The Examination of and the Improvement on the FPA Unit Model"). The inclination-angle variation of microcantilevers caused by per unit temperature rising

will be interpreted as variation of resistance value which is represented by a temperature-sensitive resistor  $R_s$  in this model by analogy in this article, and the resistance value and the resistance-temperature coefficient of this equivalent temperature-sensitive resistance be calculated according to the law of relevance between temperature variation and inclination angle of microcantilevers as illustrated in formula 2:

$$\Delta\theta = 6(a_1 - a_2) \left( \frac{d_1 + d_2}{d_1 K} \right) L * \Delta T \equiv \alpha_{XY} * L * \Delta T \quad (4)$$

As to a common temperature-sensitive resistor, the relevance between its resistance and the temperature is:

$$\Delta R = \alpha_{TCR} * \Delta T \quad (5)$$

Hereinto, the resistance-temperature coefficient of the equivalent temperature-sensitive resistor is represented by  $\alpha_{TCR}$ . According to comparison of formula (4) and formula (5), the thermal deformation of microcantilevers can precisely be equivalent to a linear variation of temperature-sensitive resistor resistance value affected by temperature. Such equivalence will interpret the thermal deformation process of microcantilevers as familiar temperature-resistance variation.

### Equivalence of Optical Readout

The principle of knife-edge filtering noninterference optical system raised by this research can be simplified as follows: cantilevers' inclination causes displacement of energy spectrum of the reflection of incident light at the lens' back focal plane, then the spectrum being modulated by knife-edge filter, thus the inclination angle of cantilevers being finally transformed into light intensity variation upon CCD imaging plane. Through the photoelectric coupling model, this complex process can be simplified into an equivalence as: light intensity of light emitting diode LED gets modulated by voltage between the two ends of temperature-sensitive resistor  $R_s$  via junction field-effect transistor M. The steady current generated by constant flow source A in Fig. 15 is interpreted as the light of an illuminant in readout system by analogy, which transforms resistance variation of equivalent resistor  $R_s$  into voltage variation between its two ends and keeps junction field-effect transistor M in this equivalent system operating at constant current region, thus the current through M under linear control of voltage of  $R_s$  so as to regulate light intensity variation of LED. The photosensitive diode in this equivalent model actually represents the photosensitive cell of CCD in a real system. Photosensitive diode D and light emitting diode LED together composes a photoelectric coupling component. After amplification and quantization, photo voltage induced by photosensitive diode D is transformed into output signal of the system.

The advantage of this circuit analogy model which introduces photoelectric coupling component to represent the relevance between variation of light intensity and voltage in optical readout system is its preciseness in presenting the noise-isolation

feature of optical readout system, and actual components such as LED and photosensitive diode are presented in it, which allow the model to describe its noise.

During the process of light intensity being modulated by the deformation inclination-angle of cantilever, Formula (3) can be deformed as:

$$\Delta I = -\frac{4\pi}{\lambda} * I_s * L_Y * R * \Delta\theta \quad (6)$$

Thereinto,  $\Delta I$  is the received light-intensity variation of photosensitive diode D of CMOS CCD image sensor,  $\Delta\theta$  the variation of cantilever inclination angle, R a constant when the system operates normally,  $I_s$  the length of cantilever, and  $\lambda$  the read-out optical wavelength. The formula of light intensity of light emitting diode LED in the equivalent circuit model given by this article is as follow:

$$\Delta I = \xi * g_m * I_A * \Delta R_S \quad (7)$$

Hereinto,  $\xi$  is the luminous efficiency of light emitting diode,  $g_m$  the low-frequency transconductance of junction field-effect transistor M, and  $I_A$  the read-out current generated by constant flow source. By comparison of formulas (6 and 7) and formulas (4 and 5) in the previous section, we learn that the process of making an equivalence of the inclination-angle variation of cantilever to the resistance variation of thermistor, controlling output light-intensity of LED with voltage between the two ends of thermistor and having it and photosensitive diode D a coupling will precisely reflect detecting process of inclination angle of FPA microcantilevers by optical readout module; thus, this equivalent model can ideally represent modulation process of read-out light intensity by microcantilever inclination angle of the real system.

As light emitting diode LED and photosensitive diode D of CMOS CCD image sensor in this model together compose a photoelectric coupling component (Zhuang 1993), when the relevance between output voltage of CMOS CCD image sensor and light intensity is presented that  $\Delta V = \rho \Delta I$  as  $\rho$  being conversion ratio, the transmission efficiency of photoelectric coupling component can be represented by the product of image sensor's photoelectric conversion efficiency and photosensitive diode's luminous efficiency as  $\rho \cdot \xi$ .

According to the model above, the output response of photoelectric coupling model based on optical readout uncooled infrared imaging system can finally be described as:

$$\begin{aligned} \Delta V &= \rho \cdot \Delta I = \rho \cdot \xi \cdot g_m \cdot I_A \cdot \Delta R_S \\ &= \rho \cdot \xi \cdot g_m \cdot I_A \cdot H \cdot \alpha'_{tcr} \cdot L \cdot \Delta T_S \\ &= \rho \cdot \xi \cdot g_m \cdot I_A \cdot C \cdot \Delta T_S \\ &= \mathfrak{R} \cdot \Delta T \end{aligned} \quad (8)$$

Here,  $C$  is the corresponding sensitivity of a microcantilever.  $\mathfrak{R}$  is the conversion efficiency of the system, i.e., the output voltage change of the system due to the temperature change of the target.

The equivalent circuit succeeds in transforming the complex systematic process into a relatively simple and clear circuit mode. Besides, it gives a formulaic expression of the system performance. In the equivalent, the introduction of the photoelectricity perfectly shows the specific performance of the noise insulation in the optical readout system and has the LED and CCD in the system express their function, which will provide the convenience in the subsequent noise analysis.

## Noise Analysis of the System

The parameter NETD is employed when evaluating the performance of an uncooled IR detector. The concept of NETD can be expressed as the temperature difference between the object and background with a signal-to-noise ratio of unity when detecting an IR source. According to the formula of the system model in the last section, its expression is:

$$NETD = \frac{V_{\text{noise}}}{\mathfrak{R}} = \frac{V_{\text{noise}}}{\Delta V} \Delta T_s \quad (9)$$

where  $V_{\text{noise}}$  is the noise voltage of the system.

From the definition of NETD, the degree of noise is just as important as the sensitivity of a detector concerning the performance of an uncooled IR imaging system. The mathematical expression of the sensitivity of a detecting system was shown above, now we will show the analysis of noise which is divided into FPA noises and optical readout noises.

## FPA Noises

### Temperature Fluctuation Noise

This noise is inevitable in all kinds of thermal IR detectors. It proceeds from the heat exchange between the detector elements and the environment in the micro scale because of the temperature fluctuation of the detector elements (Sarid 1991; Saulson 1990; Kruse 2001). Thus, the kind of noise is usually regarded as the intrinsic noise and limit noise. The noise appears as the random fluctuation of the detector temperature. Prof. Kruse (Kruse 2001) has noted that the random temperature fluctuation can be expressed as:

$$\langle \delta T_f^2 \rangle^{1/2} = \frac{2\sqrt{k_B B T}}{G_T^{1/2} \sqrt{1 + \omega^2 \tau^2}} \quad (10)$$

Here,  $K_B$  is the Boltzmann constant,  $T$  the absolute temperature (K),  $B$  the measured bandwidth,  $G_T$  is conductivity of detector elements,  $\omega$  the operating frequency, and  $\tau$  the thermal time constant.

Given that for the detector the temperature fluctuation is attached to the temperature change due to absorbing radiation, the fluctuation can be converted into the voltage fluctuation of the ultimate output signal as:

$$\langle \delta V_f^2 \rangle^{1/2} = \frac{2\sqrt{k_B B T}}{G_T^{1/2} \sqrt{1 + \omega^2 \tau^2}} \cdot \alpha'_{\text{ter}} \cdot L \cdot \rho \cdot \xi \cdot g_m \cdot \beta \cdot I_A \quad (11)$$

### Background Fluctuation Noise

This noise is generated from the radiant heat exchange between the background and the elements, which can also be applied to all kinds of thermal IR detectors as the temperature change. Its appearance is also the temperature fluctuation of the elements and unrelated to IR radiant signal (Saulson 1990). The temperature fluctuation can be expressed as:

$$\langle \delta T_B^2 \rangle^{1/2} = \frac{2\sqrt{2k_B \sigma_T B A_d (T_D^5 + T_B^5)}}{\eta L} \quad (12)$$

Here,  $\sigma_T$  is the radiation coefficient of FPA at the temperature  $T$ ,  $A_d$  the IR radiation absorber area of FPA,  $T_d$  the FPA temperature,  $T_B$  the background temperature,  $L$  the lever length of equivalent deformation of FPA, and  $\eta$  the IR absorptivity of FPA.

According to the above discussion, the output voltage fluctuation caused by background fluctuation noise is similar to temperature fluctuation noise, which can be expressed as:

$$\langle \delta V_B^2 \rangle^{1/2} = \frac{2\sqrt{2k_B \sigma_T B A_d (T_D^5 + T_B^5)}}{\eta L} \cdot \alpha'_{\text{ter}} \cdot L \cdot \rho \cdot \xi \cdot g_m \cdot \beta \cdot I_A \quad (13)$$

### Thermo-Mechanical Noise

In the indicated optical readout IR imaging system based on MEMS microcantilever, a specific noise has to be introduced due to the specific nature of mechanical structure of the cantilever, i.e., thermo-mechanical noise. Because FPA cantilever brings in an oscillator when absorbing energy and deforming, there is still heat exchange between the detector elements and the environment when the cantilever is in a state of thermal equilibrium. The persistent heat exchange process is defined by the fluctuation-dissipation theorem. Each portion of heat energy value is  $k_B T$  in the heat exchange process which results in the spontaneous vibration of the cantilever. Prof. Sarid calls this noise the thermo-inducing lever noise (Saulson 1990). In the equivalent model, the spontaneous vibration of the cantilever can be regarded as the resistance fluctuation of thermistor  $R_s$ .

According to Prof. Sarid's theory, the volume of this noise can be expressed as:

$$\langle \delta R_{TM}^2 \rangle^{1/2} = \sqrt{\frac{4k_B T B}{Qk} \frac{\omega_0^3}{(\omega_0^2 - \omega^2)^2 + \omega^2 \omega_0^2 / Q^2}} \quad (14)$$

Here,  $Q$  is the quality factor of the cantilever structure, which is measured via experiment and defined as the ratio of resonance frequency to resonance peak width.  $\omega_0$  is the resonance frequency of the cantilever structure, and  $k$  the elasticity coefficient of the cantilever structure.

When the operating frequency of the microcantilever is much less than the resonance frequency, this formula can be expressed as:

$$\langle \delta R_{TM}^2 \rangle^{1/2} = \sqrt{\frac{4k_B T B}{Q k \omega_0}} \quad (15)$$

When the operating frequency is equal to the resonance frequency, the formula (14) can be simplified as:

$$\langle \delta R_{TM}^2 \rangle^{1/2} = \sqrt{\frac{4k_B T B Q}{k \omega_0}} \quad (16)$$

Because the operating frequency of the uncooled IR imaging system is 30 Hz and the natural frequency of the microcantilever in a vacuum condition as a rule, substituting formula (15) into formula (8) can express its effect on output voltage of the system.

$$\langle \delta V_{TM}^2 \rangle^{1/2} = \sqrt{\frac{4k_B T B}{Q k \omega_0}} \cdot \rho \cdot \xi \cdot g_m \cdot \beta \cdot I_A \quad (17)$$

The above three kinds of noises have a same trait that the volume of the noise depends on the structure and parameter of detector elements and outputted magnifying through readout magnifying system. The three kinds of noises are directly reflected on the equivalent thermosensitivity  $R_s$  in the equivalent model, including the temperature fluctuation of thermosensitive resistance and the resistance value.

Unlike the electrical readout system, the optical readout system, as a noncontact data readout system, would not introduce noise into sensitive elements when reading out signals. In other words, the volume of the noise of the equivalent resistance  $R_s$  is only concerned with its intrinsic design parameter and technical art, not with the peripheral readout system. Based on this trait, we call the three kinds of noises generated from FPA the internal noises of the system, the noises introduced by optical system, and the external noises of the system.

### Optical Readout Noises

As indicated in the equivalent circuits, all the circuits exist in the objective elements of the optical readout system except the virtual elements encircled by the dotted lines, so they can all introduce noises. Here, the power source VDD of LEDs can introduce the power source fluctuation noise; LEDs can introduce the light intensity



fluctuation noise; CCD, composed of LEDs and light sensitive diodes, can introduce shot noise; the follow-up readout circuits can introduce a range of noises, such as current noise, 1/f noise, and quantization noise.

### LED Fluctuation Noise

In this section, the LEDs are used as the noninterference light source. The virtual structure of LED is a PN junction of semiconductors. Its operating principle is that the PN junction with a forward voltage can bring in a small number of current carriers and then the radiative recombination of current carriers produces photons which have the LED radiate. It is usually thought that only when impurities and line defects of heavy metals are guided to the deep level in the forbidden band of semiconducting elements can obvious G-R noise be observed (Zhuang 1993). 1/f noise is a universal phenomenon due to the fluctuation of particles' collective motion and the reaction of internal trait of the system. Thus, in this section 1/f noise is:

$$\langle \delta I_{\text{LED}}^2 \rangle^{1/2} = \sqrt{\frac{\alpha q I_D}{f \tau_e}} \quad (18)$$

Here,  $\alpha$  is the Hooge's constant,  $q$  the electron charge,  $I_D$  the input current intensity,  $f$  the frequency, and  $\tau_e$  the life span of the particles. It should be noted that the power stability of the light source appears as the fluctuation of the input current  $I_D$ . The ultimate output effect on the system by this noise can be expressed as:

$$\langle \delta V_{\text{LED}}^2 \rangle^{1/2} = \sqrt{\frac{\alpha q I_D}{f \tau_e}} \cdot \rho \cdot \beta \quad (19)$$

In general, this noise affects the system together with the fluctuation noise of the power source, which makes its effect difficult to be calculated. Therefore, the intensity of this noise is usually got in an experimental way. The 1/f noise of LED is small. If the front LED of GH5520Z, a domestic military CCD, is taken as an example, its fluctuation of light intensity is only  $10^{-9}$  in the condition of 16 mA direct current (Jinhu et al. 2006). So, the volume of the LED fluctuation noise mainly depends on the ripple coefficient of its driving power source.

### Shot Noise

This noise is generated from the fluctuation of the number of photons shot onto the CCD board under the natural statistical law when CCD absorbs the detecting light (Saulson 1990).

The output voltage change of CCD caused by this noise can be expressed as:

$$\langle \delta v_s^2 \rangle^{1/2} = \sqrt{\frac{q C B I_l}{2}} \quad (20)$$

Here,  $q$  is the electron charge,  $C$  the sensitivity of the light-detecting element, and  $I_l$  the readout light intensity. Its effect on the ultimate output signal can be expressed as:

$$\langle \delta V_s^2 \rangle^{1/2} = \sqrt{\frac{qCB I_l}{2}} \cdot \beta \quad (21)$$

This noise appears along with the above LED fluctuation noise, which makes it difficult to differentiate between them.

### Current Noise

A range of noises, such as the current noise,  $1/f$  noise, measured noise (Zhuang 1993), inevitably exist in all the circuits of the CMOS CCD part in the readout system. They simultaneously appear in the irregular fluctuation of output voltage, which is expressed as  $\langle \delta V_{\text{current}}^2 \rangle^{1/2}$ . The commercial CCD in actual system is used, so the noise volume completely depends on its performance.

From the above introduction, it can be concluded that the external noise volume inputted by the optical readout system is only concerned with the performance of the constant current power supply, LED, CCD, etc., not with performance of the key element FPA.

### Effects of the Noises on System Performance

If substituting the above noises into the formula (9), the respective effects of all the noises on system performance can be got.

#### Effects of the Internal Noises

If substituting the formulas (11), (13), and (17) into the formula (9), the effects of the temperature fluctuation noise can be got, the background fluctuation noise and the thermo-mechanical noise respectively on system performance, which can be expressed as:

$$NETD_f = \frac{2\sqrt{k_B B T}}{H \cdot G_T^{1/2} \sqrt{1 + \omega^2 \tau^2}} \quad (22)$$

$$NETD_B = \frac{2\sqrt{2k_B \sigma_T B A_d (T_D^5 + T_B^5)}}{H \cdot \eta \cdot L} \quad (23)$$

$$NETD_{TM} = \sqrt{\frac{4k_B T B}{Q k \omega_0}} \cdot \frac{1}{H \cdot \alpha'_{lcr} \cdot L} \quad (24)$$

In terms of the expressions of the effects of the external noises on system performance, we can get that unlike the internal noises, the effects of all the internal noises can be effectively repressed via the improvement of the performance of FPA and the reaction sensitivity of the readout system.

## Experimental Measuring of the Effects of System Noise

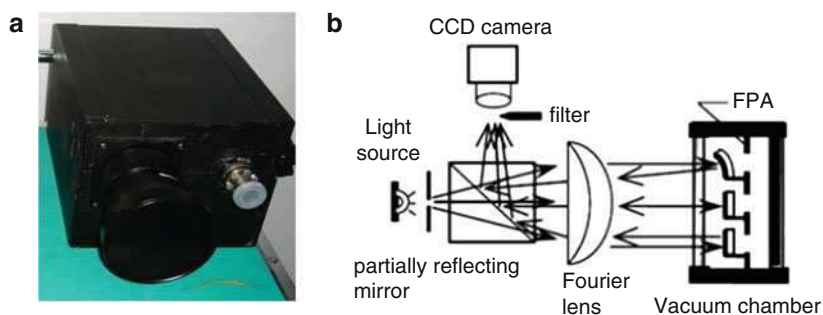
The above model demonstrates the sources of all kinds of noises and their effects on the system. In order to confirm the analysis of the model and test the volume of the noises in the system, the analysis via experiment is continued.

Based on the optical path (Fig. 12), an IR imaging device is fabricated as shown in Fig. 16. In the device, the longest relative aperture of the IR lens is  $4/3$  and the focal length is 50 mm. The image-capturing device is a CCD produced by Canadian Dalsa Corporation, whose type is DALSAADS-11-01 M30 and resolution is 12-bit. Using the commercial CCD, the gray levels of the output image signals of this CCD are regarded as quantized output signals.

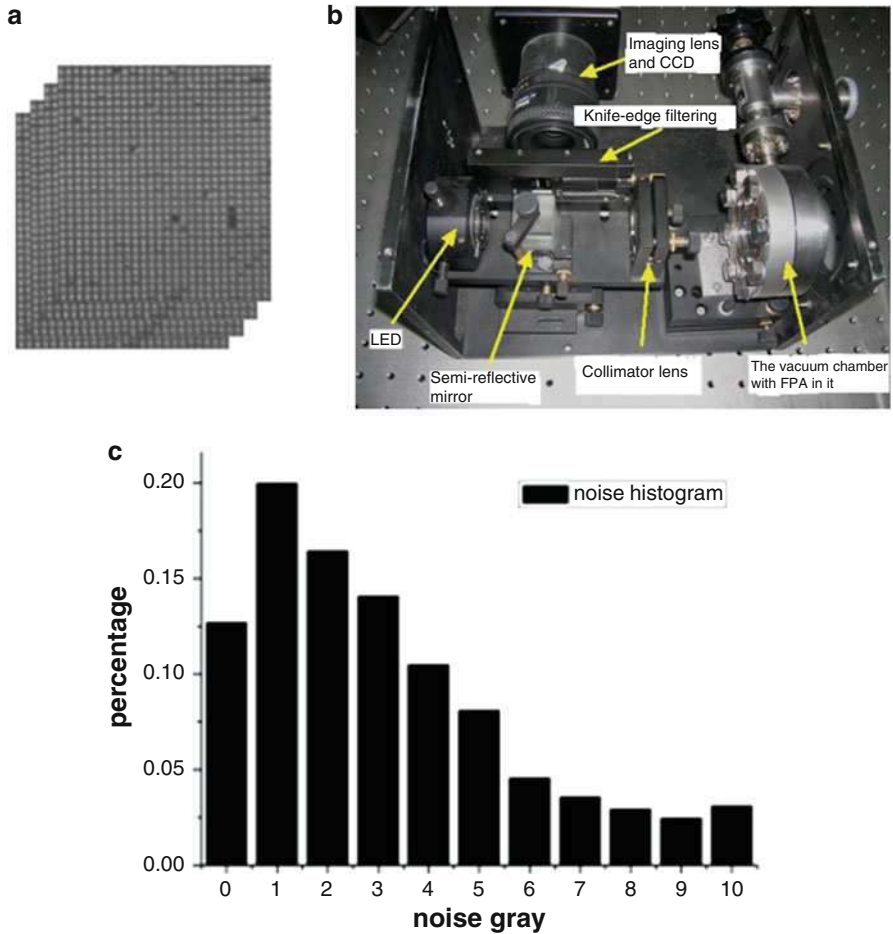
First, the noise of the imaging system equipped with microcantilever FPA is tested. The following are the testing approaches. Figure 17a indicates that 100 images are captured without IR source. Figure 17b indicates the result. By counting the gray of the captured image, a noise histogram of the system is obtained (Fig. 17c). Because the noise is distributed around the gray value of the maximum probability, it can be regarded as the system noise.

Based on the trait of noise insulation of the optical readout system, the noise test by substituting a mirror for FPA in the system (Fig. 18) is continued. The mirror cannot introduce the internal noises in the system, so this measured noise can be the external noises introduced by optical readout system. In the experiment, a mirror for the FPA is substituted while other parameters are not changed. The contrast between two noise histograms of the test results indicates that the volume of the two noises corresponds basically.

Table 1 indicates the parameters and performance of the FPA in the contrast experiment. By substituting the parameters into formulas (11), (13), and (17), the estimated value of the internal noises introduced by the FPA can be obtained, which is expressed as gray level (Table 2). The estimated results demonstrate that the internal noises introduced by FPA are much smaller than the external noises, which correspond to the result of the contrast experiment.

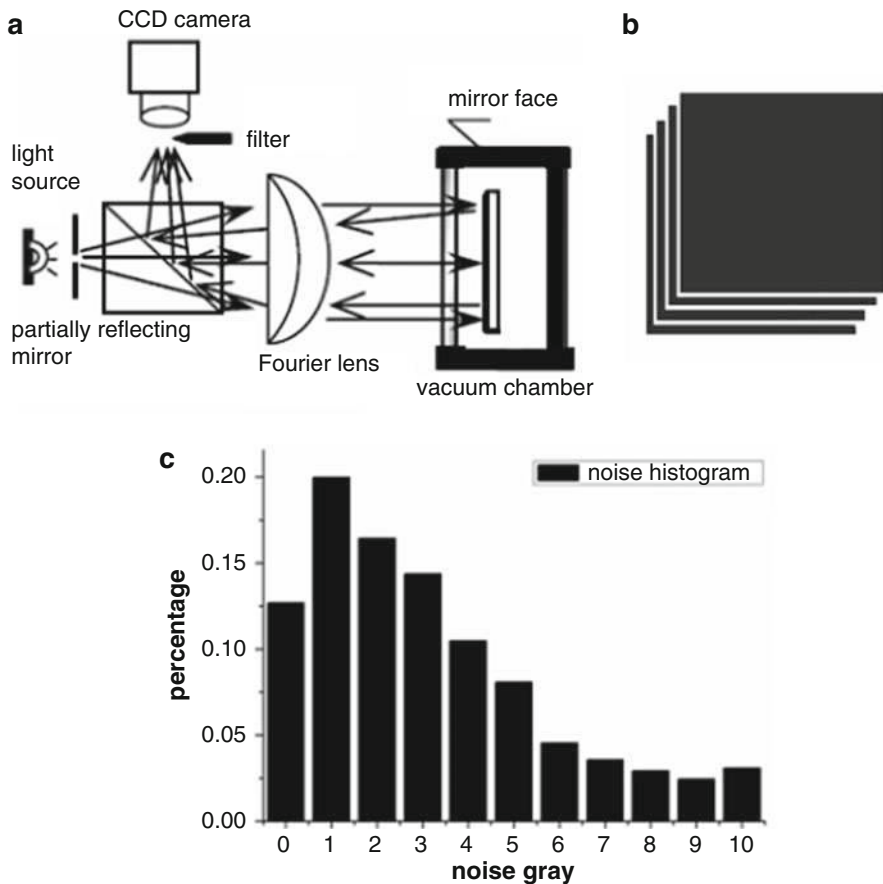


**Fig. 16** IR imaging device. (a) Fabricated IR imaging device. (b) System components



**Fig. 17** Noise test of the system with FPA. (a) Approach of measuring system noises with FPA. (b) FPA images captured by CCD. (c) Noise histogram of gray level of the system with FPA

The set of test results demonstrates that the noises of the uncooled IR imaging system using optical readout method mainly originate from the optical readout system, and the volume is a constant 8 gray levels. In addition, the noises introduced by FPA are so small that they can be neglected. According to the definition of NETD of the IR imaging system, if it needs to improve NETD of the uncooled IR imaging system using optical readout method without changing the optical readout system, the reaction sensitivity of the system should be improved to the best. If it needs to improve the reaction sensitivity of the system, the most important method is to improve the performance of the key element, FPA. Thus, the subsequent content goes around the improvement of the FPA performance.



**Fig. 18** Noise test of the system with the mirror. (a) Approach of measuring system noises with the mirror. (b) Mirror images captured by CCD. (c) Noise histogram of gray level of the system with the mirror

**Table 1** The parameters of FPA in the experiment ( $\mu\text{m}$ )

Unit size	Reflector size	Leg length	Leg width	Au film thickness	SiNx film thickness
120 × 120	110 × 63	330	2	0.2	1.2

**Table 2** The internal noises introduced by FPA (gray level)

Temperature fluctuation noises	Background fluctuation noises	Thermo-mechanical internal noises	RMS
0.223	0.035	0.197	0.274

## Section Summary

The model of the uncooled IR imaging system using optomechanical method, which is on the systematic scale, is fabricated in this section. At first, the operating fundamentals are discussed in depth of the uncooled IR imaging system using optomechanical method, and an equivalent circuit model is fabricated where the CCD indicates the noise-insulation trait of the optical readout system. Next, based on the model, the noise volume of all parts and their effects on the system is analyzed, combined with the specific trait of FPA with cantilever structure. The analysis indicates that the effect of the internal noises introduced by the existing FPA on NETD is just on the mK scale, much smaller than the noises of the optical readout system. At last, the analysis of the noise source in the model is demonstrated by conducting an experiment. It is confirmed that the main noise source of the existing uncooled IR imaging system using optomechanical method originates from the system of optical path, not concerned with FPA.

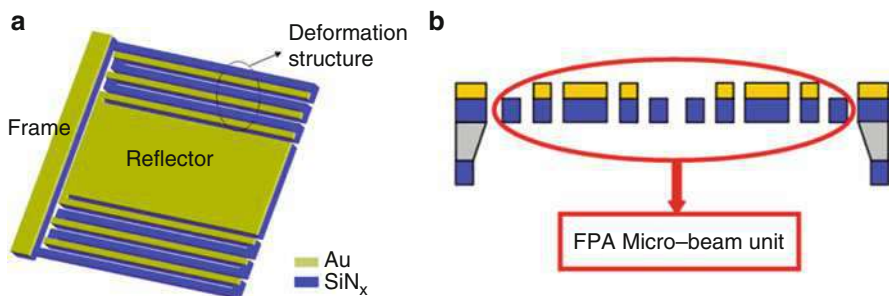
---

## The Examination of and the Improvement on the FPA Unit Model

The focal plane array is the key element of the uncooled IR imaging system, whose performance decides the system performance. During the previous work, the substrate-free structure is put forward, which succeeds in solving the problem of IR reflection caused by the structure with Si substrate on the sacrificial layer. The FPA uses the multifold interval metalized legs, which effectively improve the temperature reaction sensitivity of the microcantilever and obtain a relatively clear human IR image at the end. Although good images were obtained in previous work, its theory model is still imperfect. In this section, the previous structure and theory model is carefully analyzed. The special problems of the substrate-free FPA are found through experiments, and the model is corrected.

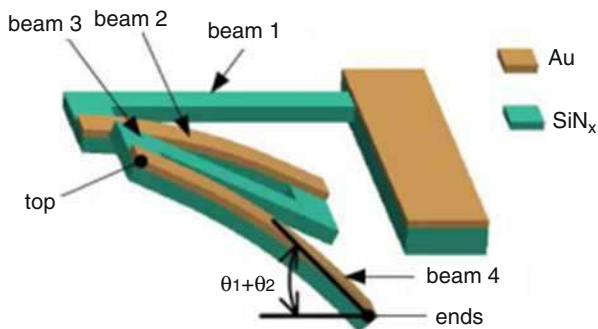
### Structure of the Single-Layer FPA

The substrate-free FPA with multifold interval metalized legs is designed and fabricated, as indicated in Fig. 19. The bi-material leg of large area between the fold legs of both sides is infrared-absorbing part (reflector). Its  $\text{SiN}_x$  layer can effectively absorb the IR radiation on it and transform the radiation into temperature change. Its Au layer can reflect the readout light to have the optical readout system obtain signals. The magnifying principle of the multifold IML structure is demonstrated in Fig. 20, where the twofold structure is taken as an example. Here, the thermal isolation legs (Leg 1 and 3) are single-material legs, which cannot produce thermal tilt angle when temperature changes. The deformation legs (Leg 2 and 4) are bi-material legs, which can produce thermal deformation when the microcantilever absorbs IR radiation. At that time, Leg 2 is deformed and the tilt angle of its end part relative to Leg 1 is  $\theta_1$ . Because Leg 3 does not produce a tilt angle, the tilt angle of



**Fig. 19** Structure of the previous FPA. (a) Stereogram of the previous FPA. (b) Cross-section diagram of the previous FPA

**Fig. 20** Multifold IML magnification principle



the end part of Leg 3 relative to Leg 1 is also  $\theta_1$ . The thermal tilt angle of Leg 4 relative to Leg 3 is  $\theta_2$ , which means that the tilt angle of Leg 4 relative to Leg 1 is  $\theta_1 + \theta_2$ . With the help of this structure, FPA can produce more tilt angle change under the same temperature change.

## Early Model of the Substrate-Free Single-Layer FPA

In the earlier work, the model which can present the reaction performance of a substrate-free FPA had been fabricated, according to the model of an optical readout uncooled FPA fabricated by Prof. M. Mao at Berkeley (Mao et al. 1999a, b; Norton et al. 2000; Miao et al. 2006). The model divides the operating process of FPA into two sections: the thermal one which absorbs IR radiation and produce temperature change, and the thermo-mechanical one which produces deformation due to temperature change.

### Thermal Model of the Microcantilever

Assuming that the radiation intensity of a black-body radiation source is  $P$  under the temperature of  $T_s$ , the absorbed part by the microcantilever can be expressed as (Liu 1992; Yilin 1987):

$$q = \frac{A_d \tau \varepsilon \pi}{4F^2} P \quad (25)$$

where  $A_d$  is the effective IR-absorbing area of the microcantilever units,  $\varepsilon$  the IR-absorbing coefficient of the microcantilever units,  $\tau$  the transmittance of the IR imaging system ( $=0.9$ ), and  $F$  the f-number of the lens.

Under the original temperature of  $T_0$  ( $\sim 300$  K), the temperature  $T_c$  of the microcantilever unit after absorbing IR radiation meets the equation as (Yang 1987):

$$\rho V c \frac{dT_c}{dt} = -G_{\text{total}}(T_c - T_0) + q \quad (26)$$

where  $\rho$ ,  $V$ , and  $c$  are respectively the density, volume, and heat capacity of the microcantilever,  $T_0$  is the environment temperature, and  $G_{\text{total}}$  is the total thermal conductance of the microcantilever. There are mainly three ways of heat exchange between the microcantilever and the environment, i.e., by way of thermal radiation ( $G_{\text{rad}}$ ), by way of heat conduction of the legs ( $G_{\text{leg}}$ ), and by way of heat conduction and heat convection of air ( $G_{\text{air cond}}$  and  $G_{\text{air cov}}$ ). Thus, the total thermal conductance can be expressed as:

$$G_{\text{total}} = G_{\text{rad}} + G_{\text{leg}} + G_{\text{air, cond}} + G_{\text{air, cov}} \quad (27)$$

When the left side of the formula (26) equals 0, the temperature reaches the balance, so that the temperature rise of the microcantilever is obtained as:

$$T_c - T_0 = \frac{q}{G_{\text{total}}} \Rightarrow \delta T_c = \frac{\delta q}{G_{\text{total}}} \quad (28)$$

In terms of the formulas (25) and (26), the energy conversion efficiency can be obtained as:

$$H = \frac{\delta T_c}{\delta T_s} = \frac{\delta T_c}{\delta q} \cdot \frac{\delta q}{\delta P} \cdot \frac{dP}{dT_s} = \frac{1}{G_{\text{total}}} \cdot \frac{A_d \tau \varepsilon \pi}{4F^2} \cdot \frac{dP}{dT_s} \quad (29)$$

Within the wave band of 8~14  $\mu\text{m}$ , the radiation rate of a unit temperature of a blackbody under the temperature of 300 K ( $dP/dT_s$ ) equals  $0.63 \text{ W m}^{-2} \text{ K}^{-1} \text{ sr}^{-1}$ . According to the above formulas, the key issue of thermal design is that one of the thermal conductance  $G_{\text{total}}$  can be obtained.

First, the thermal conductance of the microcantilever unit's radiation is discussed. According to Kirchhoff's laws, the net radiation power of the microcantilever unit's surface can be obtained as:

$$q_{\text{rad}} = \sigma A_{\text{pixel}} (\varepsilon_{\text{SiNx}} + \varepsilon_{\text{Au}}) (T_c^4 - T_0^4) \quad (30)$$



where  $\sigma$  is the Stefan-Boltzmann constant,  $A_{\text{pixel}}$  the area of the microcantilever unit. Because the temperature  $T_c$  of the microcantilever unit is similar to the original temperature (room temperature), the radiation thermal conductance can be approximately expressed as:

$$G_{\text{rad}} = \frac{\delta q_{\text{rad}}}{\delta T_C} = 4\sigma A_{\text{pixel}}(\varepsilon_{\text{SiN}_x} + \varepsilon_{\text{Au}})T_0^3 \quad (31)$$

In terms of the formula (31), if the material and the size of the microcantilever unit is definite, the radiation thermal conductance is definite basically, which decides the energy conversion efficiency in the end.

Then, the heat exchange with the environment by way of the leg conduction ( $G_{\text{leg}}$ ) is discussed. The thermal conductance of thermal isolation legs is:

$$G_{\text{leg}} = \frac{A_{\text{leg}}k_{\text{leg}}}{L_{\text{leg}}} \quad (32)$$

Here,  $k$  is the thermal conductivity coefficient of the material and  $A$  the cross-sectional area of the material. The thermal conductance of bi-material legs can be calculated via the following formula:

$$G_{\text{leg}} = \left( \frac{L_{\text{leg}}}{\sum_i k_i A_i} \right)^{-1} \quad (33)$$

The subscript  $i$  represents the  $i$  material. Thus, as to the FPA structure put forward, the thermal conductance through the legs can be expressed as:

$$G_{\text{leg}} = 2 \left[ \frac{L}{k_{\text{SiN}_x} A_{\text{SiN}_x}} + \frac{L}{k_{\text{Au}} A_{\text{Au}}} \right]^{-1} \quad (34)$$

where,  $L$  is the length of the single-fold leg.

Next, the effect of air is discussed. The condition of air's convection can be confirmed by the Rayleigh criterion. The Rayleigh number of the air gap ( $D$  represents its thickness) between the microcantilever unit and the wall of the vacuum chamber is:

$$Ra = \frac{g(T_C - T_0)D^3}{\alpha \nu} \quad (35)$$

Here,  $g$  is the acceleration due to gravity,  $D$  approximately 2~20 mm,  $\alpha$  the thermal diffusivity of the air, and  $\nu$  the kinematic viscosity of the air. At room temperature and pressure,  $\alpha = 22.5 \times 10^{-6} \text{m}^2 \text{s}^{-1}$ ,  $\nu = 15.89 \times 10^{-6} \text{m}^2 \text{s}^{-1}$ . Here, the Rayleigh number is estimated. Assuming that the temperature change  $\Delta T = T_c - T_0 = 1 \text{K}$

and  $D = 20$  mm, then  $Ra = 731$ . Because the condition of starting natural convection is that the Rayleigh number is greater than 1708, the heat convection of air cannot take place. The thermal conductance of air can be expressed as:

$$G_{\text{air}} = \frac{k_{\text{air}}A_{\text{pixel}}}{D} \quad (36)$$

Here,  $k_{\text{air}}$  is the thermal conductivity coefficient of the air. At room pressure ( $P_0 = 1.0 \times 10^5$  Pa), its value  $k_{\text{air},0}$  is  $26.3 \times 10^{-3} \text{Wm}^{-1} \text{K}^{-1}$ . If estimating in terms of the unit of  $100 \mu\text{m} \times 100 \mu\text{m}$  ( $D = 20$  mm), the value of  $G_{\text{air}} = 1.3 \times 10^{-8} \text{WK}^{-1}$ . From the following analysis, the value of  $G_{\text{rad}}$  and  $G_{\text{leg}}$  is approximately  $10^{-8} \text{WK}^{-1}$ . Thus, the heat conduction of the air at room pressure can affect the total thermal conductance of the microcantilever units. Even so, an experiment is still performed with FPA put into room pressure and successfully gets the IR image of an object whose temperature is greater than  $120^\circ\text{C}$ . The result was reported in the literature (Dong et al. 2006).

To eliminate the air's effect, the FPA is put into a vacuum chamber. At different pressures, the thermal conductivity coefficient of the air meets the relation as:

$$k_{\text{air}} = k_{\text{air},0} \frac{Pl_m}{P_0l_{m,0}} \quad (37)$$

where,  $l_m$  is the mean free path of the air molecules. At room temperature, the mean free path of the air molecules and the pressure (using Pa as the unit) approximately meet the relation as (Liu 1991):

$$l_m = 5/P \quad (\text{mm}) \quad (38)$$

At an atmospheric pressure and the vacuum degree of 0.1 Pa, the mean free path is respectively 50 nm and 50 mm. On the other hand, the distance between the FPA microcantilever units and the chamber wall is approximately 5 mm ( $< 50$  mm). Thus, at 0.1 Pa, the effective mean free path of the air molecules is  $D$  and formula (49) can be rewritten as:

$$k_{\text{air}} = k_{\text{air},0} \frac{PD}{P_0l_{m,0}} \quad (39)$$

Here,  $k_{\text{air}} = 2.63 \times 10^{-3} \text{Wm}^{-1} \text{K}^{-1}$ ,  $G_{\text{air}} = 1.3 \times 10^{-9} \text{WK}^{-1}$ . Thus, the heat conductance of the air is negligible. Then, the formula (41) can be rewritten as:

$$H = \frac{1}{G_{\text{leg}} + G_{\text{rad}}} \cdot \frac{A_d \tau \epsilon \pi}{4F^2} \cdot \frac{dP}{dT_S} \quad (40)$$

From the above analysis, the FPA's total thermal conductance  $G_{\text{total}}$  can be got, the area of the IR absorber  $A_d$  and the IR-absorbing coefficient  $\epsilon$  is the key to the thermal performance of the FPA microcantilever units.

### Thermal-Mechanical Model of the Microcantilever Units

The thermal conductivity coefficient of Au is two orders of magnitude greater than that of SiN<sub>x</sub>. According to the formula (34), the contrast can be got with the thermal isolation legs of SiN<sub>x</sub>, the thermal resistance of the bi-material deformation legs is negligible. Thus, it can make a sensible assumption that temperature is evenly distributed on the bi-material legs and it is linearly distributed on the thermal isolation legs, which can be confirmed by the result of finite element modeling shown in Fig. 21.

Figure 21 shows the modeling conducted by the software ANSYS and the Au/SiN<sub>x</sub> bi-material deformation legs and thermal isolation legs in the microcantilever unit are both threefold. When 1 K temperature change takes place on the reflector, the temperature gradient takes place on the threefold thermal isolation legs and is distributed linearly, while the temperature change of threefold bi-material deformation legs (beginning from the one next to the reflector) is respectively 1 K, 2/3 K, and 1/3 K.

The array element can produce mechanical deformation after the radiation of the IR heat source. If the array element is regarded as a simple vibration system and the model shown in Fig. 22 is regarded as a zero-input system, then

$$M \frac{d^2x}{dt^2} + c \frac{dx}{dt} + kx = 0 \quad (41)$$

The RCL-modeled circuit model equivalent to the mechanical vibration model (shown in Fig. 22b) can be expressed as:

$$L \frac{d^2q}{dt^2} + R \frac{dq}{dt} + \frac{1}{C} q = 0 \quad (42)$$

The above two types are of the same form, the same with the mechanical vibration model RCL simulation of equivalent circuit model, they also have the effect of clear physical meaning and are easy to calculate. Mass M, damping coefficient c, elastic coefficient k, and speed of mechanical vibration model are corresponding to inductor L, resistance R, electricity play and electric current of RCL analog circuit model. Considering the FPA is encapsulated in a vacuum, the damping is ignored, that is to say, c = 0, R = 0. The inherent vibrational frequency of array after simplification is:

$$f_v = \frac{1}{2\pi} \sqrt{\frac{2k}{M}} \quad (43)$$

Among them, the elasticity coefficient k is defined as the load of beam and ratio of deformation. The load subjected by the free end of simplified array (Fig. 22a) is W = Mg/2. The calculated inherent vibrational frequency of array is 4211 Hz. The ratio of this beam's gravitational and defamation is 3.43 × 10<sup>-2</sup> N/m by finite

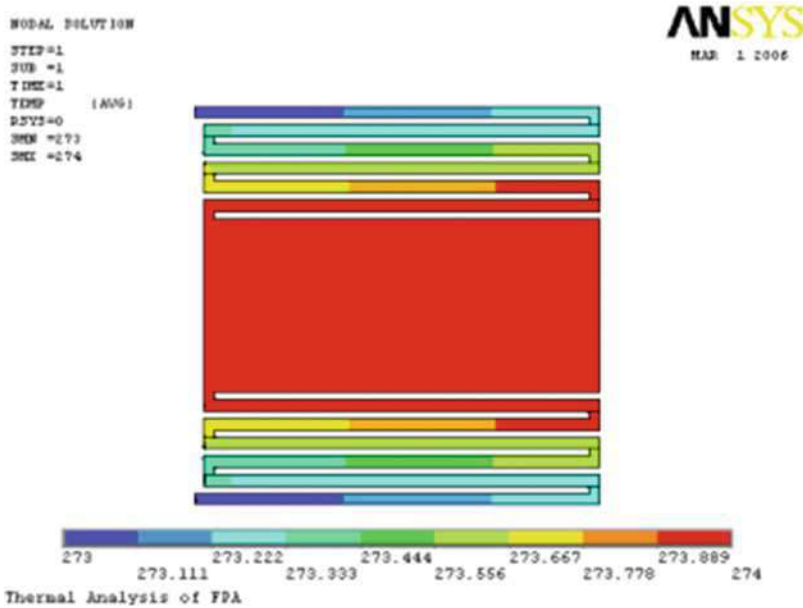


Fig. 21 Modeled temperature distribution on the legs of the microcantilever unit

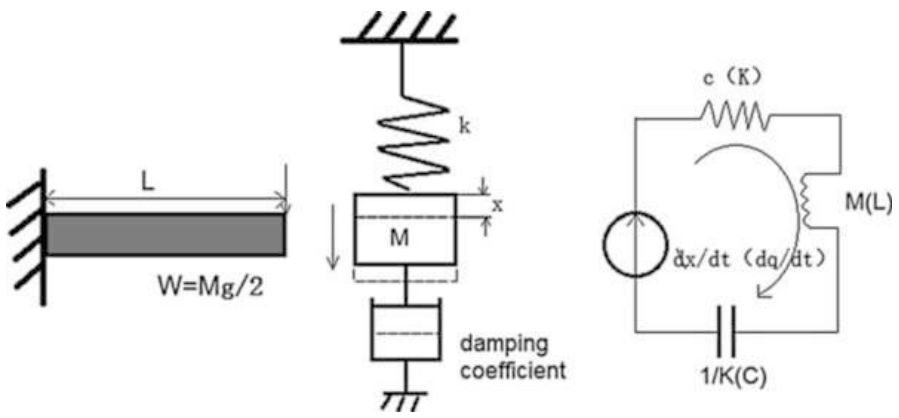


Fig. 22 (a) Equivalent cross-section diagram. (b) Equivalent mechanical vibration model. (c) RCL-modeled circuit model equivalent to the mechanical vibration

element analogy. The inherent vibrational frequency of array is simulated by finite element method, and displacement assumption of array's fixed end is set 0 as boundary condition, the load which is inertial acceleration  $g$  is applied in the whole array, then the inherent vibrational frequency of array by simulation is 4303 Hz. Calculated value and the simulation value is very close.

Based on the above hypothesis, the angle  $\theta$  of microcantilever unit which has multiple inflexion deformation structure by formula is (2):

$$\begin{aligned}\theta &= 6(\alpha_{\text{Au}} - \alpha_{\text{SiNx}}) \left( \frac{t_1 + t_2}{t_1^2 K} \right) L \Delta T \left( \frac{1}{m} + \frac{2}{m} + \dots + \frac{m}{m} \right) \\ &= 6(\alpha_{\text{Au}} - \alpha_{\text{SiNx}}) \left( \frac{t_1 + t_2}{t_1^2 K} \right) L \frac{m+1}{2} \Delta T\end{aligned}\quad (44)$$

In the formula,  $m$  represents inflexion times of multiple inflexion structure.

By changing formula (44), the thermo-mechanical response of microcantilever unit is:

$$S_T = \frac{\theta}{\Delta T} = 6(\alpha_{\text{Au}} - \alpha_{\text{SiNx}}) \left( \frac{t_1 + t_2}{t_1^2 K} \right) \cdot \frac{m+1}{2} \cdot L = \alpha'_{\text{tr}} L \quad (45)$$

From the above formula, it can be seen that choosing a bi-material which has large difference in coefficient of thermal expansion benefits the maximization of  $S_T$ . When the material is selected, optimization design of thermo-mechanical response just depends on structure design and dimension parameters. The longer the length  $L$  or the smaller the  $t_1$  of the deflection beam, the greater the thermo-mechanical response. What's more,  $S_T^\infty (n+1)/k$ .

Due to the structural parameter  $K$  is also a function of the thickness ratio  $n$ , thus optimizing thickness ratio  $n$  can make the best of the value of  $S_T$ .

### The Efficiency of System Transformation $\mathfrak{R}$

The relationship between gray variation  $\Delta I$  of infrared thermal imaging and temperature rise  $\Delta T_s$ , which is responding sensitivity of system  $\mathfrak{R}$ , is:

$$\mathfrak{R} = \frac{\Delta I}{\Delta T_s} = \frac{\Delta I}{\Delta \theta} \cdot \frac{\Delta \theta}{\Delta T_C} \cdot \frac{\Delta T_C}{\Delta T_s} = g_m \cdot I_A \cdot \xi \cdot \beta \cdot \rho \cdot S_T \cdot H = D \cdot S_T \cdot H \quad (46)$$

Among them,  $D$  represents the ability of system's changing cantilever deformation to final quantized gray variation, which is the readout sensitivity of opticator.

### Thermal Response Time of Micro Beam Unit

The answer of differential equation (47):

$$T_C = T_0 + \frac{q}{G_{\text{rad}} + G_{\text{leg}}} \left\{ 1 - \exp\left(-\frac{G_{\text{rad}} + G_{\text{leg}}}{\rho V c} t\right) \right\} \quad (47)$$

So the time required for thermal equilibrium—thermal response time is:

$$\tau = \frac{\rho V c}{G_{\text{rad}} + G_{\text{leg}}} \quad (48)$$

From the above formula, it can be seen that the larger the thermal response time of micro beam unit.

$$\tau \propto 1 / (G_{\text{rad}} + G_{\text{leg}}) \circ (G_{\text{rad}} + G_{\text{leg}}), \quad (49)$$

the more beneficial to get a small thermal response time, which is beneficial to improve sampling frequency of dynamic image. On the other hand, the enlargement of  $(G_{\text{rad}} + G_{\text{leg}})$  reduces the thermal transition efficiency  $H$  of micro beam unit, thus affecting the detection performance of the micro beam unit.

## The Model Validation

In order to verify the validity of the above model, we designed FPA composed of four different cells, the whole picture is shown in Fig. 23. Microphotograph of each area of FPA is shown in Fig. 24. The amount of microcantilever array of each area is  $100 \times 100$ , and each size of micro beam unit in every area is different; detailed sizes are shown in form 25. Form 26 lists corresponding performance parameter which is based on above models in each area. Putting this FPA in the optical readout system which is introduced in section “[Modeling and Analysis of MEMS Based Optical-Readable Thermal Imaging Technology](#),” we can know parameters like responding sensitivity or responding time which are corresponded to four differently designed units. Figure 25 shows infrared image of human’s hand which comes from this FPA and infrared lens whose largest relative aperture is 4/3 and focal distance is 50 mm.

### Validation of Response Sensitivity $\mathfrak{R}$

In order to evaluate sensitivity  $\mathfrak{R}$  of four different regions, we design a temperature adjustable object as an infrared source (Fig. 26), and the temperature adjustment accuracy is 0.01 K. When the temperature of infrared source approaches to room temperature, it changes at a certain temperature interval (such as 2 °C), and author collected 30 thermographs at each temperature point. The deformation of FPA micro beam unit at each temperature can be represented by average gray degree of 30 pictures of CCD pixel which is corresponded with reflector. The grayscale response of micro beam unit can be calculated by gray variation of CCD pixel which is corresponded with reflector that causes the rise of temperature of target.

We did a test to FPA of these four areas at the condition that room temperature is 22 °C and the pressure of vacuum chamber is 0.5 Pa. In the test, the temperature range of measuring target is 22 °C~42 °C. And we made a measure every 2 °C, then each area’s responding sensitivity can be known one by one after averaging these 11 temperature points’ response. Figure 27 shows the response of four areas in the target infrared source at 42 °C.

Through the analysis of these 11 temperature points’ response, we got the response’s histogram of four areas. Figures 28, 29, 30, and 31 are sensitivity’s histogram of area1 to area4. From the figure, it can be seen that the largest response

**Fig. 23** The whole picture of FPA test lens



of four areas are: 18 gray/K, 44 gray/K, 108 gray/K, 48 gray/K. Set area 3's response as benchmark, the responses are 0.167, 0.407, 1, 0.444. This result is close to the result of Fig. 20 which is resulted from model (0.11, 0.33, 1, 0.51). So the description of the property of FPA response sensitivity is reliable.

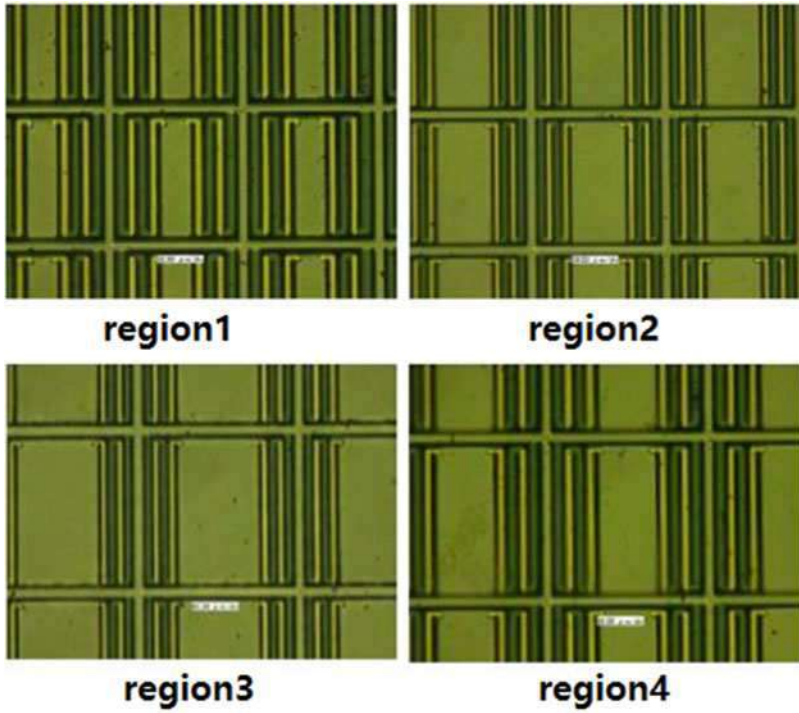
### Validation of Response Time

Verification tests of FPA's thermal response time is at the condition of vacuum ( $1 \times 10^{-2}$ Pa) and temperature of 24.8 °C. Experimental facility is shown in Fig. 32, temperature-controlled infrared target is 1.5 M away from imaging system, and a baffle is set between them. From experience, we set the temperature of temperature-controlled infrared target at 42 °C. When the temperature of infrared source is being stable, we put a baffle between infrared target and imaging system, at the same time, recording the gray change of infrared image.

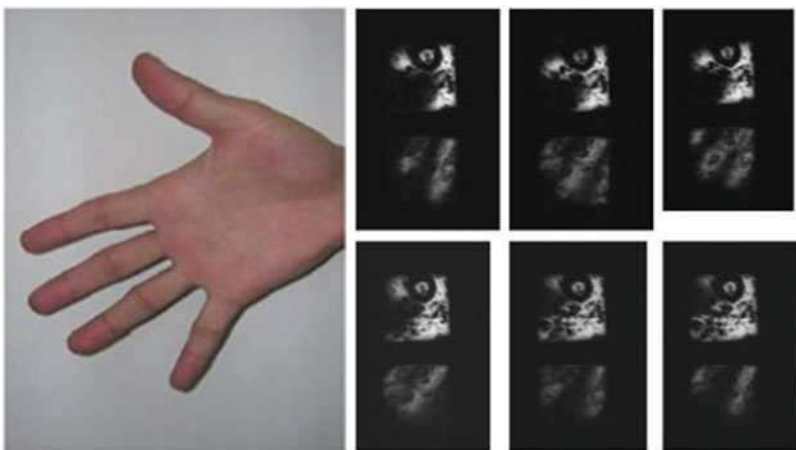
Figure 33 shows the graph that gray level of four areas changes as the time goes. In the test, characteristic unit of each area is chosen, and the result of the graph is

$$I = I_0 - \Delta I \cdot e^{-\tau} \quad (50)$$

Among them,  $I_0$  is the original gray level,  $\Delta I$  is the total variance of gray level, and  $\tau$  is the time. By numeration, it can be known the constant of thermal response of each area are: area1, 0.33 s; area2, 0.53 s; area3, 0.72 s; area4, 0.49. This result is



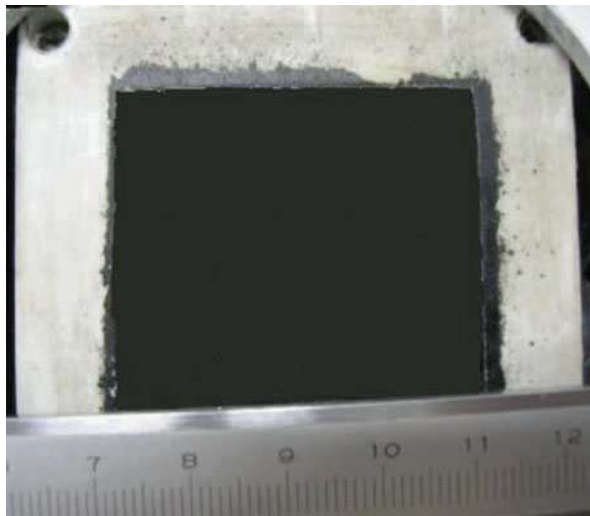
**Fig. 24** The microphotograph of FPA's each area



**Fig. 25** Human's hand and its thermal imaging



**Fig. 26** The infrared used for measuring grayscale response



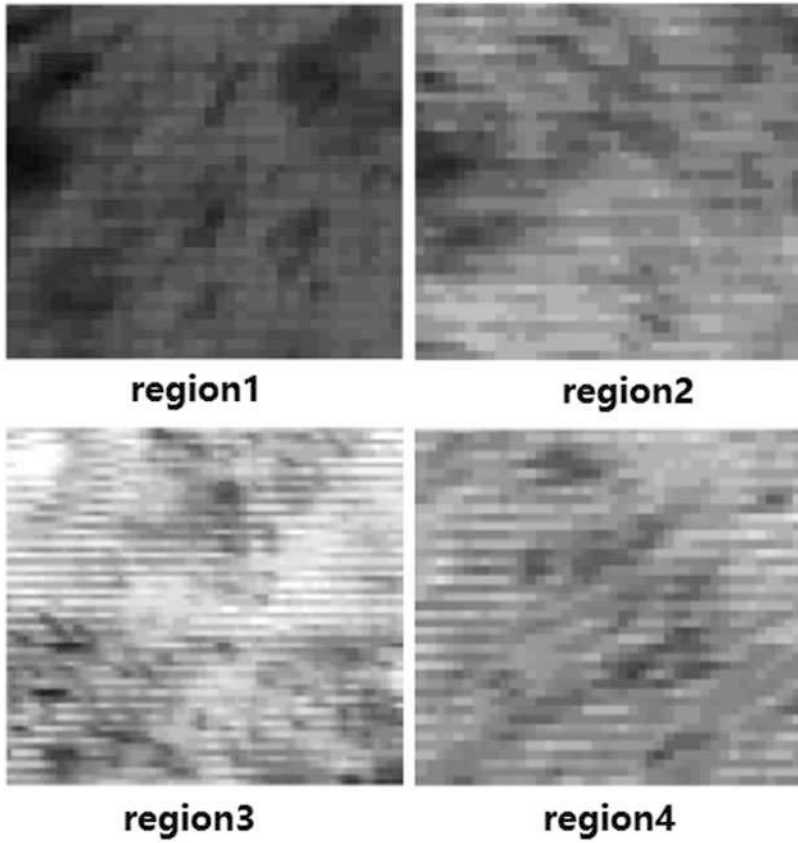
quite different from previous result, therefore, previous model which is used to infer FPA's thermal response time is not correct.

### Correction of the Model

In the model validation test, it can be found that the description for FPA's thermal response time of previous model is quite different from truth. The previous model refers to a model proposed by Professor Mao. Actually, hollow FPA designed by us has big difference from Mao's design which has silicon on the bottom. FPA's temperature rises when it is contaminated by radiation, then FPA which has silicon on the bottom proposed by Prof. Mao may introduce heat to the bottom, as is shown in Fig. 34. Silicon has a good thermal conductance ( $135 \text{ W}/(\text{m}\cdot\text{K})$ ), a big thermal capacity ( $700 \text{ J}/(\text{kg}\cdot\text{K})$ ), and mass, so silicon has a constant temperature compared with FPA's microcantilever array, therefore, the structure basically will not produce thermal perturbation between pixels.

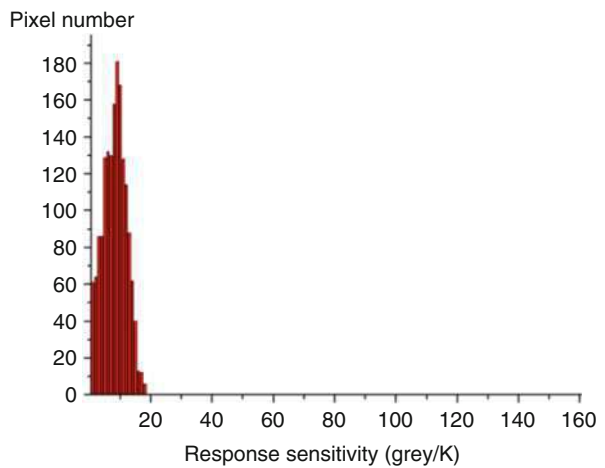
As to the hollow FPA proposed by us, pixel is different from that of FPA which has a bottom, but connect to reticular brace summer. In this reticular structure, heat flux can only propagate around and cause thermal interference between pixels. By finite element simulation, the transmission of energy can be clearly presented; Fig. 35 shows the transmission process by Ansys. The figure clearly shows that when FPA pixel is radiated, the energy propagates along brace summer to the around, and affect pixel around.

In order to describe and analyze the thermal response time of grid structural full hollow-out FPA, a circuit model for analysis can be proposed (Shi 2007). When we set off from the thermal equilibrium equation (26) and take the heat flux generated by infrared heat radiation to the array element as  $q_s$ , it can be seen from the right side of equation that it is dissipated and stored in two different ways: (1) The dissipative part

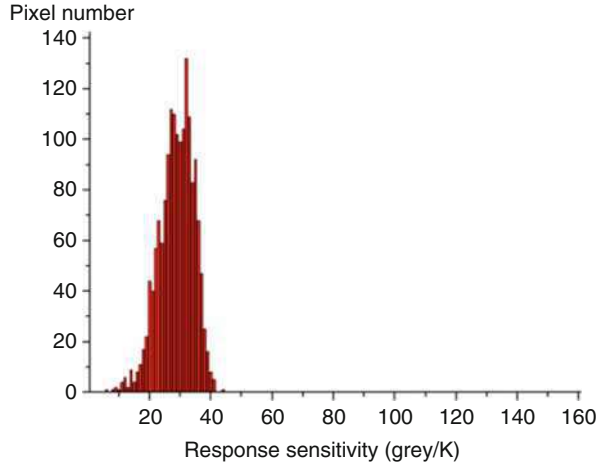


**Fig. 27** The response of four areas in the target infrared source at 42 °C

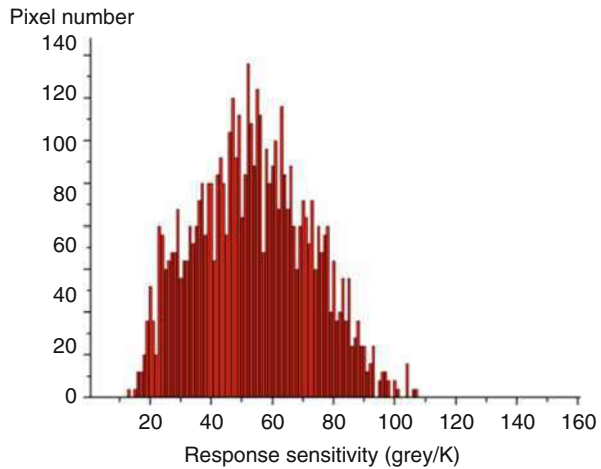
**Fig. 28** Sensitivity's histogram of areal



**Fig. 29** Sensitivity's histogram of area2



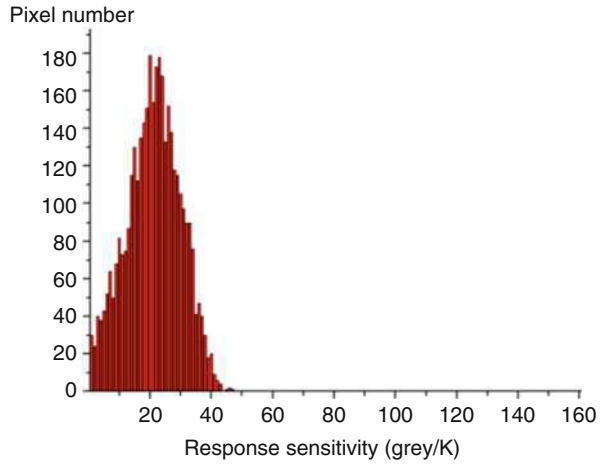
**Fig. 30** Sensitivity's histogram of area3



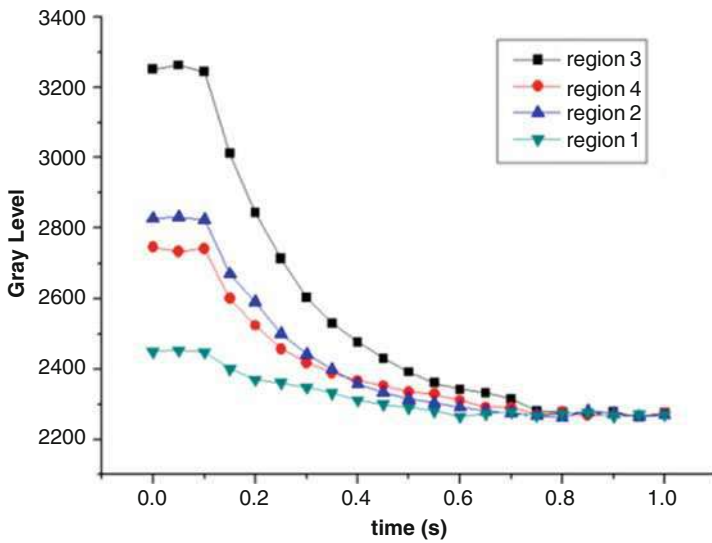
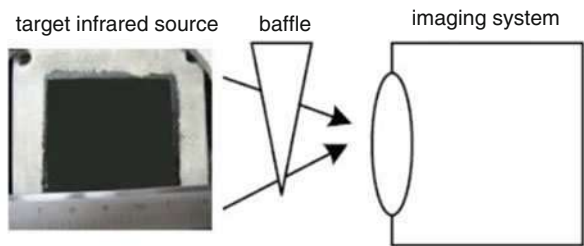
is the heat exchange between array element and the environment  $-G_{\text{Total}}(T_c - T_a)$ ; (2) the storage part is  $\rho V c \frac{dT_c}{dt}$ . Figure 36 is the equivalent circuit model corresponding to array element thermal model. Circuit balance equation  $i_s = -\frac{1}{R}(U_c - U_a) + C_c \frac{dU_c}{dt}$  just corresponds to the heat balance equation (26); the heat flux  $q_s$  is corresponding to the current  $i_s$ ; the total thermal conductivity  $G_{\text{Total}}$  of array element is corresponding to the conductivity  $\left(\frac{1}{R_{\text{leg}}} + \frac{1}{R_{\text{rad}}}\right)$ ; the total heat capacity of array element is corresponding to the capacitor  $C_c$ .  $R_{\text{leg}}$  is composed by the resistance of bi-material part in the array element cantilever beam  $\sum_i R_{\text{SiNx,Au},i}$  and the resistance of the SiNx material part  $\sum_i R_{\text{SiNx},i}$ .

These two models are strictly corresponding both in the physical sense and the equation form. The establishment of equivalent circuit model makes it more favorable to calculate the thermal response time, which makes the physical meaning more

**Fig. 31** Sensitivity's histogram of area4

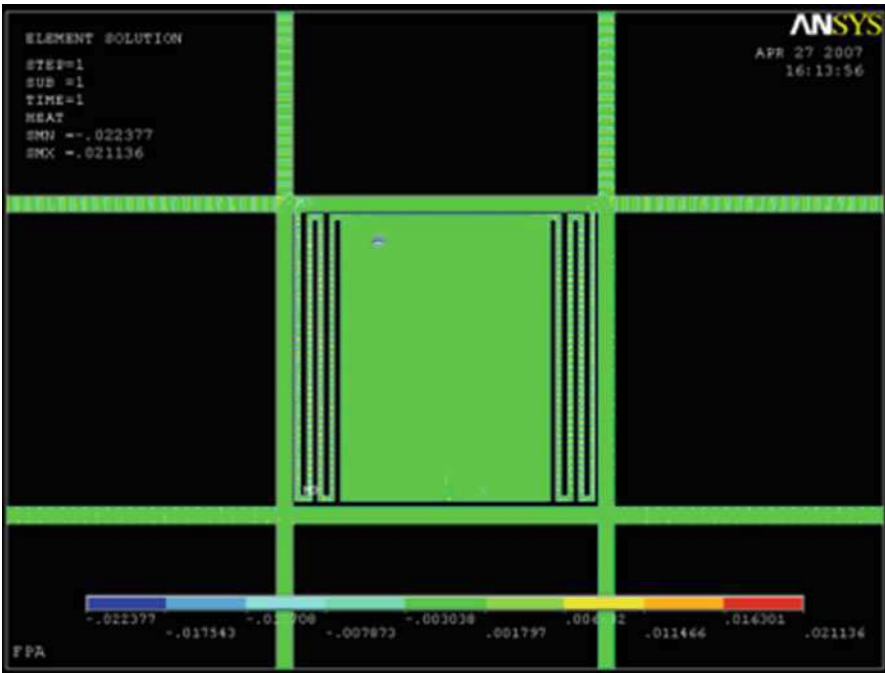
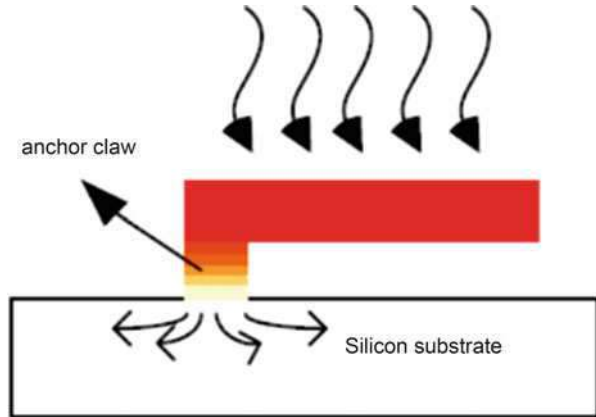


**Fig. 32** Experimental facility of thermal response time of FPA



**Fig. 33** Gray level of four areas changes as the time goes

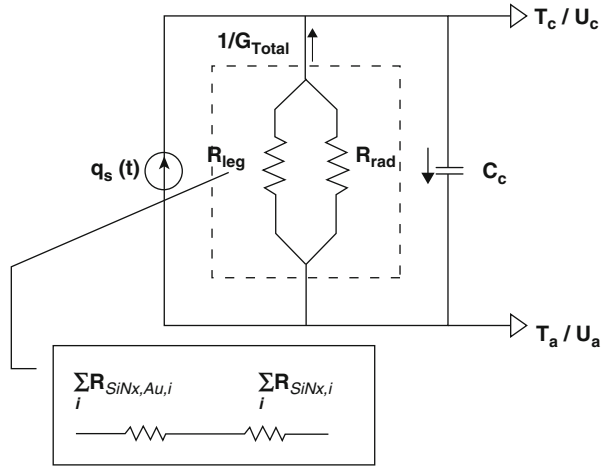
**Fig. 34** FPA introduces energy through the anchors into the body silicon



**Fig. 35** The energy propagation of substrate-free full hollow-out FPA

clearly. Derived from the equivalent circuit model, the thermal response time of single pixel is  $\tau = C/G_{total}$ , the result is same as the previous model. But the whole thermal response time of FPA (defined as  $\tau_{FPA}$ ) includes not only the single element's thermal response time but also the time heat transferred from single element to the FPA boundary (defined as  $\tau_{frame}$ ), as well as the thermal crosstalk time

**Fig. 36** The equivalent circuit model corresponding to array element thermal model



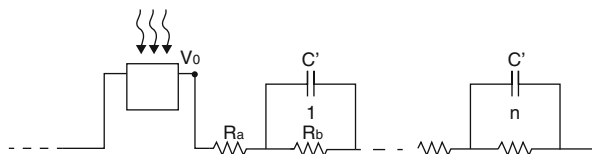
between array elements (defined as  $\tau_{induct}$ ). Therefore, the whole thermal response time of FPA should be  $\tau_{FPA} = \tau_{th} + \tau_{frame} + \tau_{induct}$ . The thermal crosstalk time refers to: when an array element generates heat, the other array elements around are sensitive to heat, the time required to reach thermal equilibrium with each other.

In order to estimate the thermal crosstalk time, one-dimensional model is used to calculate, Fig. 36 for the circuit model of thermal crosstalk. This model assumes that the FPA system is not affected by thermal radiation thermal conductivity. In the model, the leftmost element is the radiation element. we assume that the radiation element is in the center of the grid, so the left and right cases are symmetric, and only need to analyze the results of one side. In the figure,  $V_0$  is the voltage at the radiation element,  $R_c$  is the resistance between elements, and  $R_0$  is the resistance of the grid parallel connected with pixel. Through this model, both ends energy of  $n$ -th pixel away from the radiation pixel can be got as

$$Q_n = \frac{VR_0c'}{n(R_c + R_0)} \tag{51}$$

It is assumed that when the affected pixel energy is less than 10 of the radiation pixel energy as not affected, then the number of pixels  $n$  that could be affected by the radiation pixel can be calculated through the formula 51. Take the parameters in the region 3 from the four regional FPA designed, it can be calculated that  $n$  is equal to 3. Consider that the mesh is actually a two-dimensional structure, namely the energy spread in four directions as shown in Fig. 38. In a two-dimensional mesh, a radiation pixel will be induced by the surrounding 48 pixels. Therefore, the thermal crosstalk time  $\tau_{induct}$  is considered 49 times of the array element thermal response time  $\tau_{th}$  (0.02 s). The value of  $\tau_{frame}$  in the regional 3 is calculated as  $10^{-4}$  s, so it can be ignored. Then the thermal response time of whole FPA regional 3 as  $\tau_{FPA}(\tau_{FPA} = \tau_{th} + \tau_{frame} + \tau_{induct})$  is 0.98 s. This result is slightly larger than 0.72 s measured by experiment, because the effects of thermal radiation are not considered

**Fig. 37** A thermal crosstalk circuit model of full hollow FPA



in the model. The existence of thermal radiation makes the actual unit crosstalk effect smaller than the model estimation so that the crosstalk time is shorter (Fig. 37).

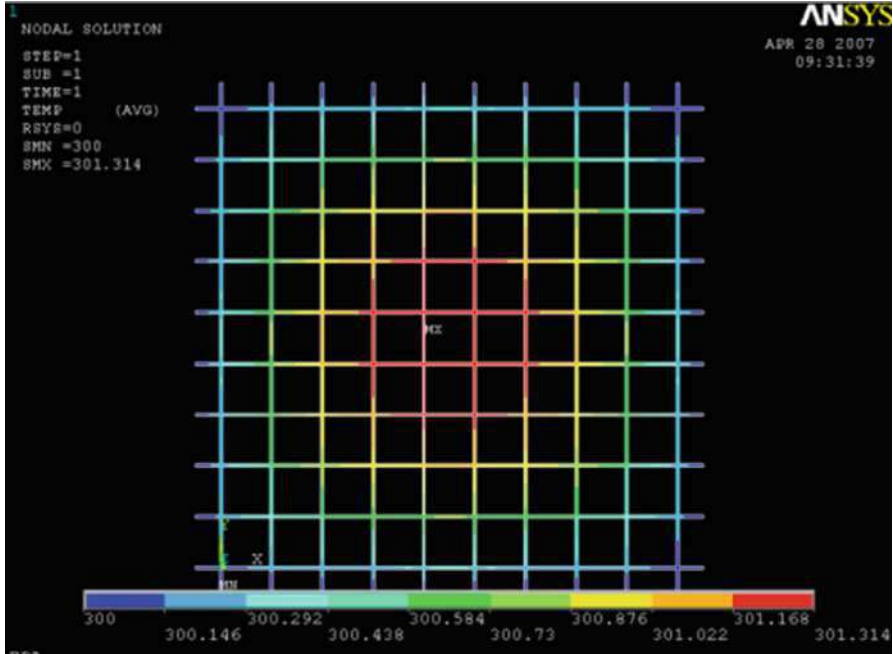
In models and calculations, it can be found that the response time of full hollow structure FPA is different with silicon substrate type FPA, and the response time of full hollow structure FPA is determined by the thermal crosstalk time rather than pixel response time. As can be seen from the formula 51, an effective means to reduce the thermal crosstalk between pixels is to increase the thermal resistance between the pixels.

### Failure Analysis of the System Under High g Inertial Load

Most infrared imaging systems are facing wide application in military, so it is very important to analyze the failure of the system under high g inertial load. The uncooled infrared imaging optical readout system has achieved the manpower static imaging under room temperature. The principle based on is: By detecting changes in the output spectra of square micro-mirror, optical detection part can decipher the angle change and distribution of the cantilever beam, and the temperature field of the object measured is displayed in the form of light intensity image finally. Under high g inertia load, the thermal induced mechanical rotation of the cantilever beam also including the rotation angle caused by the inertia load. In the range of elastic deformation, the rotation angle caused by the inertia load increases linearly with the linear increase of the acceleration. (It can be seen in the finite element simulation.) The optical readout section can be based on a certain algorithm; the linear increase of the rotation angle caused by the inertia load is subtracted from the total angle, so it will not affect the imaging results. It is worth noting that, when the angle caused by the inertial load transforms from elastic deformation (linear increase area) to plastic deformation (nonlinear increase area), the required imaging cannot be reflected correctly by optical readout system. In this transformation, the failure stress will appear in the FPA structure, which is an important parameter to measure failure.

### Experimental Device and Verification

The impact resistance of FPA can be validated by impact loading the Hopkinson bar, but first of all FPA should be vacuum packaged, which is a problem to be solved in the future. Currently available experimental device diagram is shown in Fig. 38. Installing the encapsulated chip in the tail of Hopkinson bar, firing a bullet by compressed air, coaxial impacting at the beginning of Hopkinson bar, an approximate half sine compressive strain pulse will be generated in the Hopkinson bar, and longitudinally propagate along the Hopkinson bar.



**Fig. 38** Ansys simulation results of energy transmission in FPA on the mesh

When the stress wave is propagated in the slender guide rod, the attenuation and dispersion can be neglected, and the strain gauge at the center of the rod can be measured. The mass of the chip can be ignored when the strain pulse is transferred to the interface between the Hopkinson rod and the packaged chip, and the velocity of the interface particle can be obtained according to the one-dimensional stress wave theory as:

$$v_1(t) = 2C\varepsilon(t) \tag{52}$$

In the formula, C is the wave velocity in the bar and  $\varepsilon$  is measured by the strain gauges attached to the central of Hopkinson bar.

Since the encapsulated chip is tightly stuck to the tail of Hopkinson bar, and its quality is small enough to be ignored, the speed of infrared FPA chip should be equal to the particle velocity at the tail of the Hopkinson bar.

$$v_2(t) = v_1(t) \tag{53}$$

The acceleration of chip is

$$a(t) = \frac{dv_2(t)}{dt} = 2C \frac{d\varepsilon(t)}{dt} \tag{54}$$



The peak acceleration is approximately

$$a \approx 2C \frac{\Delta \varepsilon}{\Delta t} \quad (55)$$

In the formula,  $\frac{\Delta \varepsilon}{\Delta t}$  is the average strain rate for the strain rising stage.

In order to determine the corresponding relationship between the voltage signal of strain gauge and strain, first of all, it needs to calibrate the ultrahigh dynamic strain gauge. The calibration step height is  $X$ , and the corresponding strain is  $X_g$ . In impact experiment, if the strain gauge output voltage signal is  $U_g(t)$  after amplified by the ultrahigh dynamic strain gauge, the corresponding strain will be

$$\varepsilon(t) = \frac{X_g U_g(t)}{X} \quad (56)$$

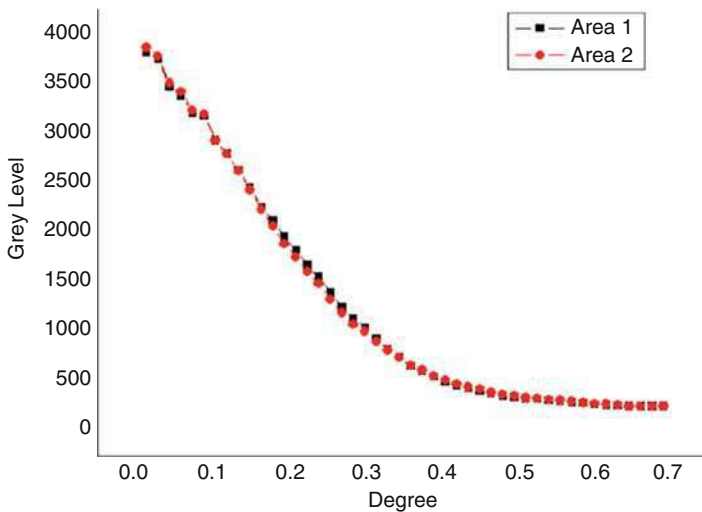
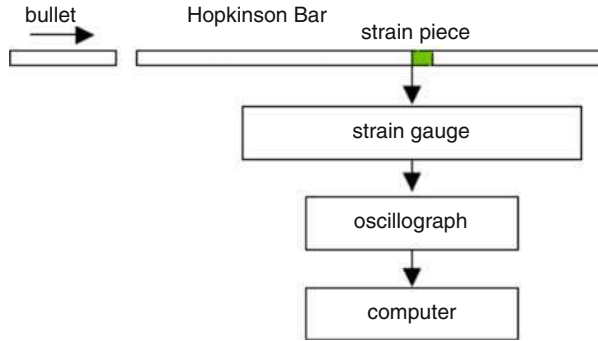
From the above two types available, taking  $\Delta U_g$  and  $\Delta t$  respectively as half of strain voltage signal peak and strain pulse width, the acceleration of chip can be estimated according to the above formula (Fig. 39).

Through the preceding analysis, the optically readable uncooled infrared system used in 2715 g acceleration under inertial load needs optical readout system that should be able to withstand the same 2715 g acceleration. However, the acquisition frequency of CCD we used is 30 Hz, linear range of optical detection sensitivity measured in the experimental is  $\pm 0.3$  degrees (Fig. 40). These constraints limit the impact resistant ability of the whole system to 448 g (finite element simulation results).

## Simulation Design of Full Hollow Double-Layer Infrared Focal Plane Array Devices Based on the Sacrificial Layer

Continuously optimized technology can provide a basis for designing more perfect devices. To pursue the optimization of the device's performance, it is necessary to expand the technology and make it more standard and perfect. Uncooled infrared imaging system based on optical readout always moves in the direction of small pixels, large arrays, high fill factor, high sensitivity, and low noise equivalent difference in temperature. Thus, based on the single-layer structure given in the previous section, it is proposed that full hollow double-layer infrared focal plane array devices are based on the sacrificial layer. The upper layer of the device includes reflectors and support beams, both of which are SiNx/Au bi-material; thermo-mechanical deformation generates after absorption of infrared heat. The lower layer is heat absorbing plate and insulation beam, both of which are SiNx material, which functions as a thermal isolation. Meanwhile, the upper and lower boards form a resonant cavity (through the sacrificial layer technology), which can increase infrared absorption; substrate-free full hollow structure reduces the loss of infrared heat. To achieve this, it is necessary to break through the bottleneck of sacrificial

**Fig. 39** Impact loading experimental device using Hopkinson bar



**Fig. 40** Relationship between gray level and array element maximum turning degree

layer technology to avoid the release and adhesion between upper and lower layers. We take that into account in the design, connecting the upper and lower boards of the resonant cavity through holes, and staggered designing of the deformed beam and heat insulation beam, so that in this aspect of design, we can try to avoid adhesion problems.

**Overview of Structure and Thermodynamic Models**

Figure 41 is the structure of full hollow double-layer infrared focal plane array based on the sacrificial layer, the colored figure on the left makes it easier for the readers to see clearly the connection between upper and lower layers, especially the connection between deformation beam and insulation beam.

Figure 42 is the thermodynamic model, substantially same as the full hollow single-layer infrared focal plane array model previously discussed; the difference is

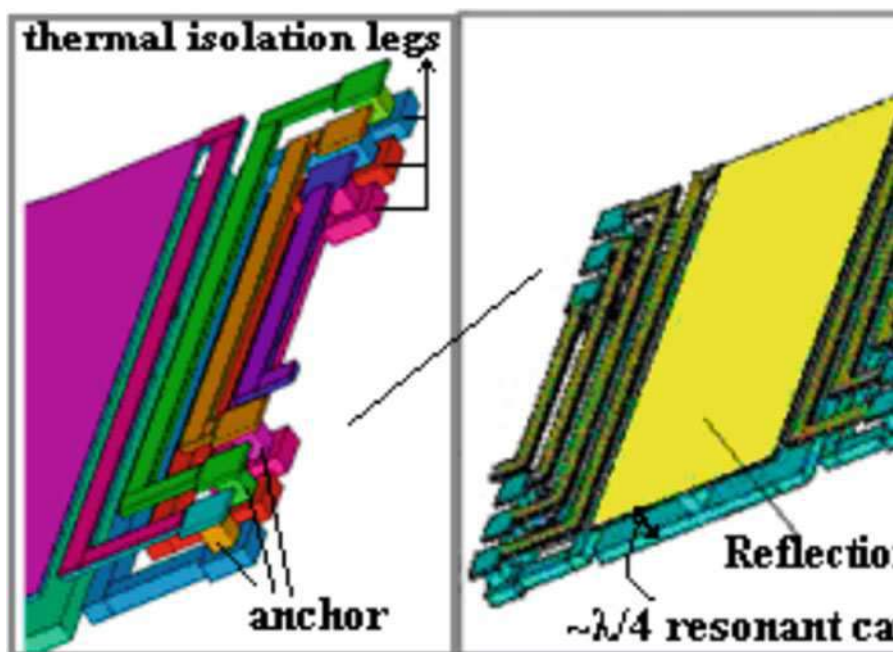
that, due to the more complex structure, if the device is not packaged in a vacuum, the thermal crosstalk in the structure should be taken into consideration, including the thermal conduction, thermal convection between the upper and lower plate in a resonant cavity, and thermal convection between upper and lower beams. As the simulation objects are vacuum packaged, there is no need to consider the thermal crosstalk (Fig. 42).

### Numerical Analysis of Thermal Performance and Thermo-Mechanical Properties

Based on the sacrificial layer, the thermal property of full hollow double-layer infrared focal plane array is similar to full hollow single-layer infrared focal plane array previously discussed but different in the expression of temperature rise. When the depth of the infrared resonance absorption chamber is  $\sim\lambda/4$  ( $\lambda$  for the infrared wavelengths absorbed), the infrared heat absorbed is twice as much as that of full hollow single-layer structure, so the temperature rise as:

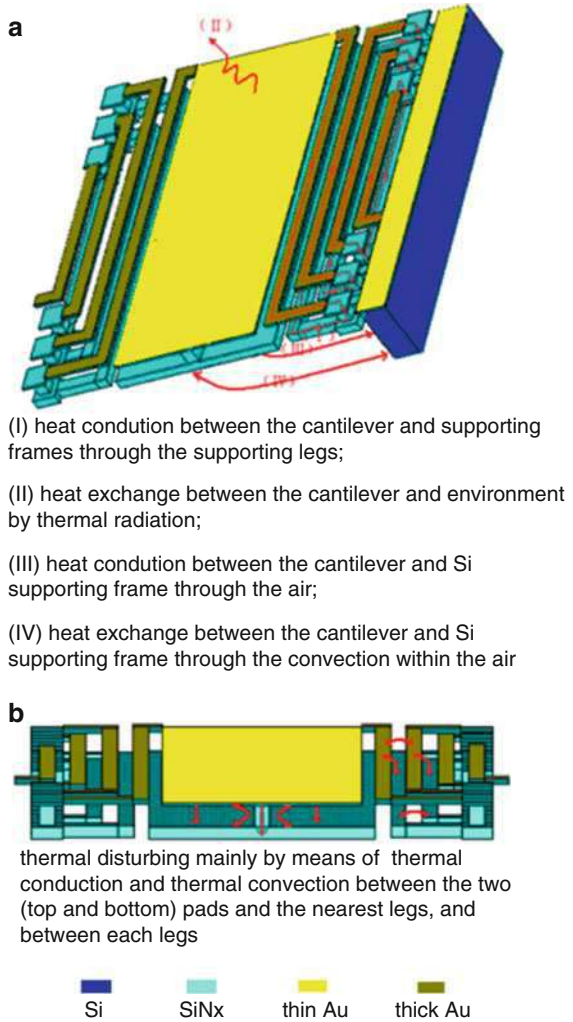
$$H = \frac{\delta T_c}{\delta T_s} = \frac{1}{G_{\text{Total}}} \cdot \frac{A_{ab} \tau \epsilon \pi}{2F_{no}^2} \frac{dL}{dT_s} \quad (57)$$

The thermal isolation structure in this device made the temperature rise mainly gathered in the first deformed beam near the reflector, so the deformation of FPA also comes mainly from the beam.



**Fig. 41** The structure of full hollow double-layer infrared focal plane array

**Fig. 42** Thermodynamic models



$$S_T^* = \frac{\Delta\theta_{\max}}{\Delta T_c} = \frac{\sum_{i=1}^n \theta_i}{\Delta T_c} = 6(\alpha_1 - \alpha_2) \left( \frac{n+1}{h_2 K} \right) \int \Delta T(x) dx \quad (58)$$

Suppose the temperature distribution in the upper reflector and the deformed beam is homogeneous, the lower thermal insulation beam is linear. Similarly, elements influencing the thermal mechanical sensitivity are differences of thermal expansion coefficient between Au and SiNx, SiNx film thickness, thickness ratio of Au film and SiNx film, Young’s modulus, and so on. Due to the improvement of structure, n is to be taken as 0.67 in the light of  $S_T(S_T^*) \sim \left(\frac{n+1}{K}\right)$ . Detailed analysis is as follows.

The thickness of the lower SiN<sub>x</sub> is designed as 2 μm, the upper SiN<sub>x</sub> 0.3 μm, and the upper Au 0.2 μm. Like this, it can not only meet the need of performance but also satisfy the need of structural strength.

## Summary

The section designs and creates an FPA with four different design parameter regions. Testing the feedback parameters in FPA's different regions, the section verifies the previous presented FPA design model, within this process; the estimation of model's effects on FPA's sensitivity is verified, finding that there is a large gap (upto 1 order of magnitudes) between the reality and the estimation of FPA's response time according to the previous model. Through careful analysis of the structure of the hollowed-bond FPA, it is found that FPA's response time is not determined by that of individual pixels. Instead, it is by the crosstalk phenomenon among pixels. In the study of heat conduction analysis, a kind of method is put up to estimate reckon on the crosstalk phenomenon, which is to take electric current as equivalence to the process of heat flux spread in FPA's array. By this method, the estimate-reckoned number of tested FPA's pixels is about 49, which is about the same result with the experimental value. The aforementioned work lays the theoretical foundation of the improvement of FPA's performance.

---

## The Optimization and Improvement of Monolayer FPA

In the previous work, two hollowed-out FPA had been obtained, with pixel size of  $200 \times 200 \mu\text{m}^2$  and  $120 \times 120 \mu\text{m}^2$ , as well as a series of imaging results of normal temperature objects. Though the results are exciting, the imaging results of those two FPA are still far from the practical result in respects of distinguishability, sensitivity, and response time. In this section, based on the verification and improvement of the previous model, we designs a  $57 \times 57 \mu\text{m}^2$  pixel size hollowed-out FPA by carefully analyzing of various design parameters that influence the performance of FPA and the characteristics of hollowed-out FPA. In the design and machine process of new FPA, the low-sensitivity problem caused by adopting multilayer film stress-matching technology was solved, and an almost practical imaging result was finally achieved.

## The Shortage of the Previous Monolayer FPA

In the previous work, we have created two hollowed-out FPA with pixel size of  $200 \times 200 \mu\text{m}^2$  and  $120 \times 120 \mu\text{m}^2$ , as shown in Fig. 43 (Li et al. 2006a; Jiao et al. 2007). From the imaging result, it could be found that: (1) Infrared imagery is clear but its distinguishability is obviously low. Nowadays, in the market, the size of uncooled infrared FPA's pixel is about  $50 \times 50 \mu\text{m}^2$ ; some of the high-end products could achieve  $25 \times 25 \mu\text{m}^2$  (Ueno et al. 2005). With the decrease of the size of

pixels, FPA's sensitivity decreases obviously. (2) FPA's sensitivity is not enough to guarantee the final NETD of the system under the commercial practical standard of 100mk. (3) The crosstalk phenomenon among FPA's pixels is quite serious; the imaginary-line-marked area is the crosstalk between fingers. The analysis of the FPA model in the last section shows this crosstalk determines FPA's response time.

The above three shortages could be concluded into following two questions: in the situation of pixels' size scaling down, how to achieve higher sensitivity and how to solve the crosstalk.

## Higher Sensitivity Gained in the Situation of Pixels' Size Scaling Down

As known to all, infrared probe's sensitivity declines with the decrease of its size. In order to keep the previous sensitivity or even achieve higher sensitivity after FPA scaling down, it needs to optimize FPA in both parameterized and structural aspects.

### Optimization of FPA's Graphic Design

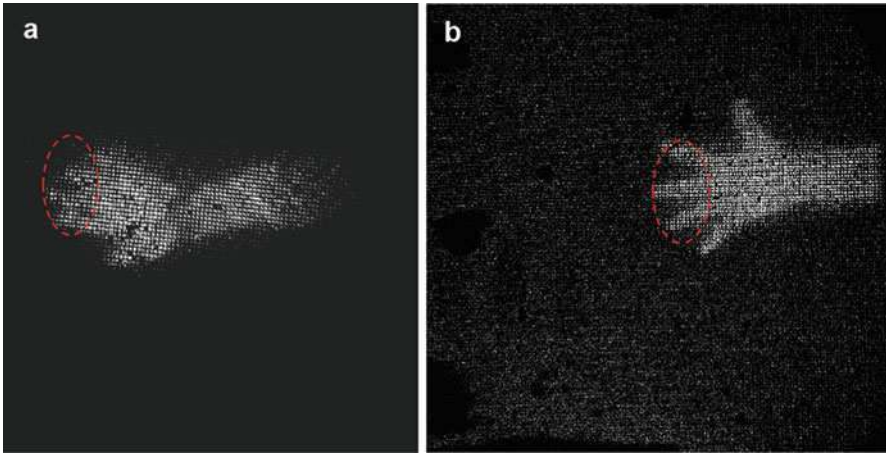
In the situation of fixed pixels' size, monolayer FPA's graphic parameter includes: the width of finite deformation beam and thermal isolation beam  $w$ , the interval  $g$  between beams, and the beam's inflection time  $m$ , shown as Fig. 44. Within it, parameters like the width of finite deformation beam and thermal isolation beam  $w$  and  $g$  are expected to be maximized in the technical process design, in order to improve FPA's fill factor and thermal isolation performance. Besides, the beam's inflection time  $m$  affects a series of performance of FPA, such as infrared thermal absorption transformation. Thus, the optimization of  $m$  is helpful to improve FPA's performance. In this case, the FPA of  $57 \mu\text{m} \times 57 \mu\text{m}$  pixel size is taken as an example.

When calculating the thermal transformation efficiency  $H$  of the micro beam element, in formula (29), factor  $A_d$ ,  $G_{\text{leg}}$  are functional attribute of inflection time  $m$ ,  $G_{\text{leg}}$  could get from formula (34):

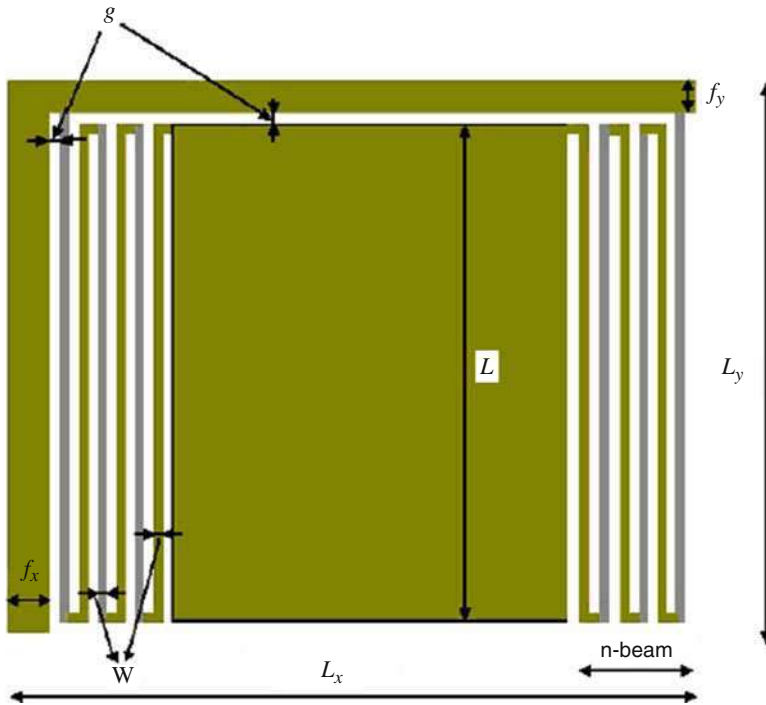
$$G_{\text{leg}} = 2 \left[ \frac{mL}{k_{\text{SiNx}}A_{\text{SiNx}} + k_{\text{Au}}A_{\text{Au}}} + \frac{mL}{k_{\text{SiNx}}A_{\text{SiNx}}} \right]^{-1} \quad (59)$$

Figure 45 shows the relation between the thermal transformation efficiency  $H$  and  $m$ . From the picture,  $H$  is not  $m$ 's monotone function. At the beginning, with the increase of  $m$  (from 1 to 2), though the area of absorber plate is decreasing,  $G_{\text{leg}}$  and  $G_{\text{rad}}$  are both decreasing, leading to the total thermal conductance decreasing more rapidly, thus  $H$  increasing instead of decreasing. When  $m$  further increases ( $m$  from 2 to 3),  $A_d$ 's decrease become the predominance, thermal transformation efficiency the decrease.

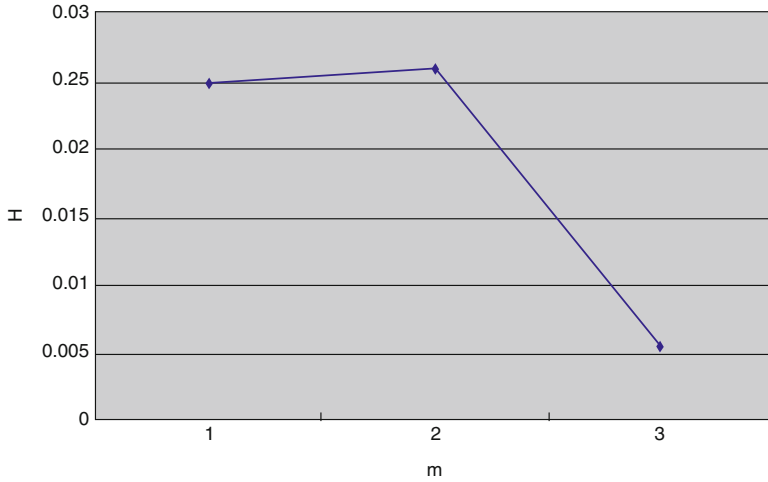
In the last section, formula 41 shows the thermal transformation sensitivity  $S_t$  is in direct ratio with  $m$ . Combined with the micro beam element thermal transformation efficiency  $H$ , it is possible to optimize  $m$ , so as to maximize its probing performance. Figure 46 shows the transformation of micro beam element thermal mechanical



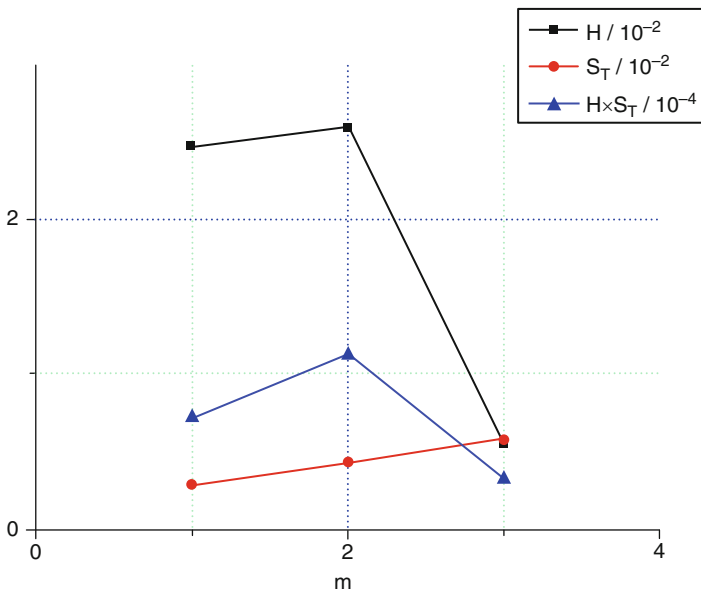
**Fig. 43** Imaging result by infrared sensor in original FDA. (a)  $200 \times 200 \text{ um}^2$  pixel. (b)  $120 \times 120 \text{ um}^2$  pixel



**Fig. 44** Pixel sketch map of  $m$  inflection



**Fig. 45** Relation between thermal conversion H and inflection M



**Fig. 46** Optimization of inflection m

response  $S_t$ , H, and  $H \times S_t$  with the charge of m. From the picture, in situation of the element size of  $57 \times 57 \text{ um}^2$ , the optimal inflection of transformal enlarged beam is 2. At  $m=2$ , the detection performance of the microbeam element is 1.56 times as much as that of the single folded beam, and is 3.5 times that of the 3 folded beam.



## Optimization of FPA's Thickness Design

### Optimization of SiNx Layer's Thickness

As to monolayer FPA, SiNx layer has three tasks: first, absorbing infrared; second, forming thermal isolation structure and guaranteeing the heat conductivity; third, forming double-deck material beam with Au layer, so as to create thermal transformation. Thus, in order to get the optional thickness of silicon nitride film, it is necessary to understand the influence of its thickness in these three processes.

#### (a) Effect of SiNx layer thickness on infrared absorption

According to Bouguer's law (Mo 1991) (P.Bouguer, 1729), when light travels in some media, the intensity of light decreases with the increase of the distance to the medium. Bouguer's formula on the light intensity variation by change of incident depth is

$$I = I_0 e^{-\alpha l} \quad (60)$$

In this formula,  $I_0$  and  $I$  were respectively the light intensities when the incident depth is 0 and  $l$ ;  $\alpha$ 's dimension is the inverse of length, and  $\alpha$ 's physical meaning is the thickness of the media through which the light travels while reduced to  $e^{-1} \cong 36\%$  of the original due to absorption. And according to Formula (60), the ratio of the thickness to the radiation absorption is

$$\frac{I_a}{I_0} = 1 - e^{-\alpha l} \quad (61)$$

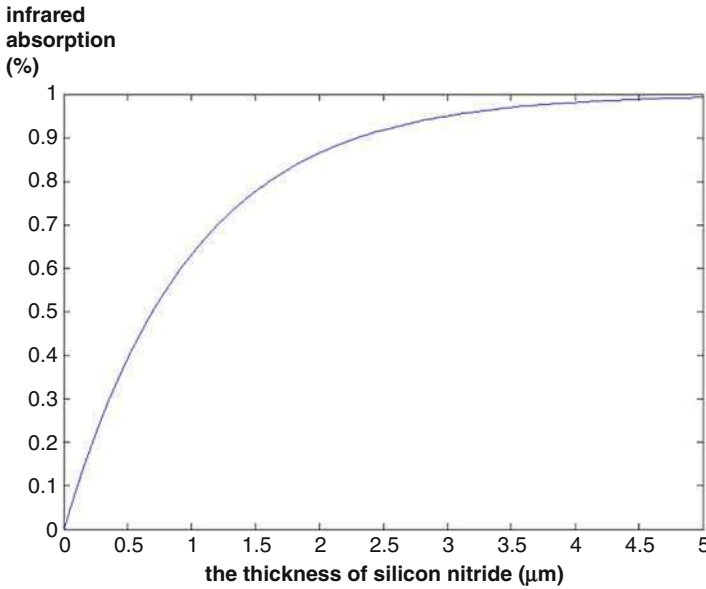
In this Formula,  $I_a$  is the light intensity absorbed (Fig. 47).

#### (b) Effect of SiNx layer thickness on FPA thermal conductivity

As the outcome of discussion about the model in the last section, the thermal conductivity of FPA pixel depends on the thermal conductivity of thermal isolation microcantilever, while the thermal conductivity of thermal isolation microcantilever is affected by SiNx layer thickness, as given in Formula 34. So that, according to Formula 27, Formula 31, and Formula 34, it can be concluded that the total thermal conductivity of FPA pixel is affected by the thickness of SiNx film.

$$G_{\text{total}} = G_{\text{leg}} + G_{\text{rad}} = 2w \left[ \frac{L}{\sum_i k_i t_i} + \frac{L}{k_{\text{SiNx}} t_{\text{SiNx}}} \right]^{-1} + G_{\text{rad}} \quad (62)$$

#### (c) Effect of SiNx layer thickness on thermal deformation



**Fig. 47** Infrared absorption varies with the thickness of silicon nitride

By transformed Formula 44, the effect of the thickness of SiNx layer on thermal deformation is

$$\theta = \left. \frac{d\delta}{dz} \right|_{x=L} = 6(\alpha_2 - \alpha_1) \left( \frac{1+n}{t_{SiNx}K} \right) L\Delta T \tag{63}$$

In this formula, n is the ratio of Au layer thickness to SiNx layer thickness ( $n = t_{Au}/t_{SiNx}$ ).

And according to Formula 29, the deformation capacity of the bi-material micro-cantilever is inversely proportional to the thickness of the SiNx layer, when n is a constant value.

From the above discussion we can conclude that the comprehensive relation of FPA response sensitivity  $H \times ST$  to the thickness of SiNx layer is

$$H \cdot S_T \propto (1 - e^{-\alpha t_{SiNx}}) \cdot \left[ \left( \frac{L}{\sum_i k_i t_i} + \frac{L}{k_{SiNx} t_{SiNx}} \right)^{-1} + G_{rad} \right]^{-1} \cdot \left( \frac{1+n}{t_{SiNx}} \right) \tag{64}$$

In this Formula, SiNx’s transmission depth  $\alpha^{-1}$  is 1 μm when the infrared target wave is 10 μm; the thermal conductivities of SiNx and Au are respectively 5.5 W/m·K and 296 W/m·K; Grad is about  $5 \times 10^{-9}$  W/K when the pixel size is about 50–60

square micron. And so the optimum value of SiNx thickness of FPA is about 0.310  $\mu\text{m}$ .

### The Optimization of Au Layer Thickness

From the microcantilever light readout FPA, Au layer plays such roles as follows. Firstly, it consists of bi-material beam together with SiNx layer and provides heat rotation angle for the micro beam depending on the huge thermal expansion coefficient difference. Secondly, the reflection layer is formed to reflect the readout light which reads out the change of thermal rotation angle of the micro beam. The following are discussed from the two aspects of the optimization of the Au layer thickness.

#### (a) Thickness optimization of bi-material beam

The purpose of bi-material beam is to provide maximum thermal rotation angle. Therefore, it is required to analyze the effect of the relation of the two materials' thicknesses on the deformation capacity of bi-material beam, before determining the thickness of the Au layer. According to Formula 41, the effect of the relation of the two materials' thicknesses on the temperature rotation angle is as follows.

$$\theta = 6(\alpha_2 - \alpha_1) \left( \frac{1 + n}{t_{\text{SiNx}} \left( 4 + 6n + 4n^2 + \frac{E_2}{E_1} n^3 + \frac{E_1}{E_2 n} \right)} \right) L \Delta T \quad (65)$$

As the Young modulus of Au is 73GPa, and the Young modulus of SiNx is 180GPa, so that the optimum value of n is 0.75. That means the Au layer thickness should be set as 0.225  $\mu\text{m}$ , when the SiNx layer thickness is the optimum value of 0.3  $\mu\text{m}$  (Fig. 48). Figure 23 shows the relation of the micro beam's deflection capability to the ratio of the thicknesses, when the SiNx layer thickness is constantly 0.3  $\mu\text{m}$ .

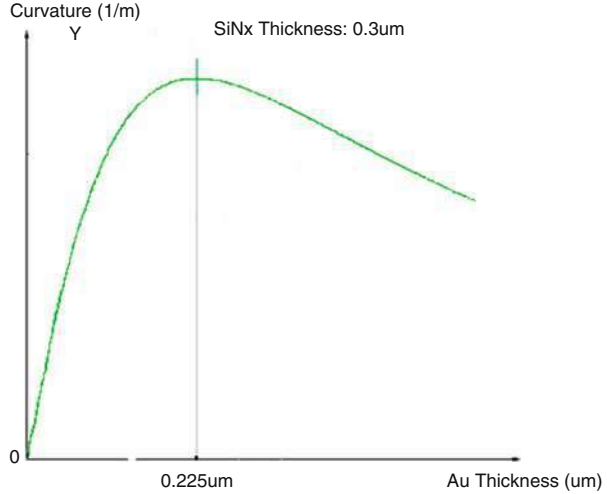
#### (b) Thickness optimization of reflector panel

The ideal state of reflector structure is as flat as a mirror, so that the readout light traveling to its surface can be effectively reflected back to the readout light path (Li et al. 2006a, b; Dong et al. 2005, 2006; Miao et al. 2006). But due to the stress mismatch between the SiNx film and the Au film, the reflector is usually bent into an arc. This deformation leads to not only spectrum widening but also significant decreases in sensitivity of readout system (Miao et al. 2007).

#### (i) Effect of reflector deformation on light readout response sensitivity

Supposing the maximum deformation migration of reflector is  $\Delta Z_{\text{max}}$ , the reflector's curvature radius can be expressed as

**Fig. 48** Effect of SiNx layer thickness on thermal deformation sensitivity, in case that SiNx layer thickness is constant



$$R = \frac{(L/2)^2}{2\Delta Z_{\max}} \tag{66}$$

In this Formula, L indicates the length of the deformation direction of the reflector. For the condition in Fig. 49, the deformation migrations of reflector  $\Delta Z_{\max}$  are respectively  $\lambda/5$  and  $3\lambda/5$ , and  $\lambda$  is the readout light wavelength (set at  $0.5 \mu\text{m}$ ). Thus, the effect of reflector deformation on light readout response sensitivity is calculated and explained as below.

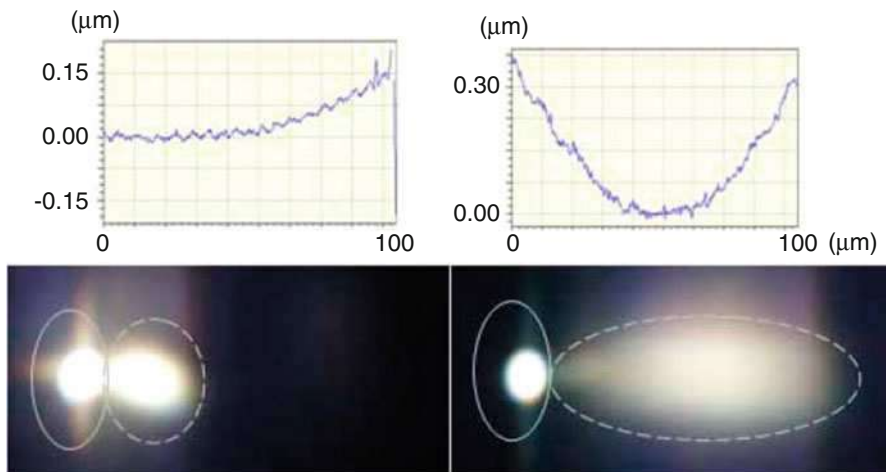
In the spectral plane knife edge filtering system which is adopted by our experimental group, the field distribution of the deformed reflector on Fourier lens posterior focal plane is

$$U_f(x_f) = (A/j\lambda f) \int_{-\infty}^{+\infty} \text{rect}(x/L) \exp(j2\pi x^2/\lambda R) \exp(-j2\pi x x_f/\lambda f) dx \tag{67}$$

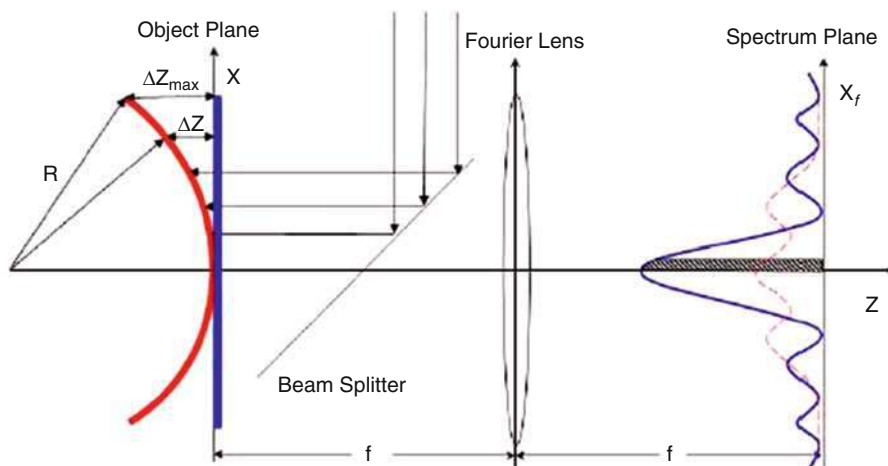
In this Formula,  $L = 100 \mu\text{m}$ ,  $j = -$ . Thus, the light distribution here is square of the absolute value. And this Formula shows close relation between the field distribution on Fourier lens posterior focal plane and the deformation quantity of reflector  $\Delta Z_{\text{ffl}}(x)\text{ffU}(x)_{\max}$ . Figure 50 shows the light distributions on Fourier lens posterior focal plane of deformed reflector and nondeformed reflector.

For vividly displaying the effect of the deformation on the optical readout sensitivity, we take four different kinds of deformation for example and calculation, with the values of  $\Delta Z_{\max}$  respectively as  $0, \lambda/5, 3\lambda/5$ , and  $8\lambda/5$ . By FFT calculating with Matlab, the light distributions of four examples are shown in Fig. 51.

According to the working principle of general plane knife edge filtering system, the trend of light readout sensitivity in different deformation conditions can be calculated (as in Fig. 52), when the filter blade is placed in the middle of spectrum.



**Fig. 49** Effect of reflector deformation on reflection spectrum



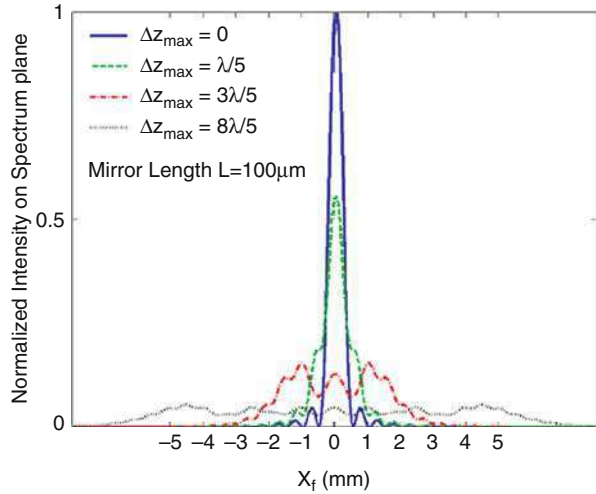
**Fig. 50** Light distributions on Fourier lens posterior focal plane of deformed reflector and nondeformed reflector

For example, when the reflector is an ideal mirror ( $\Delta Z_{max} = 0$ ), the readout sensitivity ( $k_0$ ) is 2.81/deg. And when the deformations of reflector are respectively  $\lambda/5$  and  $3\lambda/5$ , the response sensitivities  $K_1$  and  $K_2$  decreased. Thus, when the deformation is  $3\lambda/5$ , optical readout sensitivity is almost down to 1/8 of that in condition of ideal mirror.

(ii) Modification of reflector deformation

For a bi-material cantilever with one end clamped, the deformation caused by the internal stress can be analyzed generally with the structure as in Fig. 53 (Min and

**Fig. 51** Light distributions on Fourier lens posterior focal plane in different conditions of deformation



Kim 2000). In order to facilitate the analysis, it is supposed that the widths of two layers of film are both  $W$ , the stress distribution of upper layer film was  $\sigma_0$ , the stress distribution of lower layer film is  $\sigma_1$ , and the upper layer thickness ( $t_0$ ) is far less than the lower layer thickness ( $t_1$ ), which can keep the bi-material cantilever flat basically.

Here, we make an equivalence of upper-layer thin-film material to lower-layer thick-film one to accomplish deformation analysis as illustrated in Fig. 54 such that the upper-layer thin film whose width is  $w$  and Young modulus  $E_a$  is equivalent to a thin film whose width is  $w'$  ( $abEw'wE=$ ) and Young modulus  $E_b$ . The equivalent axis of this equivalence bilayer composite thin film is located at:

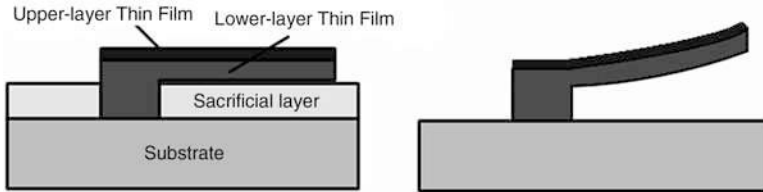
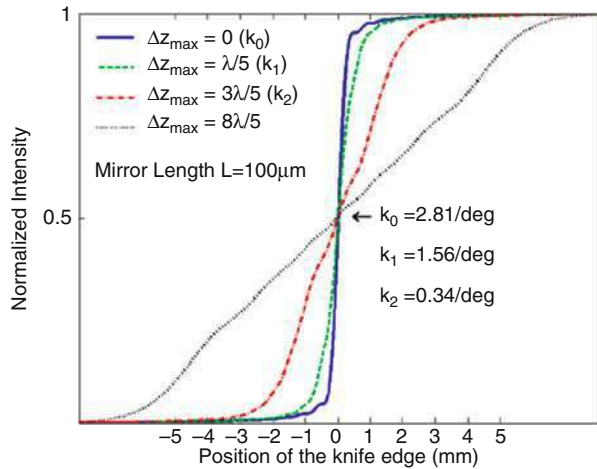
$$y_c = \frac{y_b w t_b + y_a w' t_a}{w t_b + w' t_a} \tag{68}$$

Hereinto,  $y_b$  and  $y_a$  are respectively natural axis of lower-layer thick film and upper-layer thin film. According to the equivalence given above, composite double-layer-material cantilever's torque  $M$  caused by its internal stress can easily be presented as:

$$\begin{aligned} M &= \int_{-y_c}^{t_b - y_c} \sigma_1 z dz + \int_{t_b - y_c}^{t_b - y_c + t_a} \sigma_0 z dz \\ &= (t_b^2 - 2y_c t_b) \sigma_1 + (t_a^2 + 2t_b t_a - 2y_c t_a) \sigma_0 \end{aligned} \tag{69}$$

The composite cantilever does not deform when the composite double-layer-material cantilever's torque  $M$  caused by its internal stress is 0. It can be inferred from the formula above that when  $\sigma_0$ ,  $\sigma_1$ , and  $t_b$  remain constant, variation of  $t_a$  can

**Fig. 52** Response sensitivity under different deformation conditions



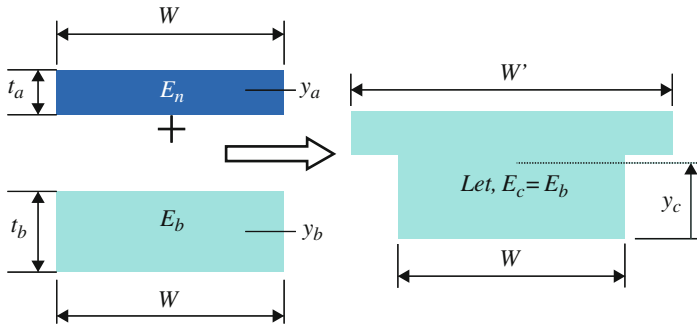
**Fig. 53** The deformation caused by the internal stress can be analyzed generally with the structure

make  $M = 0$ , namely the deformation of bi-material composite cantilever caused by its internal stress is 0.

As it is very difficult to precisely measure the residual internal stress of thin film during micromachining process, here an experimental method is adopted: setting the thickness and deposition condition of silicon-nitride thin film as constants, then matching the residual internal stress by changing the thickness of Au layer. Figure 55 illustrates bi-material cantilever’s deformation varied with different thickness of Au layer when the thickness of silicon nitride is set at  $0.3 \mu\text{m}$ . Through experiments, it is defined that under existing process conditions bi-material cantilevers will be substantially flat when the thickness of silicon nitride is  $0.3 \mu\text{m}$  and a  $25 \text{ nm}$  electron-beam evaporated Au layer is adopted.

### Improvement of Thermal Interference Among Pixels

About modification of response time model in section “[The Examination of and the Improvement on the FPA Unit Model](#),” the differences have been mentioned in energy transmission between the hollow structure FPA and silicon substrate FPA. Full hollow FPA pixel’s energy can only be transmission within the grid plane, so



**Fig. 54** Make an equivalence of upper-layer thin-film material to lower-layer thick-film one to accomplish deformation analysis

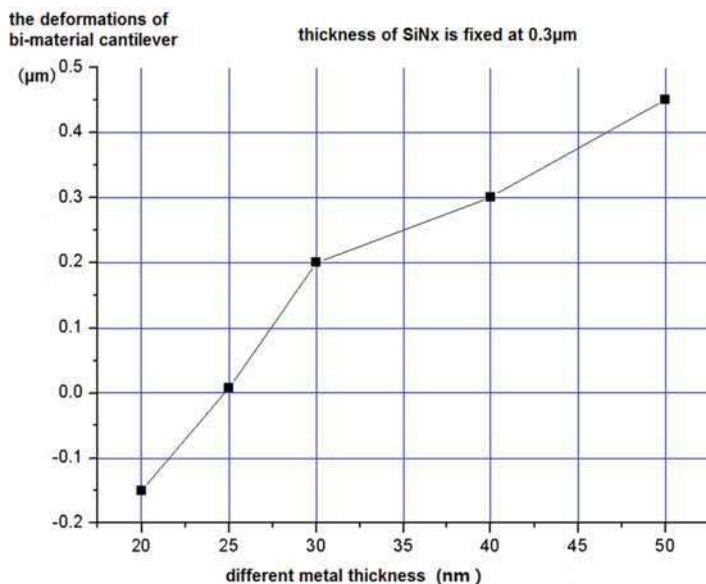
that the thermal interference among pixels as in Fig. 32. If we estimate the number of pixels that would be affected by the thermal interference by method of one-dimensional model which is described in Fig. 34, the thermal interference can affect 48 pixels around, in condition that the SiNx cantilever of  $3\ \mu\text{m}$  width,  $0.7\ \mu\text{m}$  thickness, and covered by Au film of  $0.025\ \mu\text{m}$  thickness. Formula 48 indicates that the pixel thermal interference is related to the thermal resistance (Rc) of supporting beam between two pixels and that the number of affected pixels reduces when the value of Rc rises. The physical meaning of this result is that the better the thermal isolation, the less the thermal interference. Also, same with silicon substrate FPA, it is another effective resolution to make the support beam structure both thermal conductivity and heat capacity large enough, so as to keep it at an even temperature while the pixel's temperature changes.

Because a full hollow grid structure has been adopted, it is relatively difficult to obtain a basically even-temperated framework with high thermal conductivity and high heat capacity. Therefore, thermal interference by enlarging the inter-pixel thermal resistance is tried to be reduced, that is, increasing the inter-pixel thermal isolation. At  $m=2$ , the detection performance of the microbeam element is 1.56 times as much as that of the single folded beam, and is 3.5 times that of the 3 folded beam. In order to achieve this goal, we reduce the width of the support beam of the original  $120\ \mu\text{m} \times 120\ \mu\text{m}$  pixel unit from  $10\ \mu\text{m}$  to  $3\ \mu\text{m}$  in the current structure, and cancel the Au film on the supporting beam grid structure. Through this improvement, the number of 1148 thermal-interfered pixels in  $57\ \mu\text{m} \times 57\ \mu\text{m}$  FPA is estimated as 24.

## Designing of Optimized Single-Layer FPA

For an overall consideration of those factors and the constraint of work craft, the FPA with the pixel area of  $57 \times 57\ \mu\text{m}^2$  is designed. Table 3 shows the designed parameters of the plane structure, Table 4 of parameters of FPA's longitudinal thickness, and Table 5 of FPA's performance parameters.





**Fig. 55** When the thickness of SiNx is fixed at 0.3  $\mu\text{m}$ , the deformations of bi-material cantilever with different metal thickness

Figure 56 shows FPA pixel layout designed by L-edit layout editing tools.

### Test of Optimized Single-Layer FPA

Figure 57 shows the whole photograph of FPA pixel at size of  $57 \times 57 \mu\text{m}^2$ . Figure 58 shows FPA pixel microscope photo. The entire array consists of  $160 \times 160$  micro-cantilever beam elements, which are supported by 3  $\mu\text{m}$  wide frames.

### Test of FPA Morphology

Figure 59 shows the shape of the FPA reflecting plate beam collected by Vecoo 3D profile instrument, and the reflector which is optimized by the internal stress matching can be seen to obtain an ideal flatness. Mean radius of curvature of the reflector in Fig. 59 is greater than 10 cm.

### FPA Imaging Result

Figure 60 shows a human hand image by this FPA. It can be seen that five fingers' outline is very clear, and each finger shows no "flippers" phenomenon due to interference. The boundary of the wrist and the sleeve of the figure are very clear, which reflect the gray difference caused by the temperature. Compared to the previous FPAs at sizes respectively of  $200 \times 200 \mu\text{m}^2$  and  $120 \times 120 \mu\text{m}^2$ , the optimized FPA pixel size of  $57 \times 57 \mu\text{m}^2$  has been greatly improved in terms of spatial resolution and the successful restraining of thermal interference phenomenon

**Table 3** the designed parameters of the plane structure

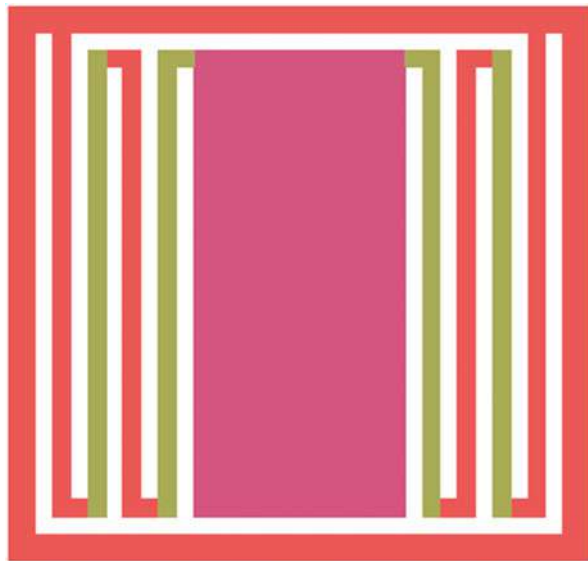
$L_x$	$L_y$	$f_x$	$F_y$	L	g	w	n
57 $\mu\text{m}$	57 $\mu\text{m}$	3 $\mu\text{m}$	3 $\mu\text{m}$	53 $\mu\text{m}$	2 $\mu\text{m}$	2 $\mu\text{m}$	2

**Table 4** parameters of FPA's longitudinal thickness

SiNx	Au(mirror)	Au (deformation leg)
0.3	0.02	0.22

**Table 5** FPA's performance parameters

$G$ (w/k)	$H$	$S_T$ (deg/K)	$D$ (deg/grey)	$\mathfrak{R}$ (grey/K)	$\tau$ (ms)
$1.1 \times 10^{-7}$	$5.5 \times 10^{-3}$	$6.4 \times 10^{-2}$	$8.5 \times 10^{-5}$	1.4	350

**Fig. 56** Optimized FPA layout

which is uniquely caused by full hollow type FPA. Due to the presence of pixel damage in the FPA production process, a small area near the center of the palm formed a blind area; accordingly, there is no thermal response in these areas.

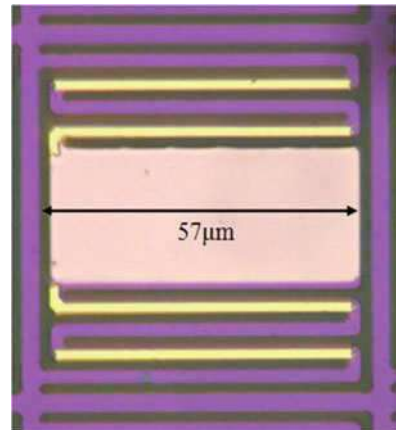
### Responding Sensitivity of FPA and NETD Test

To correctly present the thermal response sensitivity of the optimized monolayer, we test by the methods of temperature-adjusted infrared radiant source system adopted

**Fig. 57** The whole photograph of FPA pixel at size of  $57 \times 57 \mu\text{m}^2$



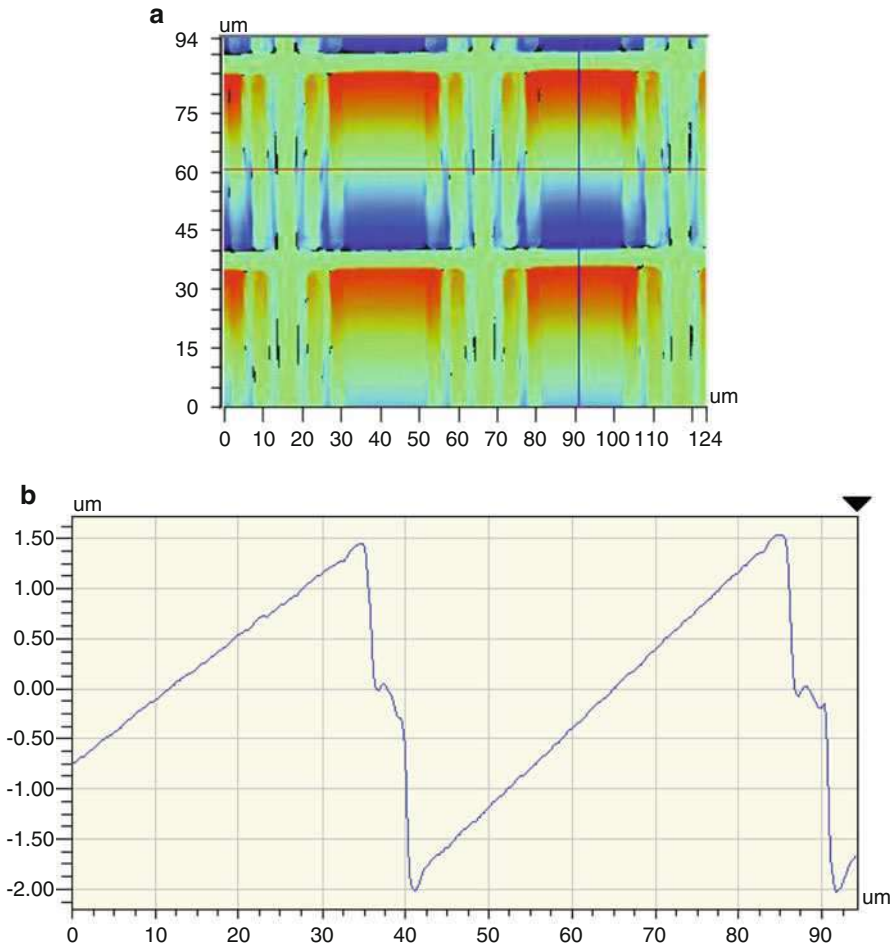
**Fig. 58** FPA pixel microscope photo optimized



in section “[The Examination of and the Improvement on the FPA Unit Model.](#)” Figure 61 is the histogram of the responding sensitivity of FPA.

Judging from the figure, the average responding sensitivity of optimized FPA reaches 71 gray/KK, even the biggest can reach 198 gray/K.

According to the definition of NETD given in section “[Modeling and Analysis of MEMS Based Optical-Readable Thermal Imaging Technology,](#)” the NETD by the use of infrared imaging system of FPA can be defined. In section “[Modeling and Analysis of MEMS Based Optical-Readable Thermal Imaging Technology,](#)” we discuss that the main noise of the uncooled infrared system read by the light is the noise of the optical path system which is not concerned with FPA (the intensity is of 8 grayscales). Combing the noise of system and the responding sensitivity into the definition of FPA, the average NETD of infrared imaging system with the help of the optimized FPA is 112 Mk. This result has reached the property of uncooled infrared imaging system for primary commercial use.



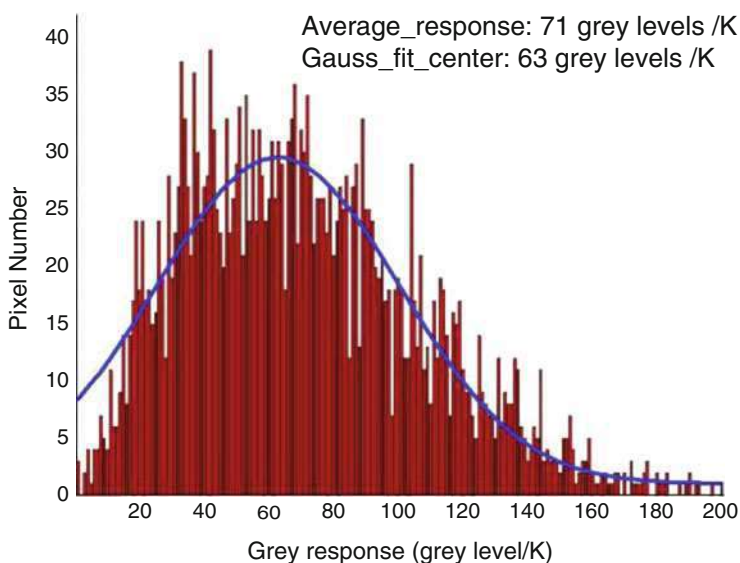
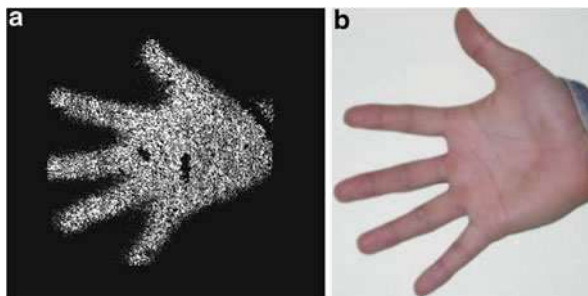
**Fig. 59** Deformation of optimized reflector. (a) 3D view. (b) Morphology along 3D pattern longitudinal lines

Due to the fact that FPA is processed in the lab of poor condition, there exists some blind pixels and the homogeneity is poor. By the methods of controlling process conditions, better homogeneity and imaging result of uncooled FPA read by the light can be gained.

### Test of the Responding Time of FPA

As a major parameter in the infrared imaging system, the responding time is also tested in the experiment. By using the testing method of responding time in section “The Examination of and the Improvement on the FPA Unit Model” (Fig. 32), the thermal response has been gained (Fig. 62).

**Fig. 60** IR thermal image of a human hand. (a) IR thermal image. (b) Visible light image of a human hand

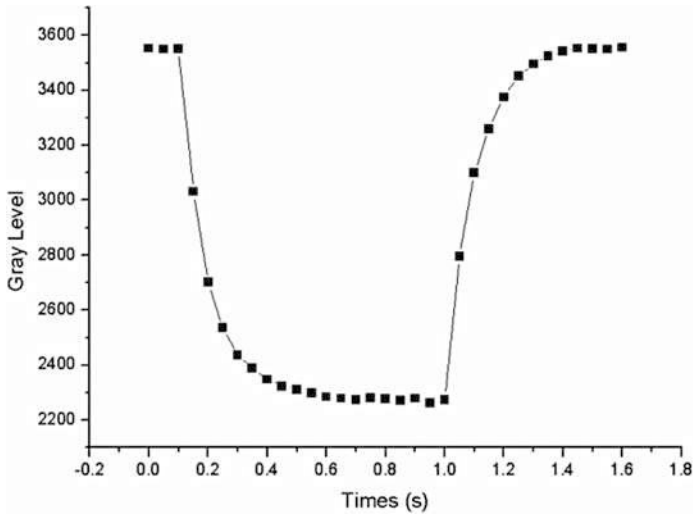


**Fig. 61** Histogram of the responding sensitivity of optimized  $57 \times 57 \mu\text{m}^2$  sized FPA

According to the test result, the FPA thermal response time is 0.53 s, slightly larger than 0.35 s which is estimated by the former model. This kind of result may be due to the fact that the number of elements affected by actual interference is greater than the estimated number of 25.

### Summary of Experimental Results

According to the previous experiment, the optimized single-layer FPA (pixel size  $57 \times 57 \mu\text{m}^2$ , NETD  $\approx 100\text{mK}$ ) overcomes the trend that the response sensitivity decreases with the pixel size, and it has reached the practical requirement in the aspect of spatial resolution and response sensitivity. This result proves the practical prospect of MEMS-based optical readout thermal imaging technique. And with the improvement of the performance of FPA devices and the suppression of the noise of the optical readout system, optical readout thermal imaging technique likely goes



**Fig. 62** FPA gray response changes over time

into commercial application which is the equivalent of electrical solitary thermal imaging system.

## Summary

This chapter discusses the optimization scheme of single-layer full hollow FPA's response performance, on the basis of validated and improved FPA models. The problems are analyzed such as the relationship between number and performance of folding beam in fixed pixel size, and the effects of thicknesses of SiN<sub>x</sub> film and Au film on their properties. Especially the effect of reflector's deformation on FPA performance was researched, and this problem was resolved by means of stress matching. In the meantime, based on the proposed pixel interference model, the inter-pixel interference is suppressed effectively by increasing the thermal isolation between the pixels. Summarizing many aspects of analysis above, a single pixel size of  $57 \times 57 \mu\text{m}^2$  full hollow FPA is designed and implemented, with the FPA pixel number  $160 \times 160$ . According to tests of the noncooled light readout system which is applying this FPA, the average NETD of FPA reached to 112 mk, with the response time 0.53 s. This FPA has reached or approached the existing commercial noncooled infrared imaging system in terms of spatial resolution and temperature sensitivity.

**Acknowledgements** This work is supported by NICT,CAS and "[http://www.baidu.com/link?url=LDwM5K5OKamg0-zuCh1A-Bnpo19wOidu0EkVRdYW80sLccZiTfduEslrb18S\\_ETcRpitwfDBPFcRe74jDBBAU6LkitoxdVeulksOPEYK1a&wd=&eqid=adf557db00013a53000000035a30e269](http://www.baidu.com/link?url=LDwM5K5OKamg0-zuCh1A-Bnpo19wOidu0EkVRdYW80sLccZiTfduEslrb18S_ETcRpitwfDBPFcRe74jDBBAU6LkitoxdVeulksOPEYK1a&wd=&eqid=adf557db00013a53000000035a30e269)"Academy of Integrated Circuit Innovation, CAS.

## References

- Barnes JR, Stepheson RJ, Woodburn CN, O'Shea SJ, Welland ME (1994) A femtojoule calorimeter using micromechanical sensors. *Rev Sci Instrum* 65(12):3793–3798
- Choi J, Yamaguchi J, Morales S, Horowitz R, Zhao Y, Majumdar A (2003) Design and control of a thermal stabilizing system for a MEMS optomechanical uncooled infrared imaging camera. *Sens Actuators A* 104:132–142
- Datskos PG (2000) Micromechanical uncooled photon detectors. *Proc SPIE* 3948:80–93
- Datskos PG, Rajic S, Datskou I (1998) Photoinduced and thermal stress in silicon microcantilevers. *Appl Phys Lett* 73(16):2319–2321
- Dong F et al (2005) Uncooled infrared focal plane array imaging system based on bi-materials cantilever microstructures-design and fabrication. *J Infrared and Millimeter Waves* 24:409–413
- Dong F, Zhang Q, Chen D, Pan L, Guo Z, Wang W, Duan Z, Wu X (2006) An uncooled optically readable infrared imaging detector. 133:236–243
- Duan Z (2003) Incoherent optics application on reconstruction of step object shape and bi-materials cantilever infrared imaging system. Master thesis, University of Science and Technology of China
- Duan Z-H, Zhang Q-C, Wu X-P et al (2003) Uncooled optically readable bimaterial micro-cantilever infrared imaging device. *Chin Phys Lett* 20:2130–2132
- Duan Z et al. Invisible light optical imaging method and system. Patent ZL03132258.1
- Ishizuya T et al (2000) US Patent 6,080,988
- Ishizuya T, Suzuki J, Akagawa K et al (2001) Optically readable bi-material infrared detector. *J Inst Image Inf Telev Eng* 55(2):304–309
- Ishizuya T, Suzuki J, Akagawa K et al (2002) 160 × 120 pixels optically readable bi-material infrared detector. In: *Proceedings of IEEE MEMS*, pp 578–581
- Ishizuya T et al. US Patent 6,339,219 B1
- Jiao B, Li C et al (2007) A novel opto-mechanical uncooled infrared detector. *Infrared Phys Technol* 51(1):66–72
- Jiao B, Chen D et al (2008) Noise modeling analysis on thermal optical uncooled infrared imaging system. *Opto Electron Eng* 35(7):130–135
- Jinhu et al (2006) Noise characterization on reliability of light emitting diode. *J Phys* 3:2261–2264
- Kruse PW (2001) Uncooled thermal imaging. Arrays systems and applications. SPIE Press, Bellingham
- Lai J, Perazzo T, Shi Z et al (1997) Optimization and performance of high-resolution micro-optomechanical thermal sensors. *Sensors Actuators A* 58:113–119
- Li C, Jiao B, Shi S, Ye T et al (2006a) A novel uncooled Substrate-free optical-readable infrared detector: design, fabrication and performance. *Meas Sci Technol* 17:1981–1986
- Li C, Jiao B, Shi S, Ye T et al (2006b) Uncooled infrared focal plane array based on MEMS technology. *J Semiconductors* 27:150–155
- Liu Z (1991) Vacuum technology and equipment. Southeast University press, Nanjing
- Liu H (1992) Mechanics of materials theory, 3rd edn. Higher Education Press, Beijing
- Manalis SR, Minne SC, Quate CF et al (1997) Two-dimensional micromechanical bimorph arrays for detection of thermal radiation. *Appl Phys Lett* 70(24):3311–3313
- Mao M, Perazzo T, Kwon O et al (1999a) Infrared vision using an uncooled thermo-opto-mechanical camera: design, microfabrication, and performance. In: *Proceedings of IEEE MEMS conference*, Orlando, pp 100–105
- Mao M, Perazzo T, Kwon O et al (1999b) Direct-view uncooled micro-optomechanical infrared camera. In: *Proceedings of IEEE MEMS*, pp 100–105
- Miao Z et al (2006) Room temperature infrared imaging system based on bi-materials cantilever microstructures. *Acta Physica Sinica* 7:132–138
- Miao Z et al (2007) An optical readout method for microcantilever array sensing and its sensitivity analysis. *Opt Lett* 32:594

- Min Y-H, Kim Y-K (2000) In situ measurement of residual stress in micromachined thin films using a specimen with composite-layered cantilevers. *J Micromech Microeng* 10:314–320
- Mo D (1991) Solid optics. Higher Education Press, Beijing
- Norton P., Mao M., Perazzo T. et al (2000) Micro-optomechanical infrared receiver with optical readout-MIRROR, Proc. of SPIE, pp 72–78
- Oden PI, Datskos PG, Thundat T et al (1996) Uncooled thermal imaging using a piezoresistive microcantilever. *Appl Phys Lett* 69(21):3277–3279
- Perazzo T, Mao M, Kwon O et al (1999) Infrared vision uncooled microoptomechanical camera. *Appl Phys Lett* 74(23):3567–3569
- Sarid D (1991) Scanning force microscopy. Oxford University, New York
- Saulson R (1990) Thermal noise in mechanical experiments. *Phys Rev D* 42(8):2437
- Shi S (2007) Technology research for design and simulation of uncooled infrared focal plane array device and system. Doctoral thesis, IMECAS
- Suzuki et al. US Patent 6,469,301 B1
- Ueno M, Kosasayama Y, Sugino T et al (2005) 640×480 pixel uncooled infrared FPA with SOI diode detectors. *Proc SPIE* 5783:566–577
- Varesi J, Lai J, Perazzo T et al (1997) Photothermal measurements at picowatt resolution using uncooled micro-optomechanical sensors. *Appl Phys Lett* 71(3):306–308
- Wachter EA, Thundat T, Oden PI, Warmack RJ (1996) Remote optical detection using micro-cantilevers. *Rev Sci Instrum* 67(10):3434–3439
- Weibing W, Ye T, Chen D et al (2003) Uncooled infrared focal plane array Imaging system based on bi-materials cantilever microstructures. *Micro-Nano Electron Technol* 7:404–409
- Weibing W, Chen D, Ye T et al (2004) Uncooled infrared focal plane array Imaging system based on opto mechanical effect. *Laser Infrared* 34:83–86
- Yang S (1987) Heat transfer, 2nd edn. Higher Education Press, Beijing
- Yilin S (1987) Mechanics of materials, 2nd edn. Higher Education Press, Beijing
- Zhang L, Yang G (2001) Design, simulation and testing on a light modulating thermal image device. *J Micromech Microeng* 11:85–93
- Zhao Y, Mao M, Majumdar A (1999) Application of Fourier optics for detecting deflection of infrared-sensing microcantilever arrays. *Microsc Thermophys Eng* 3:249–251
- Zhao Y, Mao M, Horowitz R et al (2002) Optomechanical uncooled infrared imaging system: design, microfabrication, and performance. *J MEMS* 11:136–146
- Zhuang Y (1993) Noise and low noise technology of semiconductor device. National Defense Industry Press, Beijing



---

**Part VIII**

**Power MEMS**



# MEMS Direct Methanol Fuel Cells

Yufeng Zhang, Weijian Yuan, Rui Xue, and Xiaowei Liu

## Contents

Introduction .....	1268
Design .....	1270
Modeling .....	1270
Fabrication .....	1274
MEA Preparation .....	1274
Current Collector Processing .....	1277
Encapsulation and Operation .....	1281
Encapsulation of DMFC .....	1281
A Novel Assembly Method of DMFC .....	1282
The Effect of Assembly Pressure on DMFC .....	1283
Test and Operation .....	1285
Conclusion .....	1295
References .....	1295

## Abstract

In this chapter, MEMS direct methanol fuel cell, a clean portable energy that has been widely concerned in recent years, is introduced. Starting from the working principle and basic structure of DMFC, we elaborate the design, manufacture, encapsulation, and testing of DMFC based on our previous work. An anode three-dimensional steady-state physical field-coupling model was established as an example of simulation and theoretical analysis of DMFC. The design and processing of the current collector were then demonstrated. As the core of DMFC, the fabrication of membrane electrode assembly was focused on where the

Weijian Yuan and Rui Xue contributed equally to this work.

Y. Zhang (✉) · W. Yuan · R. Xue · X. Liu

MEMS Center, Harbin Institute of Technology, Harbin, China

e-mail: [yufeng\\_zhang@hit.edu.cn](mailto:yufeng_zhang@hit.edu.cn); [vivyuanweijian@126.com](mailto:vivyuanweijian@126.com); [13B921025@hit.edu.cn](mailto:13B921025@hit.edu.cn);

[lxw@hit.edu.cn](mailto:lxw@hit.edu.cn)

preparation and measurement of Pt-based catalyst were also covered. A novel assembly method was taken as an example to illustrate the encapsulation methods and requirements for DMFC. The final part described the testing system and methods of DMFC performance, including the measurement and characterization of the catalyst. The effect of different operating conditions on the performance of DMFC was discussed. In summary, the chapter gives a comprehensive view to those who are interested in MEMS DMFC.

---

**Keywords**

Direct methanol fuel cell · Membrane electrode assembly · Current collector · Pt-based catalyst · Proton exchange membrane · Gas diffusion electrode · Anode · Cathode · Model establishment · Encapsulation

---

## Introduction

Energy is the driving force for the development of the national economy, as well as the indispensable material for defense construction. Human society cannot develop rapidly without the innovation and development of advanced new energy technology. In the past few years, the demand for the efficient and renewable energy sources is increasing rapidly with the functionalities of portable devices. As one type of polymer electrolyte membrane fuel cell (PEMFC), DMFC (direct methanol fuel cell) has been attracting more attentions due to its fast response to dynamic loads, high theoretical energy density, and quick start-up (Morse 2007).

With the development of MEMS (microelectromechanical system) technologies,  $\mu$ DMFC has obviously become the breakthrough point of the micro-power sources in the future due to its reduced size and easy integration along with other micro-devices. As a result, the traditional large-size energy system has been unable to meet the needs of the development of science and technology. Nevertheless, micro-energy technology with smaller size and higher energy density has attracted more and more attention from institutions and governments (Nguyen and Chan 2006). In recent years, the rapid development of MEMS technology provides a new approach for the processing and production of micro-energy devices. Micro-power technology based on MEMS refers to applying MEMS technology to research and develop novel type of micro-power devices or systems with high quality and high power density, on the basis of existing energy acquisition, transfer, and storage technology, called power MEMS. The advanced power MEMS technology has significant advantages in volume, weight, life, energy density, supply speed, reliability, cost, etc., which can effectively solve the energy problems of microelectronic product development (Kamarudin et al. 2009).

According to the way of energy conversion, MEMS micro-energy can mainly be divided into micro-generators, micro-solar cells, microthermal batteries, micro-nuclear cells, and micro-fuel cells based on MEMS technology. Compared with other types of micro-energy, miniature fuel cell has the advantages of higher energy conversion efficiency, higher power continuous power supply capacity, more

environmental friendly, quicker start at a lower temperature, higher reliability, and easier for integration. Power MEMS is qualified to be a new kind of micro-energy with a broad application prospect (Dillon et al. 2004).

As a new type of efficient and clean power device, fuel cell technology is considered to be one of the best choices for the power generation technology of the twenty-first century, which will be widely used in aerospace, military, energy, communications, and transportation fields. According to the different electrolytes that have been used, fuel cells are mainly divided into alkaline fuel cell (AFC), phosphoric acid fuel cell (PAFC), molten carbonate fuel cell (MCFC), solid oxide fuel cell (SOFC), and proton exchange membrane fuel cell (PEMFC). Direct methanol fuel cell (DMFC) belongs to a kind of PEMFC. Compared with other types of fuel cells, DMFC has high energy density and rich source of methanol and is low cost, easy to store and carry, and safe. What is more, DMFC system structure is simple and does not need fuel reforming and purification (Kundu et al. 2007).

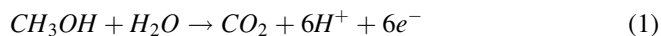
Based on these advantages mentioned above, micro-direct methanol fuel cell ( $\mu$ DMFC) has become a hotspot in research of MEMS micro-fuel cell field all around the world in recent years. As a new energy source, it is very suitable for the future of miniature weapon systems and portable electronic products.

Basically, DMFC includes two current collectors (anode and cathode) and a membrane electrode assembly (MEA) between them. The MEA is generally composed of two diffusion layers, two catalyst layers, and a proton exchange membrane (PEM).

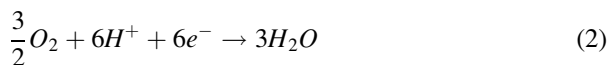
As illustrated in Fig. 1, in the  $\mu$ DMFC electrochemical process, the methanol diffuses from the anode flow field through the diffusion layer to the catalyst layer. The methanol solution is oxidized into protons, carbon dioxide, and electrons. The protons are transferred to the cathode catalyst layer through the PEM, the electrons collected by the anode current collector are transported to the cathode via the external circuit, and the carbon dioxide accumulated is removed with the methanol flow. At the same time, oxygen is converted into water combined with protons and electrons (Hsieh et al. 2008).

During the reaction of DMFC, the specific reaction equations are as follows:

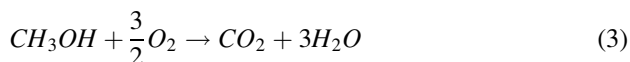
Anode reaction:

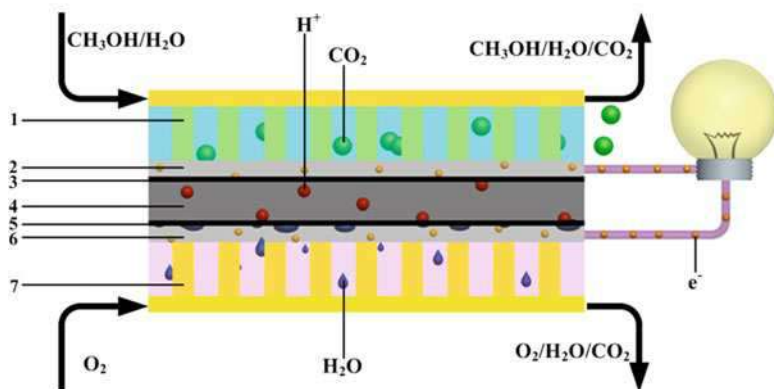


Cathode reaction:



The overall oxidation–reduction process can be defined as:





**Fig. 1** Basic principle of a  $\mu$ DMFC

## Design

### Modeling

Concerning micro-DMFC technologies, anode flow field configurations have been paid increased attention to, which is a crucial factor for micro-DMFC applications. The current collector is an important component of micro-DMFCs, which not only supplies a passage for the reactant transport but also provides structural support for the weak MEA and collects the current. The anode flow field design is critical to yield better performances of micro-DMFCs. At present, studies of anode flow fields mainly include parameter optimizations and new configuration designs.

The design of the cathode current collector significantly influences an air-breathing DMFC performance because of its effects on ohmic and mass-transport losses. The opening ratio and pore shape of DMFC are often used in the cathode current collector design. By establishing DMFC anode three-dimensional steady-state physical field-coupling model, different flow field structures and parameters on cell internal mass-transport process and the influence of heat and current distribution can be analyzed:

### Navier–Stokes (N–S) Equation

Considering that the methanol solution flow in the channel can be classified as a laminar flow, the incompressible Navier–Stokes equations can be applied to describe the momentum transport in the flow field:

$$\rho \frac{\partial \mathbf{u}}{\partial t} + (\rho \mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \nabla \cdot \boldsymbol{\tau} + \mathbf{F} \quad (4)$$

$$\nabla \cdot \mathbf{u} = 0 \quad (5)$$

$$\boldsymbol{\tau} = \eta \left( \nabla \mathbf{u} + (\nabla \mathbf{u})^T \right) \quad (6)$$

where,

$P$  – methanol concentration  
 $u$  – velocity of the liquid phase  
 $p$  – pressure of the liquid phase  
 $\boldsymbol{\tau}$  – liquid viscous stress tensor  
 $F$  – volume force  
 $\eta$  – dynamic viscosity

### Convection and Diffusion Equation

Fick law:

$$-D\nabla c = \mathbf{R} \quad (7)$$

where,

$D$  – diffusion coefficient of methanol in water  
 $c$  – methanol concentration  
 $R$  – reaction rate

The transfer velocity vector in the field is incompressible Navier–Stokes equations, while the velocity vector in the diffusion layer is the speed of Darcy’s law. Thus, the mass transfer equation in flow channels is obtained as follows:

$$\nabla \cdot (-D\nabla c) = \mathbf{R} - \mathbf{u}\nabla \cdot c \quad (8)$$

Methanol diffusion coefficient in diffusion layer is:

$$D = D_{eff} \cdot (\varepsilon)^{1.5} \quad (9)$$

$$D_{eff} = D_{ref} \cdot \exp \left( 2436 \left( \frac{1}{353} - \frac{1}{T} \right) \right) \quad (10)$$

where,

$D_{ref}$  – methanol diffusion coefficient in water at 353.15 K  
 $\varepsilon$  – porosity of porous medium  
 $T$  – working temperature

### Darcy’s Law

The velocity of the fluid in the diffusion layer can be modified by Darcy’s law:

$$\nabla \cdot \left( \rho \left( -\frac{\kappa}{\eta} \nabla p \right) \right) = \mathbf{Q}_{darcy} \quad (11)$$

$$\mathbf{u} = -\frac{\kappa}{\eta} \nabla p \quad (12)$$

where,

$\kappa$  – permeability of the liquid in porous medium.

### Ohm's Law

According to Ohm's law, electron conduction through the anode diffusion layer and electric phase potential distribution can be described:

$$\nabla(-\sigma_l \nabla \varphi_l) = -S_a \cdot i_a \quad (13)$$

While the proton transfer in anode diffusion layer is:

$$\nabla(-\sigma_i \nabla \varphi_i) = S_a \cdot i_a \quad (14)$$

where,

$\sigma_i$  – effective conductivity

$\varphi_i$  – anode potential

$S_a$  – specific area

$i_a$  – anode current density

### Butler–Volmer (B–V) Equation

B–V equation is used to describe the anode current density by methanol oxidation reaction in the catalytic layer:

$$i_a = i_0 \cdot \exp\left(0.5F \frac{eta}{(R \cdot T)}\right) \cdot \left(\frac{c}{c_{ref}}\right) \cdot (c \geq 0) \quad (15)$$

where,

$i_0$  – exchange current density

$F$  – Faraday's constant

$eta$  – overpotential

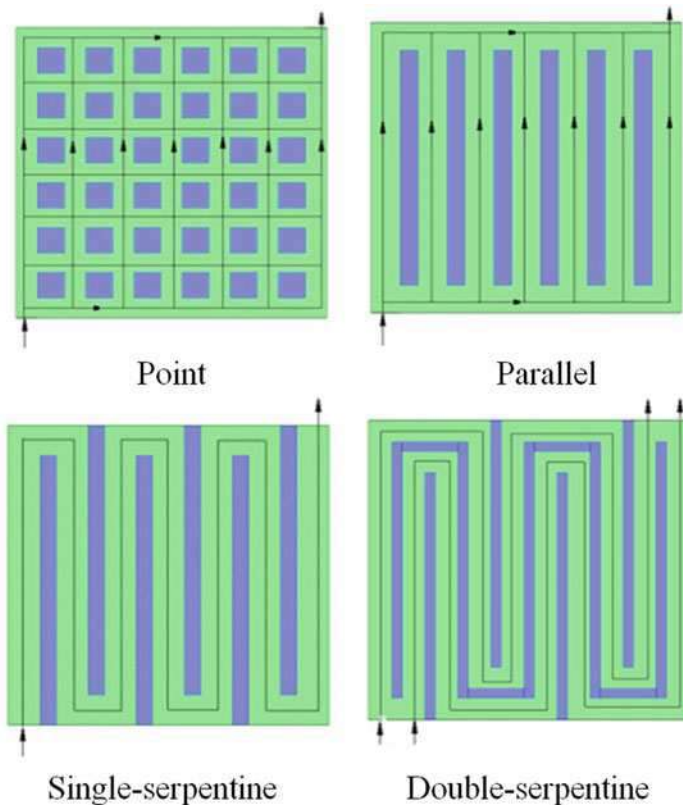
$c_{ref}$  – reference methanol concentration

The parameters are defined and listed in Table 1.

Figure 2 shows the design of the structure of four different flow fields. Different anode flow fields of air-breathing micro-direct methanol fuel cells ( $\mu$ DMFCs) are investigated to improve the cell performances. The single-serpentine flow field can effectively improve the methanol mass-transport efficiency and exhibit higher

**Table 1** Parameters used in the anode model calculations

Name	Value
Faraday's constant ( $F$ )	96,495 C/mol
Gas constant ( $R$ )	8.314 J/(mol · K)
Cell voltage ( $V_{\text{cell}}$ )	0.2 V
Permeability ( $\kappa$ )	$1.2 \times 10^{-12} \text{ m}^2$
Gibbs free energy $G_a$	$-166.27 \times 10^3 \text{ J/mol}$
Operating temperature ( $T$ )	293 K
Pressure ( $p_0$ )	$1.013 \times 10^5 \text{ Pa}$
Electron conductivity ( $\sigma$ )	10 S/m
Methanol solution density ( $\rho$ )	1000 kg/m <sup>3</sup>
Viscous coefficient ( $\eta$ )	$0.9 \times 10^{-3} \text{ kg/(m} \cdot \text{s)}$
Overtoltage ( $\text{eta}$ )	0.2 V
Enthalpy $H_a$	$-238.66 \times 10^3 \text{ J/mol}$
Porosity ( $\varepsilon$ )	0.6

**Fig. 2** Schematic of four types of flow fields



exhaust resultant ( $\text{CO}_2$ ) rates than other flow fields such as grid, parallel, and double serpentine. Additionally, the effects of open ratios and channel lengths on the cell performance are evaluated to determine the optimal anode flow field structures. The  $\mu\text{DMFCs}$  with different anode flow fields are fabricated using silicon-based micro-electromechanical system (MEMS) technologies and are tested at room temperature. The experimental results show that the single-serpentine flow field exhibits a significantly higher performance than those of other flow fields, demonstrating  $16.83 \text{ mW cm}^{-2}$  in peak power density and a substantial increase in mass-transport coefficients.

---

## Fabrication

### MEA Preparation

#### Preparation of Catalyst

For the anode and cathode reaction of DMFC, Pt exhibits the highest catalytic activity and thus plays the dominant role of all metal catalysts. However, Pt alone cannot be directly used as methanol electrooxidation catalyst. Intermediate species of methanol oxidation, such as  $\text{CO}_{\text{ads}}$ ,  $\text{COOH}_{\text{ads}}$ , and  $\text{CHO}_{\text{ads}}$ , will make Pt particles poisoned and lead to a sharp decrease of DMFC performance. To solve this problem, bifunctional mechanism has been proposed by many researchers, that is, other metal or metal oxide is added to form a binary catalyst (Shim et al. 2012; Liu et al. 2006; Liu and Nørskov 2001). Among all promoters, Ru is the most widely used in commercial catalyst (Liu and Nørskov 2001; Gojkovića et al. 2003).

To fully exploit the catalytic activity of Pt or Pt–Ru and decrease the cost of DMFC, catalyst support is essential. Catalyst support should have the advantage of large surface area for good dispersion of metal particles, high electric conductivity, excellent stability, etc. At present, metal particles with the size of several nanometers dispersed on carbon support are in common use. For instance, carbon black Vulcan XC-72R is used in the fabrication of commercial catalysts Pt/C and Pt–Ru/C for DMFC. Graphene, since discovered in 2004, has become a star material and showed amazing physical and chemical properties (Geim and Novoselov 2007). The two-dimensional monolayer structure of graphene makes it ideal for the support of DMFC catalyst. Much work has been done to investigate this issue, and experiment results show graphene-based catalysts have better catalytic activity and stability than their traditional counterparts (Shuping et al. 2013; Shixin et al. 2012).

There are three typical methods to prepare Pt-based graphene hybrid. One is impregnation reduction. Pt or other metal precursors, such as  $\text{H}_2\text{PtCl}_6$  and  $\text{RuCl}_3$ , are mixed with graphene or graphene oxide. Then a reductant, such as hydrazine hydrate, sodium borohydride, hydroiodic acid, etc., is added carefully to the mixture. The reduction will take place whether in room temperature or in heated circumstance depending on the choice of reductant. Filtered, washed, and dried under vacuum, the catalyst sample obtained is ready to be applied on DMFC. Two is electrodeposition. By applying cyclic voltammetry, square wave scanning, chronoamperometry, and

other electrochemical methods to electrolyte solution of precursors, the corresponding catalyst can be obtained. Potential, current density, and deposition time are parameters to control the synthesis procedure. However, the particles prepared by electrodeposition are large and it is difficult to fabricate on a large scale. The third way is the microwave-assisted method. It is a new catalyst preparation method developed in recent years with the advantages of simple operation, short preparation period, easy reaction control, and uniform product particles. Microwave irradiation can generate homogenous and rapid heating so the reaction only takes tens of seconds. Besides, the microwave heating can promote the formation of a large number of initial nuclei in the reaction, making the synthesized nanoparticles small and uniform.

Among them, microwave-assisted ethylene glycol (EG) method is widely used. A standard procedure is stated as follows. Firstly, a certain amount of graphene oxide is dispersed uniformly into the mixture of ethylene glycol and isopropyl alcohol (with the ratio of 4:1) by ultrasonic treatment. Then  $\text{H}_2\text{PtCl}_6$ -EG (with or without  $\text{RuCl}_3$ -EG) solution is added and stirred for 3 h. After that, the pH value of the ink is adjusted by  $\text{NaOH}$ -EG solution drop by drop until its value reached 12. Then argon gas is fed into the ink to expel oxygen. The next step is to place the beaker in the center of a microwave oven for consecutive heating time for 1–3 min. After cooling down to room temperature, the diluted  $\text{HNO}_3$  solution is added into the mixture to adjust the pH value to 4. The mixture is kept stirred for 12 h, and then the product will be washed repeatedly with ultrapure water until no  $\text{Cl}^-$  is detected. A solid catalyst sample is obtained after the drying process.

## MEA Fabrication

The design of  $\mu\text{DMFC}$  not only includes the current collectors but also the components of MEA (membrane electrode assembly), which contains anode diffusion layer, anode catalyst layer, Nafion membrane, cathode catalyst layer, and cathode diffusion layer. A novel double-catalytic layered MEA which includes a hydrophilic thin film of inner catalyst and a traditional electrode with an outer catalyst was provided. Both catalyst-coated membrane (CCM) and gas diffusion electrode (GDE) methods were employed for fabricating the new structured MEA. The results showed that the performances of DMFC with such MEA were enhanced.

The Nafion 117 membrane was pretreated in deionized water, 3 wt.%  $\text{H}_2\text{O}_2$ , 3 wt.%  $\text{H}_2\text{SO}_4$ , and deionized water again for 1 h in each solution in turn. The protocol for forming the hydrophilic inner catalyst layer was prepared through the decal transfer method. The catalyst inks (consisting of Pt-Ru/Pt black, 5 wt.% solubilized Nafion, isopropanol, alcohol, glycerol, and deionized water) were uniformly sprayed by a spray gun onto Teflon decal blanks with metal loading about 3.0 and 2.0  $\text{mg cm}^{-2}$  for anode and cathode, respectively. Then the thin film was transferred from Teflon decal blanks to the  $\text{Na}^+$  form of Nafion 117 membranes by hot pressing at 160–200 °C and 5–9 MPa for 150 s. Then the CCM were ion-exchanged to the protonated form by treating in 0.5 M  $\text{H}_2\text{SO}_4$  solution for 2 h and in 80 °C deionized water bath for 1 h. The traditional electrodes were made as follows: the gas diffusion layer (GDL) was made by spraying carbon inks (70 wt.% XC-72, 17 wt.% PTFE, and 13 wt.% Nafion) over

the hydrophobic-treated (Teflon 15 wt.% and pore-former  $\text{NH}_4\text{HCO}_3$ ) carbon paper (TGP-060, Toray Inc.). The electrode was constructed by spraying the catalyst ink over the GDL. The anode catalyst consisted of 40 wt.% Pt–Ru/C, 10 wt.% Nafion, and pore-former  $\text{NH}_4\text{HCO}_3$ . The cathode was made using a similar method with catalyst 40 wt.% Pt/C. For both the anode and cathode, the total metal loading is  $2.0 \text{ mg cm}^{-2}$ . At last, the CCM was sandwiched between the traditional anode and cathode and hot-pressed under the condition of  $135 \text{ }^\circ\text{C}$  and 5 MPa for 180 s to form the novel MEA. This MEA was labeled as M1. To make some comparisons, other two MEAs were prepared and labeled as M2 and M3. Thus, in the double-catalytic layered electrodes, three gradients were formed: catalyst concentration gradient, hydrophilicity gradient, and porosity gradient, resulting in good mass transfer, proton and electron conduction, and low methanol crossover.

Modification at both anode and cathode at the same time has already been achieved. The self-breathing  $\mu\text{DMFC}$  with the spoke cathode structure was fabricated by bulk-silicon MEMS technology. It consisted of two plates (anode and cathode) as current collectors and a membrane electrode assembly (MEA) between them, in a sandwich structure. The cathodes introduced with the different structures and parameters were fabricated using bulk-silicon technology, and with the micro-fabricated flow field, the distribution area played the role of transporting reactants to the diffusion layer of the MEA. To enhance the electronic conductivity, Au layers were sputtered on one side of the bipolar plates. The image of the spoke cathode reveals that there are some burrs existed after the fabrication process. A piece of five-layered MEA with an area of  $0.64 \text{ cm}^2$  was employed in this silicon-based  $\mu\text{DMFC}$ , which was fabricated by the catalyst-coated membrane (CCM) method. First, the hydrophilic catalyst layers were prepared using the decal transfer method to form the CCM, with an anode catalyst layer of Pt–Ru black ( $4.0 \text{ mg cm}^{-2}$ ), a cathode catalyst layer of Pt back ( $4.0 \text{ mg cm}^{-2}$ ), and a Nafion 117 membrane between them. Afterwards, a carbon paper (TGPH-090, Toray Inc.) was prepared with the hydrophobic (10 wt.% PTFE for the anode and 30 wt.% PTFE for the cathode) and pore-formed ( $\text{NH}_4\text{HCO}_3$ ) pretreatment to form the gas diffusion layer (GDL). Finally, the five-layered MEA was achieved with two GDLs hot-pressed on both sides of the CCM at  $130 \text{ }^\circ\text{C}$  and 4 MPa for 120 s. In addition, before operation, the MEA needed an activation process to achieve the optimal performance. First, to humidify the MEA, we fed  $90 \text{ }^\circ\text{C}$  deionized water to the anode side and saturated oxygen ( $80 \text{ }^\circ\text{C}$ , ambient pressure) to the cathode side, and then we maintained the cell at  $80 \text{ }^\circ\text{C}$  for 1 h. Second, to activate the catalysts, with the cell maintained at  $80 \text{ }^\circ\text{C}$ , we fed 2.0 M methanol solution with a flow rate of  $5.0 \text{ mL min}^{-1}$  to the anode and saturated oxygen at  $720 \text{ mL min}^{-1}$  to the cathode. Under these conditions, the cell was kept operating at 0.2 V for 4 h. Thus, the MEA was then completely activated. In this work, a novel encapsulation approach was also designed using polydimethylsiloxane (PDMS), which could easily contact the silicon wafer with constrictions. A polymethyl methacrylate (PMMA) mold was first fabricated and cleaned, and then a PDMS elastomer and curing agent with a vacuumized weight ratio of 10:1 was poured into the face-to-face PMMA mold in a vacuum drying oven

at 65 °C. After 90 min, the current collectors patterned with PDMS were molded. Finally, the holes and windows were created by laser cutting technology.

## Current Collector Processing

The anode current collectors with different flow patterns were fabricated using bulk-silicon MEMS technology. The silicon wafers with  $\langle 100 \rangle$  crystal orientation were utilized as anode current collectors, including grid, parallel, and single-serpentine and double-serpentine channels. First, clean the silicon wafer by ultrasonic, and a 0.8- $\mu\text{m}$ -thick  $\text{Si}_3\text{N}_4$  layer was deposited on the silicon substrates with low-pressure chemical vapor deposition (LPCVD). Photolithography was applied to pattern microchannels on the  $\text{Si}_3\text{N}_4$  layer. The flow channels depth of 240  $\mu\text{m}$  were etched using an anisotropic etching process of 40% KOH solution at 40 °C, then  $\text{Si}_3\text{N}_4$  layer was removed by reaction ion etch method. Considering portable applications, the cathode current collectors with air-breathing perforated structure were employed to fabricate the self-breathing openings with a radius of 0.3 mm on the silicon wafer by laser cut. In fact, the program radius which is set on the laser machine should be a little smaller than 0.3 mm because the remnant heat can increase the radius of opening. To collect current and minimize contact resistance, a Ti/Au (0.05/1.0  $\mu\text{m}$ ) layer was sputtered on the current collectors. During the laser ablation, the incisions on the silicon are burnt by the high energy of laser (Lumonics JK702 Nd: YAG laser, made in the UK); however, this Ti/Au layer-sputtered process can repair the electrical conductivity of incisions.

Configurations and parameters of anode flow fields have great effects on the performance of micro-DMFCs. Utilizing optimum flow fields, the power density of micro-DMFCs can be increased by about 50%. Different flow fields are generally applied in anode configurations of micro-DMFCs, such as parallel channel, serpentine channel, and spiral channel. Owing to the mass-transport shadow region of an under-rib diffusion layer from the channel to the electrode, it results in the fall of methanol transport efficiency and the deterioration of the cell performance. In fact, the pressure difference between adjacent flow channels of the ribs determines the convection and diffusion of the methanol molecules in the electrodes. If the transport rates increase with an increment of the pressure difference, the methanol transport efficiency to the catalyst layer is improved, and the resultant ( $\text{CO}_2$ ) is also exhausted from the flow channels more quickly. Based on the above considerations, we investigate different anode flow fields of self-breathing micro-DMFCs. The single-serpentine flow fields can effectively improve the methanol mass-transport efficiency and exhibit higher exhaust resultant ( $\text{CO}_2$ ) rates than other flow fields such as grid, parallel, and double serpentine. Additionally, the effects of open ratios and channel lengths on the cell performance are evaluated to determine the optimal anode flow field structures. Using silicon-based MEMS technology, the self-breathing micro-DMFCs are fabricated and tested to obtain the optimum structure parameters for single-serpentine flow fields.

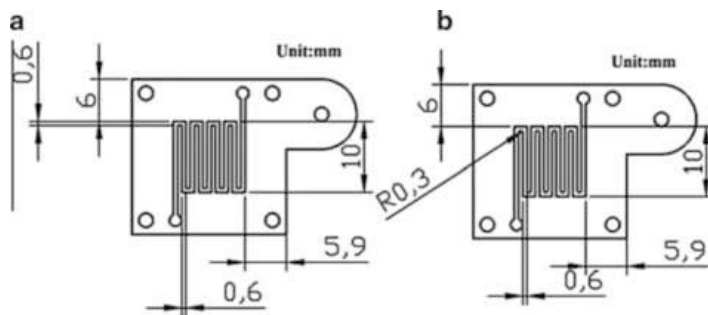
In conclusion, though the cell performance with double-serpentine flow field is a bit higher than that with others, its fabrication process is so complicated, and the improvement of the performance is not obvious. It is concluded that increasing the serpentine channels of anode flow field does not have many benefits on the cell performance.

In addition to silicon, metallic materials can also be applied as current collectors. For example, aluminum (Al) current collectors have been developed for powering portable applications. The anode and cathode current collectors with an area of  $4.5 \text{ cm}^2$  were fabricated on the Al substrates utilizing computer numerical control (CNC) technology. To obtain strong anticorrosion resistance, a 3 mm Au layer was deposited on the current collectors using chemical plating. Compared with the graphite and stainless steel, the characterization of the Au-coated Al current collector was investigated to exhibit superior characteristics in electric conductivity, weight, and electrochemical corrosion resistance. The current collector was applied to a DMFC, and the cell performance was experimentally investigated under different operating conditions. The results indicated that the Au-coated Al current collectors presented in this chapter (Yuan et al. 2012) might be helpful for the development of portable power sources applied in future commercial applications (Yuan et al. 2012).

The effective removal of carbon dioxide ( $\text{CO}_2$ ) in the anode microchannels is of great importance to the performance of a  $\mu\text{DMFC}$ . The gathered bubbles block part of the methanol mass transfer area from anode flow field to catalyst layer, causing a reverse effect on cell performance. Unlike the flow in the straight channel, the gas emission at the corner is more critical because of the corner effect. Dimensions of the plate are illustrated in Fig. 3. The work investigates the two-phase transport between  $\text{CO}_2$  bubble dynamics and micro-fluid at the corner from a mesoscopic perspective. A two-dimensional, nine-velocity Lattice–Boltzmann model coupled with surface tension, buoyancy force, and fluid–solid interaction force is adopted in this work to simulate the dynamics of micro-flow. Simulation results indicate that the emission speed of  $\text{CO}_2$  gas patterned with the circular corner is faster than that with orthogonal corner. To verify the simulation results, a visual study of the  $\text{CO}_2$  bubble behavior in the microchannels of a transparent  $\mu\text{DMFC}$  is conducted, and the cells patterned with different structures are tested. Compared with the cell with orthogonal corner, the cell patterned with circular corner exhibits a substantial increase of 21% in peak power density.

A Lattice–Boltzmann model coupled with liquid–gas surface tension, fluid–solid interaction, and buoyancy force is adopted in this work to simulate the two-component flow with different hydrophilicity channels. The simulation results show that the emission speed of  $\text{CO}_2$  gas in the cell patterned with the hydrophilic channel wall is faster than that with the hydrophobic wall. In order to mutually validate the simulation results,  $\mu\text{DMFC}$  with an active area of  $1.0 \text{ cm}^2$  is designed, fabricated, and tested. Data results show that with improved gas expelling and fuel supplying, the novel  $\mu\text{DMFC}$  with the hydrophilic flow field gets an increase of 15.7% in peak power density and a dramatically lowered polarization at high current density output.

Aluminum alloy LY12 was used to fabricate a perforated cathode current collector. The surface of the current collector, which contacts directly with the cathode gas diffusion electrode, was coated with an Au layer to reduce the contact resistance. The



**Fig. 3** Geometry of the flow fields: (a) orthogonal corner and (b) circle corner

rest surface was treated by plasma electrolytic oxidation to form an insulated, protective, and hydrophilic coating. The PEO treatment was carried out for about 30 min in a sodium silicate solution by a homemade 5 kW AC power supply with the frequency of 1000 Hz. More detailed experiment procedures can be found in some references. After PEO treatment, a gray PEO coating was formed on the surface of the cathode current collector. The phase composition and surface morphology of the coating were characterized, respectively, by scanning electron microscopy and X-ray diffraction. The X-ray pattern shows that the coating is composed of  $\text{Al}_2\text{SiO}_3$  and  $\text{Al}_2\text{O}_3$ . SEM result shows that the coating gives a typical craterlike porous structure resulting from the breakdown, and some pores distribute uniformly across the coating surface.

Because of the highly hydrophilic property of  $\text{Al}_2\text{O}_3$  and the small size of the pores, there is significantly strong capillary action that draws water into the PEO coating, which can make the PEO coating possess the superhydrophilic property. Water droplet spreads quickly as a plane when it contacts the coating surface, and it is quite difficult for water droplets to form on the coating surface. If the coating is in a vertical position, all the water that contacts the coating surface will flow automatically to the bottom due to gravity, and water droplets nearly cannot clog on the coating surface. Consequently, the PEO-treated cathode current collector can prevent the air-breathing holes from water clogging up. In this case, water flooding of the cathode current collector can be avoided. Results of data show that liquid water spreads quickly as a plane on the surface of the porous PEO coating that is composed of  $\text{Al}_2\text{SiO}_3$  and  $\text{Al}_2\text{O}_3$ . As a result, water droplet can be prevented from accumulating along the air-breathing channels of the PEO-treated cathode current collector, and thus the water flooding problem of the micro-passive DMFC is effectively solved. The cell fabricated with the novel cathode structure gives excellent stability.

To resolve the poor mass-transport efficiency problem in the anode flow channel, an N-inputs–N-outputs parallel flow pattern with rectangular convexes to reinforce methanol mass transport and reduce concentration polarization is presented. The simulation results show that the N-inputs–N-outputs parallel flow channels with the rectangle convexes improve the performance.  $\mu\text{DMFCs}$ , which have four anode flow patterns, are fabricated using MEMS (microelectromechanical system)

technology. The experimental results show that the  $\mu$ DMFC with the rectangle convexes has a performance better than previously reported systems.

One problem is the unstable performance during long-time discharge. The second problem is the short operating time. A dual-chamber anode structure is proposed in order to solve two performance problems of the conventional passive micro-direct methanol fuel cell ( $\mu$ DMFC). In this structure, low concentration chamber is filled with methanol solution with appropriate concentration for the  $\mu$ -DMFC. Pure methanol in high concentration chamber diffuses to the low concentration chamber to keep the concentration of methanol solution suitable for long-term discharge of  $\mu$ -DMFC. A Nafion-polytetrafluoroethylene (PTFE) composite membrane is inserted between the two chambers to conduct pure methanol. The experimental results during long-term discharge show that the stable operating time of passive  $\mu$ -DMFC increases by nearly 2.3 times compared to a conventional one with the same volume.

When it comes to the cathode design, water flooding is a serious problem to solve. A self-breathing micro-direct methanol fuel cell with active area of  $0.64 \text{ cm}^2$  has been developed for powering portable applications. A cathode perforated current collector with parallel flow fields is presented in order to improve the cell performance. Compared with the conventional cathode self-breathing structure, the improved one can enhance oxygen transport and reduce water flooding utilizing multiphysics simulations. The stainless steel plates with the thickness of 0.3 mm as current collectors with parallel flow fields have been machined by thermally micro-stamping. For the cathode self-breathing openings, the perforated current collector has been realized using laser drilling. A 500-nm-thick titanium nitride (TiN) layer is deposited onto the surface of current collectors by magnetron sputtering ion plating (MSIP) technology to cover the cracks and prevent corrosion. Peak power density of the  $\mu$ DMFC reaches  $27.1 \text{ mW/cm}^2$  at room temperature with 1.0 M methanol solutions of 1 ml/min. The results presented in this chapter might be helpful for the development of micro-power sources applied in future portable electronic devices.

Metallic is also applied in cathode structure to prevent water flooding. A 316 L stainless steel fiber felt with 0.25 mm thickness was used as cathode current collector. Aluminum alloy LY12 was adopted to fabricate the cathode end plate that had a perforated flow field. At the back side contacting with the current collector, there are several water-collecting channels 0.3 mm in width and 0.5 mm in depth in the active area between every two rows of the holes. At the front side, several water-collecting channels are also fabricated to go through the plate, connecting with the ones at the back side. To make the liquid water flow fluently in the channels, two additional lines of holes were fabricated, respectively, on the top and bottom of the flow field. In the design, liquid water produced inside the porous current collector is expected to be drawn into the water-collecting channels and flow down automatically under the gravity. To achieve this purpose, a highly hydrophilic coating was fabricated on the surface of the cathode end plate by plasma electrolytic oxidation (PEO) technology. The PEO treatment was carried out for about 30 min in a mixed solution of  $10 \text{ g L}^{-1} \text{ Na}_2\text{SiO}_3$  and  $1 \text{ g L}^{-1} \text{ NaH}_2\text{PO}_2$  by a homemade 5 kW AC

power supply with frequency of 1000 Hz. The prepared coating is composed of aluminum silicate and aluminum oxide and possesses a typical craterlike structure resulting from the breakdown.

An air-breathing direct methanol fuel cell with a novel cathode shutter current collector is fabricated. Compared with the conventional circular cathode current collector, the shutter one improves the oxygen consumption and mass transport. The anode and cathode current collectors are made of stainless steel using thermal stamping die process. Moreover, an encapsulation method using the tailor-made clamps is designed to assemble the current collectors and MEA for distributing the stress of the edges and inside uniformly. It is observed that the maximum power density of the air-breathing  $\mu$ DMFC operating with 1 M methanol solution achieves  $19.7 \text{ mW/cm}^2$  at room temperature.

Another kind of improvement made to promote  $\mu$ DMFC involves mesh. The configuration of the metallic  $\mu$ DMFC consists of an MEA, two sheets of stainless steel mesh, two silicone rubber gaskets, two stainless steel current collectors, and two PMMA fixtures. In this  $\mu$ DMFC design, both the anode and cathode current collectors have the same configuration. In order to enhance electronic conductivity and mechanical strength, the current collectors were fabricated on stainless steel plates with a thickness of  $300 \mu\text{m}$  by micro-stamping technology, a batch-fabricated, highly effective, easily operated, and low-cost method for machining metal parts. To enhance the stamped depth, the fabrication process was performed under warm hydromechanical conditions ( $90 \text{ }^\circ\text{C}$ ) with methylsilicone oil as the hydro-medium, in respect that this method effectively increases the limited drawing ratio of stainless steel. The micrograph of the microfabricated flow field with channels of  $1000 \mu\text{m}$  width and  $370 \mu\text{m}$  depth. It is worth mentioning that the cross section of each channel was a trapezoid with an angle of  $100 \pm 1^\circ$  that was shaped by the stamping process. SEM test was conducted. It can be seen that some cracks were present after the fabrication process. To cover them, a 500-nm-thick TiN layer was deposited onto the surface of each current collector using magnetron sputtering ion plating (MSIP) technology. The surface of the channel following application of the TiN layer also protects the metal current collector from electrochemical corrosion.

---

## Encapsulation and Operation

### Encapsulation of DMFC

A typical DMFC is mainly composed of current collectors (anode and cathode) with flow channels and membrane electrode assembly (MEA). Nowadays, improving the cell performance has been an attractive topic for many researchers. However, this improvement not only depends on operating conditions but also has a significant relation with the design and material parameters. Therefore, the cell assembly plays an important role for obtaining the highest cell performance. Different assembly pattern can generate different plate deformation, which can influence the fuel distribution largely. Besides, insufficient assembly pressure may result in fuel



leakage as well as contacting problems, which can decrease the fuel cell performance. If too much pressure is applied, diffusion layers, the most deformable components in a DMFC, would be deformed sharply (Neyerlin et al. 2005; Coppo et al. 2006).

In the literatures, several modeling works have been conducted to investigate the effect of assembly pressure on the cell performance. Most of them mainly focused on the cell with point assembly method by locking with bolts. Lee et al. (2005) used the finite element analysis method to simulate the assembly of a single cell with metallic bipolar. In order to verify the results of the analysis, experimental tests with a pressure film inserted between the current collector and MEA were also conducted. Taymaz and Benli (2010) developed a three-dimensional model to calculate the deformation effect of assembly pressure on the single-channel PEM fuel cell. Their work showed that when assembly pressure increased, contact resistance, porosity, and thickness of the gas diffusion layer (GDL) decreased accordingly, and too much assembly pressure caused the destruction of GDL. Zhou et al. (2009) conducted a study on effects of assembly pressure, operating temperature, and humidity on fuel cell deformation, contact resistance, overall performance, and current distribution. Zhou et al. (2007) also implemented a model to investigate the contact resistance between current collector and GDL in PEM fuel cells. Experimental investigations were also conducted to validate the model. Xing et al. (2010) developed a three-dimensional model to investigate the effect of clamping assembly pressure on GDL properties and PEM fuel cell performance under different operating voltages.

## A Novel Assembly Method of DMFC

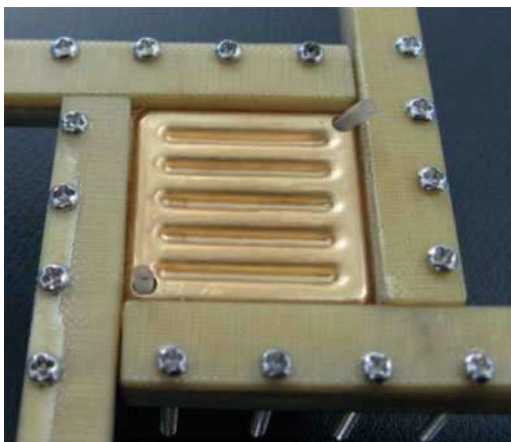
Most studies focused on the models of contact character in PEMFC; only a few attempts were made on the assembly pattern of DMFC. We have proposed a novel assembly approach using tailor-made clamps. This assembly method could control the tightness of assembly, uniformly distribute the stress in the active area, and decrease the weight and bulk of DMFC.

Two PMMA plates were carved and compactly assembled to fabricate the clamps using five bolts. The width of the clamp could be adjusted with the sum thickness of current collectors, gaskets, and MEA. It is easy to control the tightness of assembly and reduce the contact resistance using the novel method, which can improve the cell performance. The assembled cathode-self-breathing DMFC, where a piece of 6.25-cm<sup>2</sup>-five-layered MEA was employed, is shown in Fig. 4.

The performance of this novel assembly DMFC was compared with a traditional one. The assembly pressure is 1.5 MPa. Experimental data reveals that the fuel cell using clamp exhibits better performance with the maximum power density 14.82 mW\*cm<sup>-2</sup>, which is 19.5% higher than its counterpart. Besides, the cell with novel assembly method exhibits higher Faraday efficiency (77.82%) than the cell with conventional structure (71.63%).

The pressure distributions of two assembly methods are shown in Fig. 5: from the stress analysis of the cell with novel assembly (a), there is a uniform force in the

**Fig. 4** Image of sample micro-DMFC

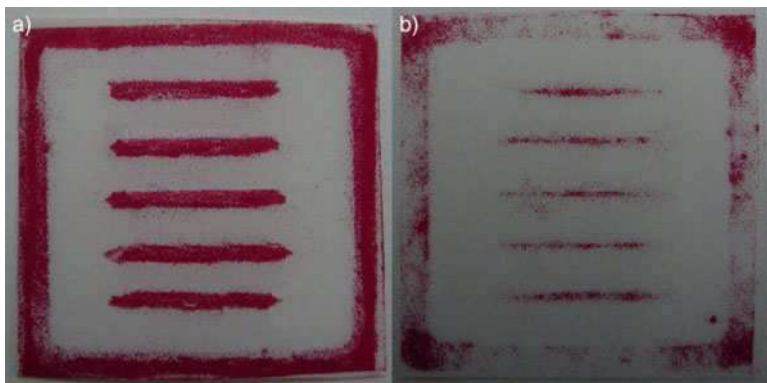


contact area between the plate and membrane electrode, and this uniform force can lead to the even fuel distribution. However, the result of the figure (b) indicates that the cell with bolts generated uneven stress distribution. The edges of the plate are distorted with the pressure of 1.5 MPa at the corner, and there is no good contact area between internal ridge and the MEA, which has a reverse effect on cell performance. At the same time, this new method avoids the phenomenon of concentration polarization caused by the insufficient supply of methanol solution or the change of contact resistance.

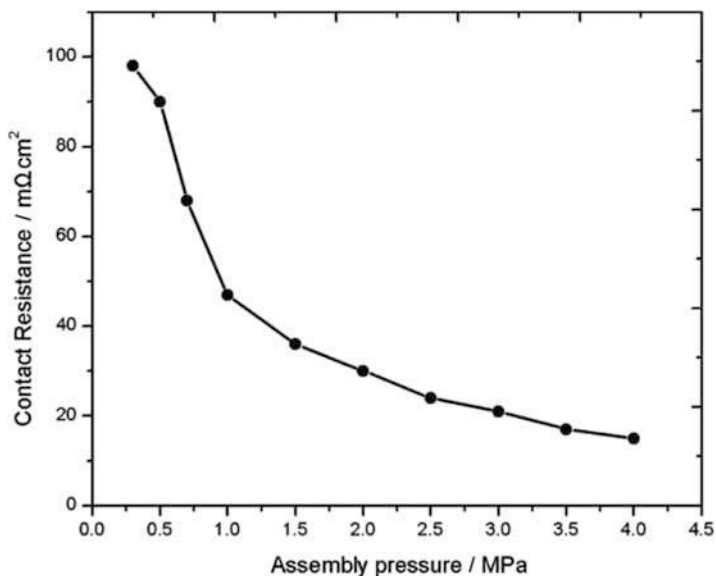
### The Effect of Assembly Pressure on DMFC

To further understand the influence of assembly, the correlation between contact resistance and assembly pressure is presented in Fig. 6. When the contact pressure is 0.3 MPa, the contact resistance is nearly 100 m $\Omega$ , but when the contact pressure is increased to 1.5 MPa, the contact resistance decreases sharply to 35 m $\Omega$ . This is due to the fact that at the early stage, two contacting parts start to contact with each other, and then the real contact area in diffusion layer increases fast. Thus, the contact resistance changes greatly with the initial assembly pressure. On the other hand, when the contact pressure is increased from 1.5 to 2.5 MPa, the contact resistance that occurred in the case of 0.5–1.5 MPa does not decrease at the same rate, which is mainly because the existing contact spots increase slightly at this level. After the contact pressure reaches 3.5 MPa, the contact resistance becomes stable because there is no increase on the contact area after this value.

Although greater assembly pressure can reduce the contact resistance, high pressure could damage GDL and obstruct gas flow. The cell performance was tested with the cell assembly pressure of 0.5, 1, 1.5, 2.0, and 2.5 MPa, respectively, and the I–P curves can be seen in Fig. 7. When the cell performance increased from 0.5 to 1.5 MPa, the maximum power density with 1.5 MPa assembly pressure could reach

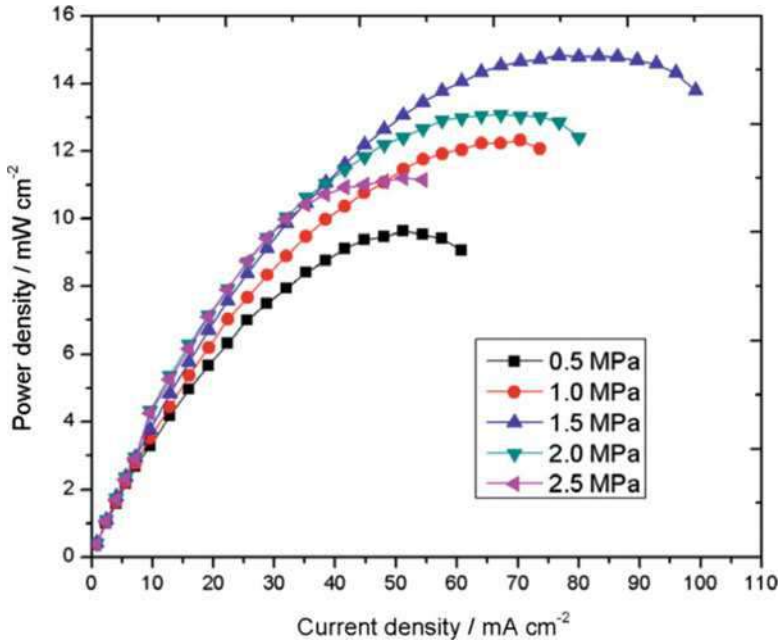


**Fig. 5** Image of stress distribution of two assembly methods: (a) novel assembly pattern and (b) traditional assembly pattern



**Fig. 6** The change of contact resistance under different assembly pressure

14.82 mW cm<sup>-2</sup>. However, the performance was decreased when assembly pressure was increased from 1.5 to 2.5 MPa. Compared with the fuel cell under 1.5 MPa, nearly 24% performance degradation was observed with the assembly pressure of 2.5 MPa. Moreover, it is interesting to notice that, at low current density, there is little difference of output performance with different levels of pressure. The phenomenon indicates that in this region, the pressure has little influence on the performance of DMFC. However, the effect of assembly pressure on performance becomes obvious with the increase of current density.



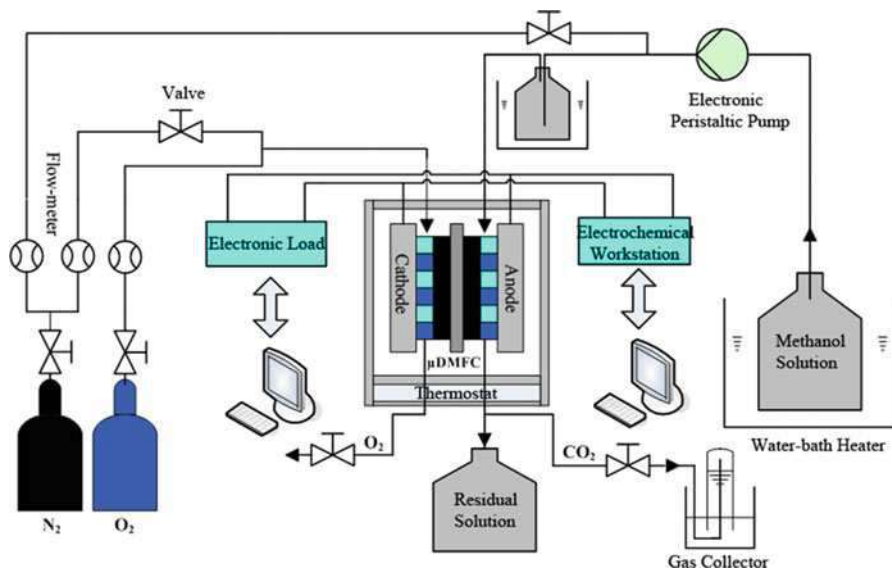
**Fig. 7** Performance comparison of the micro-DMFC with different assembly pressure

There are two aspects of the effect of pressure on DMFC performance: internal resistance and GDL porosity. The resistances of DMFC (GDL resistance and the resistance between the GDL and current collector) decrease with the increase of assembly pressure, which also reduces the internal voltage loss and enhances the output power correspondingly. Moreover, small pressure applied in the first stage is used to overcome the manufacturing tolerance of the cell components. The effect of contact resistance becomes smaller with the high level of pressure, and the effect of GDL mass transfer is gradually displayed. When higher assembly pressure is applied, the GDL is compressed into the flow channels. The exceeded pressure changes the characteristics of the pressure distribution into nonlinear state and may even damage the MEA, and the porosity of GDL would be reduced in this stage. Therefore, the middle-stage assembly pressure could cause the elasticity deformation, and the cell exhibits the best performance.

## Test and Operation

### Testing System and Methods

A DMFC testing system is a platform providing surveillance and control of operating DMFCs with the requirements of excellent reliability and operability. In order to accurately describe the output performance of DMFC, we have built up a testing system with high stability, as illustrated in Fig. 8.

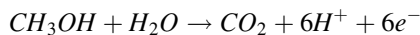


**Fig. 8** Schematic diagram of DMFC testing system

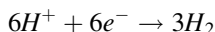
Before the formal test, the prepared MEAs need to be activated first. This is because hot pressing will make the proton exchange membrane (PEM) dehydrate seriously, which leads to a rapid decrease of the conductivity of PEM and thus a poor battery performance. The activation method employed is described as follows. First, the entire MEA was soaked in deionized water at 80 °C for 30 min so that the PEM was sufficiently hydrated. The fully immersed MEA was then assembled in an active DMFC for discharge activation. The operating conditions of activation are as follows: the anode methanol concentration is 2 mol/L with the flow rate of 10 ml/min; the cathode is fed with 720 ml/min humid oxygen. The operation voltage is 0.2 V and the activation continues for 4 h.

Output characteristics are the most direct means of DMFC performance evaluation. The basic principle is to connect the anode and cathode of fuel cells to the test system at both ends so that the power generated by electrochemical reaction can drive external electronic loads. In the output performance test, the main parameters such as open circuit voltage, discharge voltage, limiting current, discharge current, and power density are tested and evaluated. During the test process, by adjusting the operating current, the voltage dynamic response of fuel cell can be observed.

Electrochemical measurement is another major way to evaluate DMFC characteristics, which includes methanol crossover test and electrochemical impedance spectroscopy (EIS). In the methanol crossover test, the cathode is fed with humidified, high-purity nitrogen gas. The oxidation reaction of the methanol penetrated to cathode side is as follows:



whereas the reaction at anode can be illustrated as:



Methanol crossover is measured via the limiting current produced by methanol oxidation at the cathode. The methanol solution was provided to the anode, while humid nitrogen with flow rate 50 ml/min was passed through the cathode. Humidification temperature is 10 °C higher than the fuel cell temperature. Linear sweep voltammetry (LSV) under open circuit is conducted on an electrochemical workstation. The working electrode is the cathode of fuel cell, and the anode serves as the reference electrode and the counter electrode at the same time. The scanning potential ranges from 0 V to 1.0 V with a scan rate of 2 mV/s.

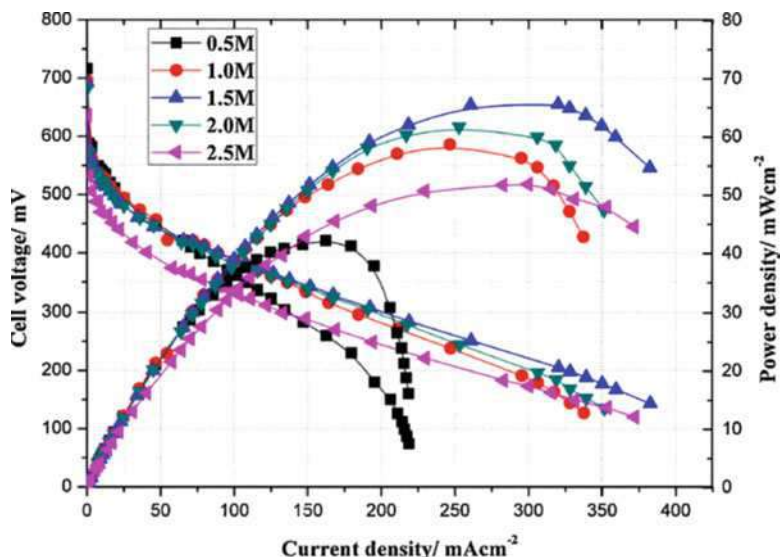
EIS is also performed on the electrochemical workstation. Different frequencies of AC signal are applied to the electrochemical system, making the electrode potential float around the equilibrium potential, and the electrochemical performance of DMFC can be characterized by the change of impedance with frequency. The EIS of DMFC is composed of high-frequency arc, intermediate-frequency arc, and low-frequency arc. The high-frequency arc is used to characterize the ohmic resistance of the cell, which reflects the difficulty of electron and proton transmission. The intermediate-frequency arc is used to characterize the dynamic properties of methanol oxidation and oxygen reduction. The low-frequency arc is used to characterize the mass transfer impedance of reactants. During the test, DC electronic load is in control of constant current discharge of DMFC. The frequency scanning range of EIS is 100 kHz–1mHz. AC signal amplitude is 10 mV.

Apart from the characteristics of MEA and current collectors, operating conditions (such as anode flow rate, concentration, and operating temperature) are also critical to the output performance of DMFC. Therefore, significant attentions have been devoted to the effects of operating parameters on DMFC performance recently (Silva et al. 2012; Tafaoli-Masoule et al. 2014).

Here we use a metallic active DMFC with MEA area of 0.64 cm<sup>2</sup> as illustration to demonstrate the impact of operational parameters on DMFC performance.

### Effect of the Methanol Concentration

The influence of the methanol concentration on cell performance is shown in Fig. 9. Related experiments were carried out with methanol solutions of five different concentrations (0.5, 1.0, 1.5, 2.0, and 2.5 M) fed at the same flow rate of 1.0 ml min<sup>-1</sup> and at 40 °C. It can be seen that the best performance was found for a methanol concentration of 1.5 M and the maximum power density was 65.66 mW cm<sup>-2</sup>, while the 0.5 M methanol solution gave the worst performance. The possible reasons are explained as follows. When a higher methanol concentration was utilized, excess methanol remained after the electrochemical reaction,



**Fig. 9** Performance curves of the DMFC with different methanol concentrations

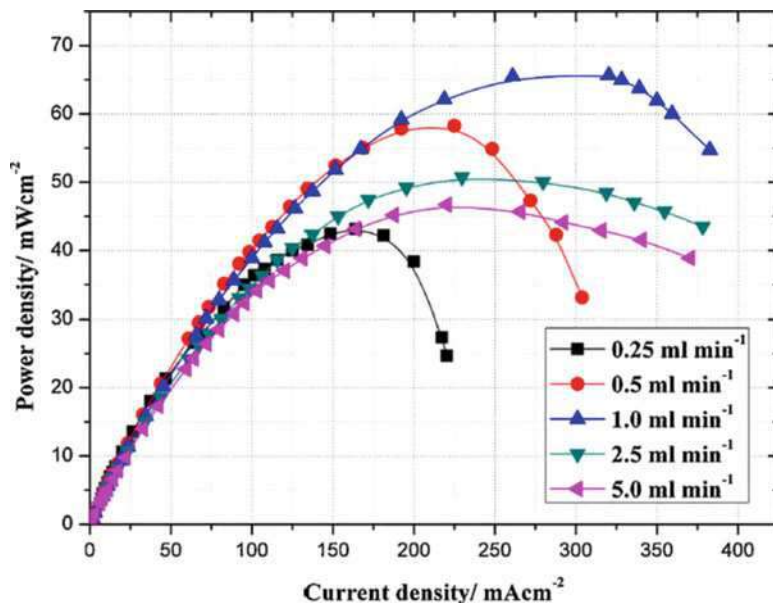
inducing methanol crossover and generating an overpotential at cathode, which degraded the cell performance. On the contrary, a lower methanol concentration could not provide sufficient reactants to join the reaction, especially at high current densities. It is also noted that the optimal methanol concentration of 1.5 M validated in this work is higher than the concentration of around 1.0 M, which was more suitable for DMFCs with larger sizes (Silva et al. 2012; Tafaoli-Masoule et al. 2014). This is mainly due to the fact that the methanol transport in micro-sized DMFCs is more difficult, so a higher methanol concentration is in demand.

### Effect of the Anode Flow Rate

To investigate the effect of the anode flow rate, the DMFC was tested by using 1.5 M methanol solution at five different flow rates (0.25, 0.5, 1.0, 2.5, and 5.0 ml min<sup>-1</sup>) and at 40 °C. The testing results are displayed in Fig. 10. As seen from this figure, when the flow rate increased from 0.25 to 1.0 ml min<sup>-1</sup>, it is clear that the cell performance had a significant improvement accordingly. The reasonable explanations include two aspects. For one thing, an increment of the anode flow rate resulted in an increment of the methanol transport rate, which was effectual to overcome the mass transfer resistance induced by both the GDLs microporous structure and the reversed flow of CO<sub>2</sub> gas bubbles. For another, since the channels were narrow, CO<sub>2</sub> bubbles could easily occupy the anode flow field and disturb the liquid flow. As the anode flow rate increased, the removal rate of CO<sub>2</sub> bubbles would increase accordingly, leading to the decrement of the quantities and sizes of bubbles. These two aspects were both beneficial for the cell performance.

On the contrary, as shown in Fig. 10, when the anode flow rate continuously increased to 5.0 ml min<sup>-1</sup>, the cell performance was found to deteriorate. In fact, as the key factors



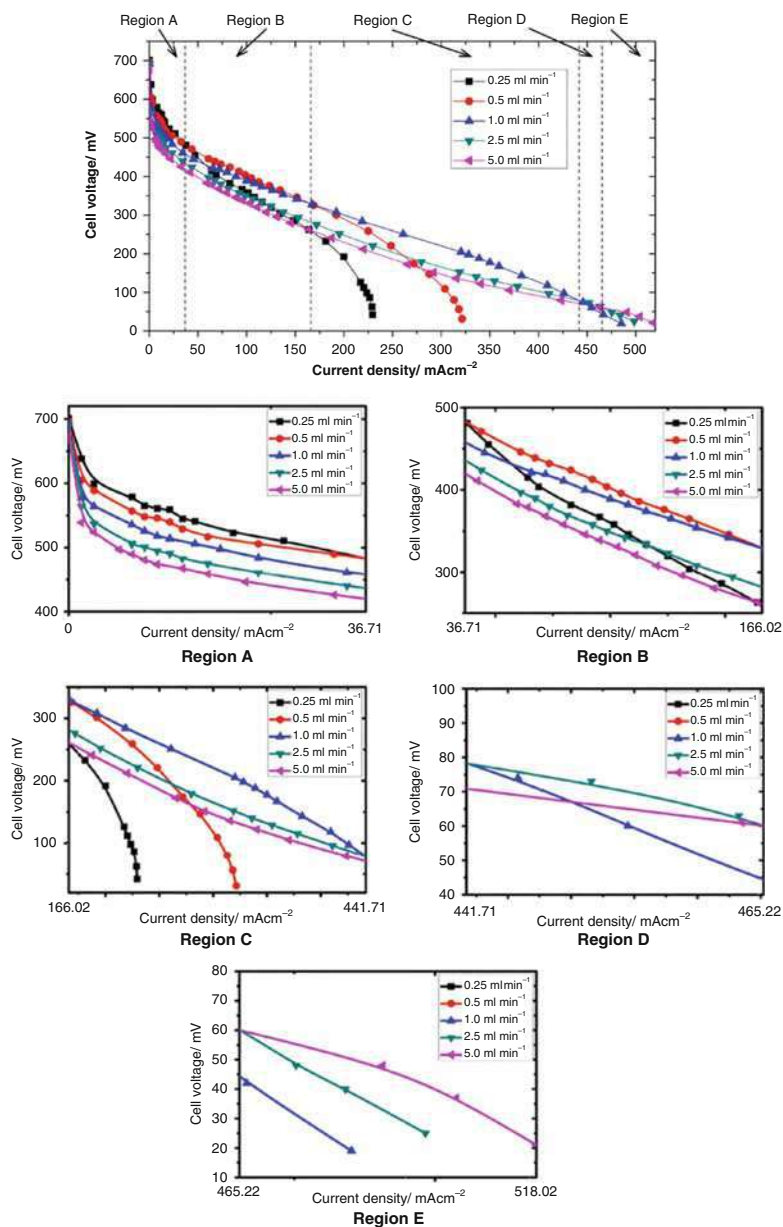


**Fig. 10** Power density curves of the DMFC with different anode flow rates

that determined cell performance, mass transport of methanol and removal of  $\text{CO}_2$  gas bubbles were replaced by methanol crossover, mass transport of oxygen, and heat loss when the DMFC is operated under higher anode flow rates. To methanol crossover, the permeating rate increased with the increment of the anode flow rate. Because the acceleration of the flow rate was ascribable to the enhancement of the static pressure which was brought by an external pump, this enhanced static pressure aggravated the methanol crossover. It is essential to note that when the static pressure became rather high, it could also provide a barrier of high resistance to the mass transport of oxygen in cathode. Moreover, because the reaction heat was released with the methanol solution, the rate of heat loss also increased with the anode flow rate, giving rise to a declination of the electrochemical kinetics. In a word, when the anode flow rate increased from 1.0 to 5.0  $\text{ml min}^{-1}$ , the methanol crossover and heat loss were both impelled, while the mass transport of oxygen was restrained; thus, the cell performance declined.

On the basis of the analysis above, we come to a conclusion that the effect of the anode flow rate is achieved by manipulating some factors that are crucial to cell performance, including mass transport of methanol and oxygen, removal of  $\text{CO}_2$  gas bubbles, methanol crossover, and heat loss. As a matter of fact, the leading factor that determines cell performance alternates with external loads, that is to say that the optimal anode flow rate is changeable at different current densities. Figure 11a shows further polarization testing results by enhancing the current close to the limiting value. It can be seen that this figure is divided into five regions (represented by A–E), which are classified according to five current density intervals dominated by the five anode flow rates, respectively. Figure 11b shows enlarged images of the five regions.





**Fig. 11** Polarization curves of DMFC under different anode flow rates with the external load increasing from unloading to nearly zero: (a) image of the entire polarization performance with five different regions and (b) five enlarged images of the regions

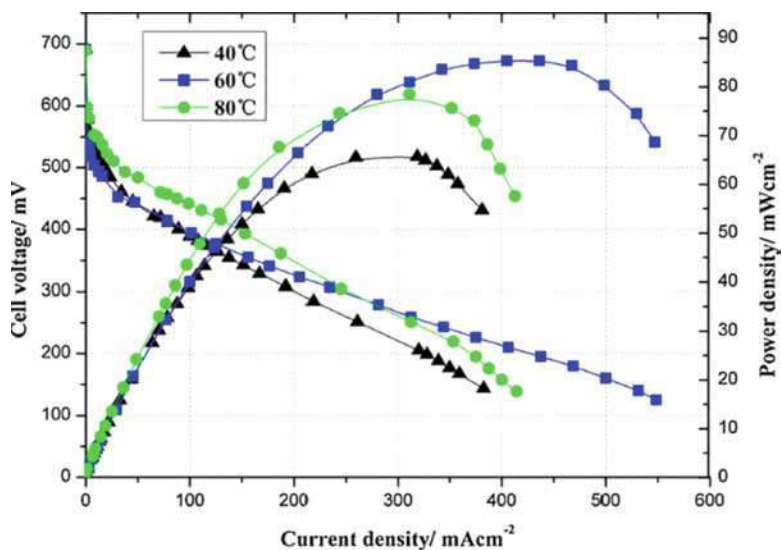
As can be seen, although  $1.0 \text{ ml min}^{-1}$  gave the best performance in general, the flow rate that brought about the highest power density was not consistent in different regions. In region A ( $0\text{--}36.71 \text{ mA cm}^{-2}$ ), in respect that the methanol crossover was very serious and the required reactants, products, and reaction heat were all of small amounts,  $0.25 \text{ ml min}^{-1}$ , which led to the lowest methanol permeating rate, performed the best. In region B ( $36.71\text{--}166.02 \text{ mA/cm}^{-2}$ ) and region C ( $166.02\text{--}441.71 \text{ mA cm}^{-2}$ ), with the weakening of methanol crossover and the strengthening of other factors,  $0.5$  and  $1.0 \text{ ml min}^{-1}$  exhibited the best performance, respectively. When the current density increased to a high magnitude belonging to region D ( $441.71\text{--}465.22 \text{ mA cm}^{-2}$ ) or region E ( $465.22\text{--}518.02 \text{ mA cm}^{-2}$ ), the violent reaction brought about large amounts of consumed reactants and produced  $\text{CO}_2$  gas and reaction heat, while little methanol permeated to cathode, and, therefore,  $2.5$  and  $5.0 \text{ ml min}^{-1}$  were in the ascendant, respectively.

### Effect of the Cell Temperature

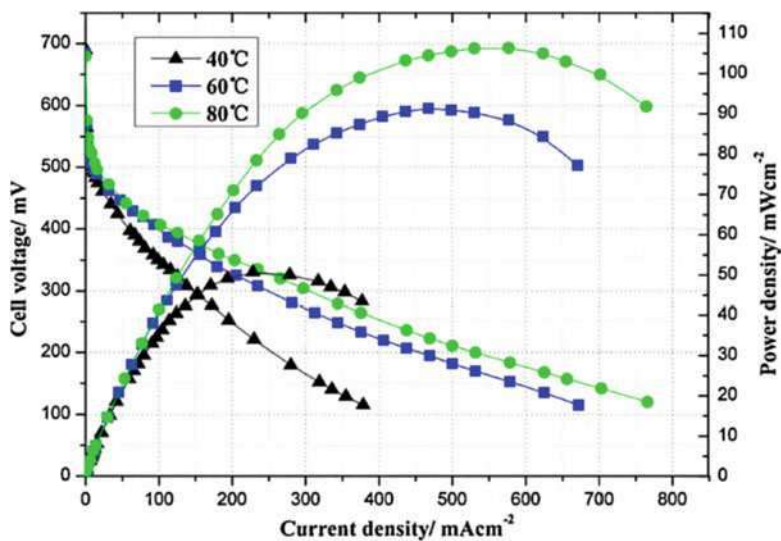
Controlling the operating temperature is critical to the improvement of DMFC performance. It is commonly believed that an increase in the operating temperature can lead to a marked increment in cell performance. The probable benefits of increasing the temperature are (1) increased activity of the catalysts, (2) reduction in the PEM resistance, and (3) increased mass transport in the cell. However, in the course of our experiments, we found that the variation in cell performance with changes in the operating temperature was not uniform. The experiments were carried out under the aforementioned optimal operating conditions, and the adopted temperatures were  $40$ ,  $60$ , and  $80 \text{ }^\circ\text{C}$ . Figure 12 shows the testing results. Based on our prior understanding of the benefits of raising the temperature, the cell performance was expected to be the best at  $80 \text{ }^\circ\text{C}$ , but, in fact, this was not so. As seen from Fig. 12, although  $80 \text{ }^\circ\text{C}$  was in the ascendant when the current density was less than  $260.92 \text{ mA cm}^{-2}$ , once the current density exceeded this value,  $60 \text{ }^\circ\text{C}$  gave better performance. It appeared likely to us that this unexpected temperature dependence was a consequence of the anode flow rates ( $1.0 \text{ ml min}^{-1}$ ) being insufficient to provide either an adequate methanol transport rate for the rapid reaction or an efficient removal rate of the large amounts of  $\text{CO}_2$  bubbles generated at high current densities. To test this hypothesis, we increased the anode flow rate to  $2.5$  and  $5.0 \text{ ml min}^{-1}$  while keeping the other operating conditions remaining constant. The testing results are revealed in Figs. 13 and 14. At a higher flow rate, the cell performance at  $80 \text{ }^\circ\text{C}$  was much better than that at  $40 \text{ }^\circ\text{C}$  or  $60 \text{ }^\circ\text{C}$ . Moreover, the higher the flow rate, the larger the performance difference. This finding is consistent with our hypothesis. We thus conclude that in order to ensure the best cell performance at different operating temperatures, the anode flow rate must be regulated. In addition, it is worth noting that, as shown in Fig. 14, the maximum power density reached  $115.0 \text{ mW cm}^{-2}$  at  $80 \text{ }^\circ\text{C}$ , a remarkable performance for DMFCs.

### Characterizations and Measurements of Catalyst

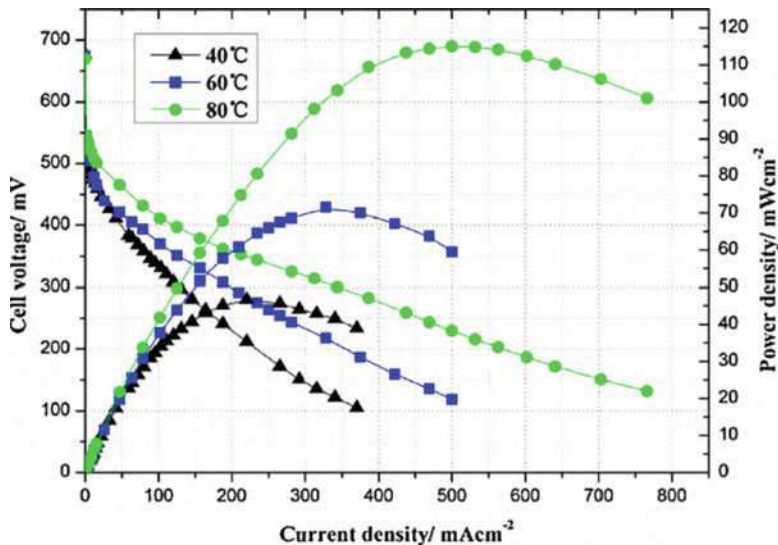
As the core of MEA, the activity of catalyst directly influences the performance of DMFC. How to boost the activity of DMFC catalyst has always been a focus of



**Fig. 12** Performance curves of DMFC at different operating temperatures with the anode flow rate of 1.0 ml min<sup>-1</sup>



**Fig. 13** Performance curves of DMFC at different operating temperatures with the anode flow rate of 2.5 ml min<sup>-1</sup>



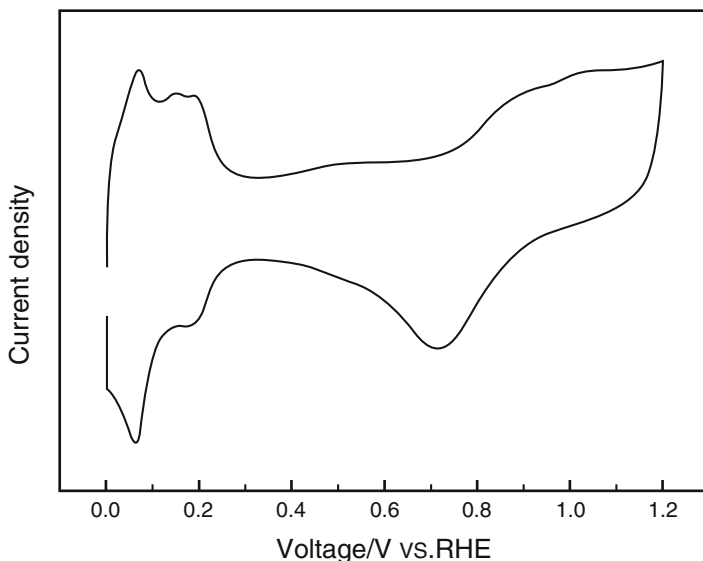
**Fig. 14** Performance curves of DMFC at different operating temperatures with the anode flow rate of  $5.0 \text{ ml min}^{-1}$

researchers. Therefore, accurate characterization and analysis of catalyst becomes a key part of DMFC study.

A three-electrode cell system is usually applied for electrochemical measurements of catalyst. The catalyst ink is loaded on the working electrode. Glassy carbon electrode (GCE) is often used as the underlying substrate and polished with alumina suspensions before using. Typically, 4 mg catalyst is dispersed in the solution of 1 mL ethanol and 20 mL Nafion (5 wt.%, DuPont) by ultrasonic treatment. A quantity of 5 mL of the dispersion is dropped onto the top of GCE, which is subsequently dried under room temperature for 2 h. The choice of the reference electrode, such as calomel electrode, Ag/AgCl electrode, and Hg/Hg<sub>2</sub>SO<sub>4</sub> electrode, depends on the solution of system in case of pollution. The counter electrode is usually served by a 1 cm\*1 cm Pt foil for Pt-based catalyst.

Cyclic voltammetry (CV) is used to evaluate the electrochemical surface area (EASA) and mass activity of catalyst. Figure 15 shows a typical CV curve in the solution of 0.5 M H<sub>2</sub>SO<sub>4</sub> of Pt-based catalyst, which is the dominant catalyst for DMFC. The curve has three regions. The left one is hydrogen area. The positive scanning curve (0–0.4 V) represents the desorption of hydrogen and the negative one (0.4–0 V) is for the absorption of hydrogen. Theoretically, a single layer of hydrogen absorption will form on the electrode. Thus, the EASA can be calculated:

$$EASA [cm^2 g^{-1} \text{ of Pt}] = \frac{Q_H [\mu C cm^{-2}]}{0.21 [\mu C cm^{-2}]^* \text{ electrode loading } [g cm^{-2}]}$$



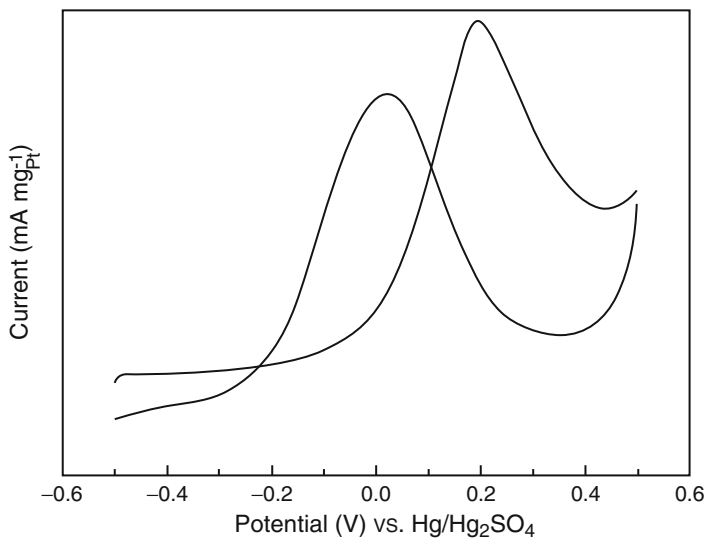
**Fig. 15** Typical CV curve of Pt in the solution of  $\text{H}_2\text{SO}_4$

where  $Q_{\text{H}}$  is the charge due to the hydrogen adsorption/desorption in the hydrogen region of the CVs and 0.21 is the electrical charge associated with monolayer adsorption of hydrogen on Pt. The middle region is double-layer region where no Faraday current occurs. The right one is oxygen region corresponding to the adsorption and desorption of oxygen.

Figure 16 reveals typical CV curve of Pt-based catalyst in the solution of 0.5 M methanol + 0.5-M  $\text{H}_2\text{SO}_4$ . There are two oxidation peaks. The one in the positive scanning direction is for methanol oxidation, and the one in the negative direction is for the oxidation of intermediate products. The mass activity can be calculated according to the peak current density. Also, the ratio of two peaks represents the ability to resist CO poisoning of catalyst.

Chronoamperometry is applied for analyzing the stability of catalyst. In the mixed solution of methanol and  $\text{H}_2\text{SO}_4$ , the potential is set to a fixed value, and a current-time relationship curve is obtained.

For physical properties of catalyst, there are a number of characterization methods to be employed. Different characterization methods should be chosen specifically for different catalysts. For instance, transmission electron microscope (TEM) and scanning electron microscope (SEM) provide direct observation of morphology of catalyst. These two methods usually come together with energy dispersive spectrometer (EDS), which is used to analyze the elemental composition, content, and distribution of samples. X-ray diffraction (XRD) spectrum consists of crystallographic information, while X-ray photoelectron spectroscopy (XPS) covers surface composition, content, chemical valence, and other information.



**Fig. 16** Typical CV curve of Pt in the solution of methanol +  $\text{H}_2\text{SO}_4$

---

## Conclusion

In this chapter, we gave a brief review of the principle, design, fabrication, encapsulation, and test for MEMS direct methanol fuel cell. Compared with traditional batteries, MEMS DMFC has significant advantages of high power density, rapid response mechanism, and long-time operation, which makes it have broad application prospects. At present, MEMS DMFC has been initially used in aerospace, military, communications, and other fields. It is foreseeable that in the near future, MEMS DMFC will play a more important role in a wider range of areas, such as individual-soldier system, micro-grid, drones, etc.

With the wide application MEMS technology, power supply of portable electronic products and systems is becoming more and more demanding in terms of integration and miniaturization. However, there are several main problems impeding commercialization of DMFC, such as poor activity of catalyst, high cost, low-energy efficiency, etc. These difficulties need more attention from researchers. We hope that this chapter brings more interest into MEMS DMFC. And we believe one day MEMS DMFC will play an indispensable role in life.

---

## References

- Coppo M, Siegel NP, von Spakovsky M (2006) On the influence of temperature on PEM fuel cell operation. *J Power Sources* 159:560
- Dillon R, Srinivasan S, Aricò AS et al (2004) International activities in DMFC R&D: status of technologies and potential applications. *J Power Sources* 127:112–126

- Geim AK, Novoselov KS (2007) The rise of graphene. *Nat Mater* 6:183–191
- Gojkovića SL, Vidakovića TR, Đurović DR (2003) Kinetic study of methanol oxidation on carbon-supported PtRu electrocatalyst. *Electrochim Acta* 48(24):3607–3641
- Hsieh SS, Huang CF, Feng CL (2008) A novel design and micro-fabrication for copper (Cu) electroforming bipolar plates. *Micron* 39:263–268
- Kamarudin SK, Achmad F, Daud WRW (2009) Overview on the application of direct methanol fuel cell (DMFC) for portable electronic devices. *Int J Hydrog Energy* 34:6902–6916
- Kundu A, Jang JH, Gil JH et al (2007) Micro-fuel cells-current development and applications. *J Power Sources* 170:67–78
- Lee SJ, Hsu CH, Huang CH (2005) Analyses of the fuel cell stack assembly pressure. *J Power Sources* 145:353–361
- Liu P, Nørskov JK (2001) Kinetics of the anode processes in PEM fuel cells – the promoting effect of Ru in PtRu anodes. *Fuel Cells* 1:192–201
- Liu F, Lee JY, Zhou WJ (2006) Segmented Pt/Ru, Pt/Ni, and Pt/RuNi nanorods as model bifunctional catalysts for methanol oxidation. *Small* 2:121
- Morse JD (2007) Micro-fuel cell power sources. *Int J Energy Res* 31:576–602
- Neyerlin KC, Gasteiger HA, Mittelsteadt CK et al (2005) Effect of relative humidity on oxygen reduction kinetics in a PEMFC. *Electrochem Soc* 152:A1073–A1108
- Nguyen NT, Chan SH (2006) Micromachined polymer electrolyte membrane and direct methanol fuel cells – a review. *J Micromech Microeng* 16:R1–R12
- Shim J, Lee J, Ye Y et al (2012) One-pot synthesis of intermetallic electrocatalysts in ordered, large-pore mesoporous carbon/silica toward formic acid oxidation. *ACS Nano* 6(8):6870–6881
- Shixian W, He Q, Zhou C et al (2012) Synthesis of Fe<sub>3</sub>O<sub>4</sub> and Pt nanoparticles on reduced graphene. *Nanoscale* 4:2478
- Shuping Y, Liu Q, Yang W et al (2013) Graphene–CeO<sub>2</sub> hybrid support for Pt nanoparticles as potential electrocatalyst for direct methanol fuel cells. *Electrochim Acta* 94:245–251
- Silva VB, Rouboa A et al (2012) In situ activation procedures applied to a DMFC: analysis and optimization study (report). *Fuel* 677(7):93
- Tafaoli-Masoule M, Bahrami A, Elsayed EM (2014) Optimum design parameters and operating condition for maximum power of a direct methanol fuel cell using analytical model and genetic algorithm. *Energy* 70:643–652
- Taymaz I, Benli M (2010) Numerical study of assembly pressure effect on performance of proton exchange membrane fuel cell. *Energy* 35:2134–2140
- Xing XQ, Lum KW, Poh HJ et al (2010) Optimization of assembly clamping pressure on performance of proton-exchange membrane fuel cells. *J Power Sources* 195:62–68
- Yuan Z, Zhang Y, Leng J et al (2012) Performance of air-breathing direct methanol fuel cell with Au-coated aluminum current collectors. *Int J Hydrog Energy* 37:2571–2578
- Zhou Y, Lin G, Shih AJ et al (2007) A micro-scale model for predicting contact resistance between bipolar plate and gas diffusion layer in PEM fuel cells. *J Power Sources* 163:777–783
- Zhou Y, Lin G, Shih AJ et al (2009) Assembly pressure and membrane swelling in PEM fuel cells. *J Power Sources* 192:544–551



# MEMS Piezoelectric Vibration Energy Harvesters

Licheng Deng, Zhiyu Wen, and Xingqiang Zhao

## Contents

Introduction .....	1298
Working Principle .....	1299
Piezoelectric Effect .....	1299
Piezoelectric Equation (PE) .....	1300
Piezoelectric Materials .....	1302
Modeling of PVEH .....	1303
Optimization of PVEH .....	1309
Effect of Damping Ratio $\zeta_m$ .....	1310
Effect of Electromechanical Coefficient $K_{ef}$ .....	1313
Fabrication of PVEH .....	1317
The Fabrication of the AlN Thin Films Based PVEH .....	1318
The Fabrication of the PZT Thin Films Based PVEH .....	1320
The Fabrication of the PZT Thick Films PVEH .....	1327
Summary and Challenges .....	1329
References .....	1330

---

L. Deng (✉)

College of Electronic and Optical Engineering and College of Microelectronics, Nanjing University of Posts and Telecommunications, Nanjing, China

Defense Key Disciplines Lab of Novel Micro-Nano Devices and System Technology, Chongqing University, Chongqing, China

e-mail: [dlc@njupt.edu.cn](mailto:dlc@njupt.edu.cn)

Z. Wen

Defense Key Disciplines Lab of Novel Micro-Nano Devices and System Technology, Chongqing University, Chongqing, China

e-mail: [wzy@cqu.edu.cn](mailto:wzy@cqu.edu.cn)

X. Zhao

School of Information and Control, Nanjing University of Information Science and Technology, Nanjing, China

e-mail: [zxq8562@163.com](mailto:zxq8562@163.com)



---

**Abstract**

Rapid advances in microelectronic devices (such as embedded sensors, medical implants, wireless communication nodes, and so on) have stimulated the development in ambient energy harvesting over the last decade. In this chapter, the fundamental theory and the typical microfabrication of MEMS piezoelectric vibration energy harvester were introduced. Firstly, the application background of the vibration energy harvester was analyzed. Then the working principle of the piezoelectric vibration energy harvester was introduced. After that, the theoretical model was deduced through Hamilton's principle and Euler-Bernoulli beam theory. The optimization of the piezoelectric vibration energy harvester was performed, and the optimization results show that the electromechanical conversion efficiency is not more than 50% and the PVEH with low damping ratio does not always have high power output, which challenges previous literature suggestion that lower damping ratio tends to higher power output. Finally, the typical microfabrication of MEMS based piezoelectric energy harvesters was described and the challenges of the MEMS PVEH were illustrated.

---

**Keywords**

MEMS · Piezoelectric effect · Vibration energy harvesting

---

**Introduction**

Wireless technology and ultra-low-power electronics such as micro-electromechanical system (MEMS) devices have made rapid progress over the last decades (Dubois-Ferrière et al. 2006). At present, such devices are usually supplied by batteries. However, in most cases, batteries cannot last the entire lifetime of the device, and periodic replacement or recharging is expensive and tedious. As a promising alternative method for batteries, harvesting energy from ambient energy in the surrounding medium has attracted increasing attention in recent years (Saadon and Sidek 2011). The potential energy sources from ambient energy include solar (Hande et al. 2007), thermal (Ujihara et al. 2007), and mechanical vibration. The solar energy is widely distributed, but it is affected by the weather. And in some applications, the sun cannot shine, such as closed engine room. The application environment with large temperature gradient is relatively less. However, vibration energy commonly exists in our living environment, so vibration energy harvesting has been widely and intensively studied. Three kinds of electromechanical transducers including electrostatic (Suzuki 2011), electromagnetic (Zorlu et al. 2011), and piezoelectric (Saadon and Sidek 2011) have been widely used for vibration energy harvesting, and piezoelectric transducers have aroused more interest because of its simpler configuration, higher conversion efficiency, and more compatibility with microfabrication techniques (Anton and Sodano 2007). Recently, the triboelectric energy harvester proposed by Zhonglin Wang group (Fan et al. 2012) has attracted extensive attention and rapid development.

This chapter mainly provides an overview of MEMS based piezoelectric vibration energy harvester (PVEH). Followed by the introduction of the working principle of PVEH, the theoretical model of PVEH, and the optimization of PVEH. And then typical MEMS-based piezoelectric energy harvesters will be described and reviewed. Finally, the summary and challenges were carried out.

---

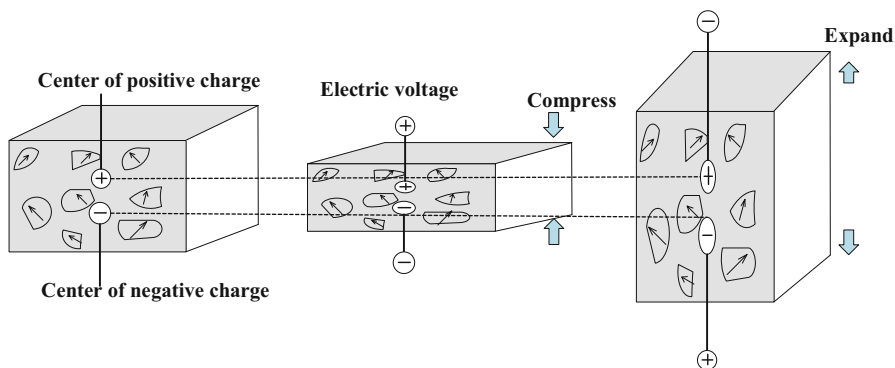
## Working Principle

### Piezoelectric Effect

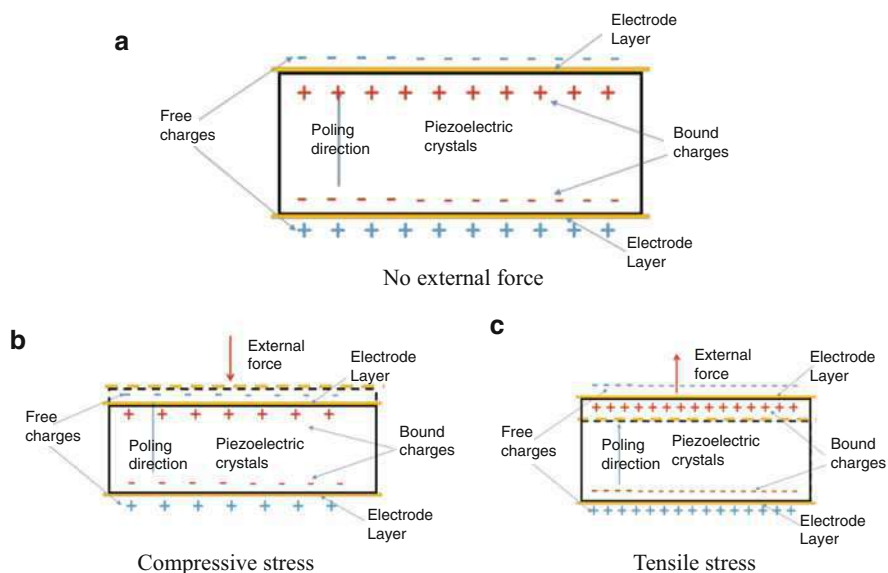
Piezoelectric effect was first discovered in 1880 by the brothers Pierre Curie and Jacques Curie. The piezoelectric effect is understood as the linear electromechanical interaction between the mechanical and the electrical state in crystalline materials with no inversion symmetry. As shown in Fig. 1, when subjected to a mechanical force, the crystals became electrically polarized. The electric voltage of opposite polarity is generated by tension and compression and is in proportion to the applied force. Subsequently, the converse of this relationship was confirmed: if the crystal is exposed to an electric field, it experiences mechanical strain according to the polarity of the field and is in proportion to the strength of the electric field. These phenomena were labeled as the direct piezoelectric effect and the inverse piezoelectric effect, respectively.

Based on direct piezoelectric effect, the piezoelectric vibration energy harvester (PVEH) achieves the conversion of vibrational energy into electrical energy, the working principle as shown in Fig. 2. Because of the inherent electric polarization of the piezoelectric crystal, a layer of free charge was attached at the surface electrode of the piezoelectric crystal, as shown in Fig. 2a. When the external force which is parallel to the polarization direction acts on the piezoelectric crystal, the external force causes the distance between the bound positive charge and bound negative charge to decrease. As a result, the electric dipole moment decreases, which causes the electric polarization strength to decrease and the surface electrode to release part of the free charge. The surface electrode presents discharge phenomenon, as shown in Fig. 2b. Similarly, when the external force causes the distance between the bound positive charge and bound negative charge increase, the electric dipole moment increases, the electric polarization strength enhances, and the surface electrode absorbs free charge from the outside. At this time, the surface electrode presents charging phenomenon, as shown in Fig. 2c. In summary, when the continuous alternating force acts on the piezoelectric crystal, the surface electrode continuously charges and discharges, as a result the mechanical energy is converted to electrical energy.

According to the relationship between applied stress direction and the piezoelectric crystal polarization direction, the operation modes of PVEH are mainly  $d_{31}$  mode and  $d_{33}$  mode, as shown in Fig. 3. When the applied stress direction and the polarization direction is orthogonal to each other, the PVEH operates in  $d_{31}$  mode. When the applied stress direction and the polarization direction are parallel to each



**Fig. 1** The schematic of the direct piezoelectric effect and reverse effect



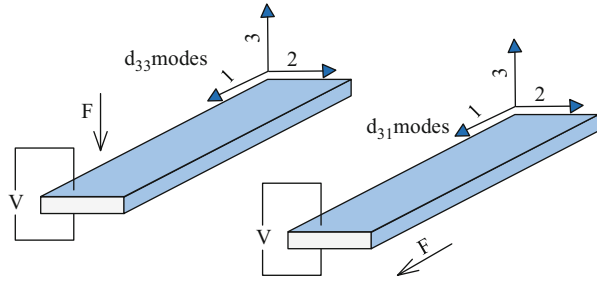
**Fig. 2** The schematic of the energy conversion with piezoelectric effect (a) No external force, (b) Compressive stress, (c) Tensile stress

other, the PVEH operates in  $d_{33}$  mode. The relationship between the mechanical and electrical variables can be described by piezoelectric equation.

### Piezoelectric Equation (PE)

Piezoelectric equation (piezoelectric constitutive equation) is the state equation of the relationship between elastic variable (i.e., stress and strain) and electrical variable

**Fig. 3** The operate mode of piezoelectric material



(i.e., electric field, electric displacement) of piezoelectric crystal. According to the boundary conditions (mechanical free, mechanical clamping, electrical short circuit, and electrical open circuit), there are four types of piezoelectric equation (Zhang and Wang 2002), as shown in Table 1:

Where subscript  $i, j = 1\sim 6$ ,  $m, n = 1\sim 3$ ;  $\epsilon(F/m)$  is dielectric constant;  $\beta(m/F)$  is dielectric isolation rate;  $s(m^2/N)$  is compliance coefficient;  $c(N/m^2)$  is stiffness constant;  $d(m/V)$ ,  $e(C/m^2)$ ,  $g(m^2/V)$  and  $h(V/m)$  are piezoelectric strain constant, piezoelectric stress constant, piezoelectric voltage constant, and piezoelectric stiffness constant, respectively. Superscript  $()^T, ()^S, ()^D$  and  $()^E$  denotes the test value at the boundary conditions with mechanical free, mechanical clamping, electrical open circuit, and electrical short circuit, respectively.

For the PVEH, the first and second types of piezoelectric equations are usually used for system modeling and simulation. For the PVEH with piezoelectric cantilever structure which operates in  $d_{31}$  mode, the thickness of the piezoelectric layer is generally much smaller than the length and width of the piezoelectric layer, i.e., the piezoelectric layer can be equivalent to thin plate. In this case, the normal stress and shear stress in the thickness direction of the piezoelectric layer can be neglected, and only the 3-direction have electric field (Erturk and Inman 2011), then:

$$T_3 = T_4 = T_5 = 0 \quad (1)$$

$$E_1 = E_2 = 0 \quad (2)$$

The matrix form of the first kind of piezoelectric equations is:

$$\begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \\ D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} s_{11}^E & s_{12}^E & s_{13}^E & 0 & 0 & 0 & 0 & 0 & d_{31} \\ s_{12}^E & s_{11}^E & s_{13}^E & 0 & 0 & 0 & 0 & 0 & d_{31} \\ s_{13}^E & s_{12}^E & s_{33}^E & 0 & 0 & 0 & 0 & 0 & d_{33} \\ 0 & 0 & 0 & s_{44}^E & 0 & 0 & 0 & d_{15} & 0 \\ 0 & 0 & 0 & 0 & s_{44}^E & 0 & d_{15} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{66}^E & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & d_{15} & 0 & \epsilon_{11}^T & 0 & 0 \\ 0 & 0 & 0 & d_{15} & 0 & 0 & 0 & \epsilon_{22}^T & 0 \\ d_{31} & d_{31} & d_{33} & 0 & 0 & 0 & 0 & 0 & \epsilon_{33}^T \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \\ E_1 \\ E_2 \\ E_3 \end{bmatrix} \quad (3)$$

**Table 1** Four types of piezoelectric equation

	Boundary conditions	PE
The first kind of PE	Mechanical free Electrical short circuit	$S_j = s_{ij}^E T_j + d_{ni} E_n$ $D_m = d_{mj} T_j + \varepsilon_{mn}^T E_n$
The second kind of PE	Mechanical clamping Electrical short circuit	$T_i = c_{ij}^E S_j - e_{nj} E_n$ $D_m = e_{mj} S_j + \varepsilon_{mn}^S E_n$
The third kind of PE	Mechanical free Electrical open circuit	$S_i = s_{ij}^D T_j + g_{mi} D_m$ $E_m = -g_{mj} T_j + \beta_{mn}^T D_n$
The fourth kind of PE	Mechanical clamping Electrical open circuit	$T_i = c_{ij}^D S_j - h_{mi} D_m$ $E_m = -h_{mi} S_i + \beta_{mn}^S D_n$

Substituting Eqs. 1 and 2 into Eq. 3:

$$\begin{bmatrix} S_1 \\ S_2 \\ S_6 \\ D_3 \end{bmatrix} = \begin{bmatrix} s_{11}^E & s_{12}^E & 0 & d_{31} \\ s_{12}^E & s_{11}^E & 0 & d_{31} \\ 0 & 0 & s_{66}^E & 0 \\ d_{31} & d_{31} & 0 & \varepsilon_{33}^T \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_6 \\ E_3 \end{bmatrix} \quad (4)$$

And then:

$$\begin{bmatrix} T_1 \\ T_2 \\ T_6 \\ D_3 \end{bmatrix} = \begin{bmatrix} c_{11,f}^E & c_{12,f}^E & 0 & -e_{31,f} \\ c_{12,f}^E & c_{11,f}^E & 0 & -e_{31,f} \\ 0 & 0 & c_{66,f}^E & 0 \\ e_{31,f} & e_{31,f} & 0 & \varepsilon_{33,f}^S \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ S_6 \\ E_3 \end{bmatrix} \quad (5)$$

The equivalent variables as shown in Table 2 where  $\nu = -s_{12}^E/s_{11}^E$  is the Poisson's ratio when the piezoelectric layer compresses in y-direction because of the piezoelectric layer stretching in the x-direction.

## Piezoelectric Materials

The selection and preparation of high-performance piezoelectric materials is the key to obtain high-performance PVEH. For the MEMS PVEH, when selecting the piezoelectric materials, we are not only to consider the piezoelectric properties of the piezoelectric material itself but also should analyze the microfabrication compatibility of the piezoelectric materials.

Among the piezoelectric material parameters, the performance of the PVEH is mainly affected by the piezoelectric stress constant, free dielectric constant, voltage coefficient, and electromechanical coupling coefficient. When operating in  $d_{31}$  mode, the voltage coefficient and electromechanical coupling coefficient can be expressed as  $e_{31,f}/\varepsilon_0 \varepsilon_{33}^S$  and  $e_{31,f}^2/\varepsilon_0 \varepsilon_{33}^S$ , respectively. Wherein the voltage coefficient

**Table 2** The expression of equivalent variable

Equivalent variable	$C_{11,f}^E$	$C_{12,f}^E$	$C_{66,f}^E$	$e_{31,f}$	$\epsilon_{33,f}^s$
Expression	$\frac{1}{s_{11}^E(1-\nu^2)}$	$\frac{1}{s_{12}^E(1/\nu^2-1)}$	$\frac{1}{s_{66}^E}$	$\frac{d_{31}}{s_{11}^E(1-\nu)}$	$\epsilon_{33}^T - \frac{2d_{31}^2}{s_{11}^E(1-\nu)}$

**Table 3** The performance parameters of piezoelectric thin film materials

	ZnO	AlN	PZT(53/47) (Ledermann et al. 2003)
$e_{31,f}$ (C/m <sup>2</sup> )	-1.0	-1.05	-14
$\epsilon_{33,f}^s$	10.9	10.5	940
$e_{31,f}/\epsilon_0\epsilon_{33}^s$ (GV/m)	-10.3	-11.3	-1.7
$e_{31,f}^2/\epsilon_0\epsilon_{33}^s$ (GVC/m <sup>3</sup> )	10.3	11.3	23.6

reflects the voltage output performance of the PVEH, the electromechanical coupling coefficient reflects the power output performance of the PVEH.

According to the above two factors, the piezoelectric film materials which are commonly used in MEMS PVEH are mainly PZT, AlN, and ZnO. Table 3 lists the performance parameters of the three piezoelectric thin film materials which operate in  $d_{31}$  mode.

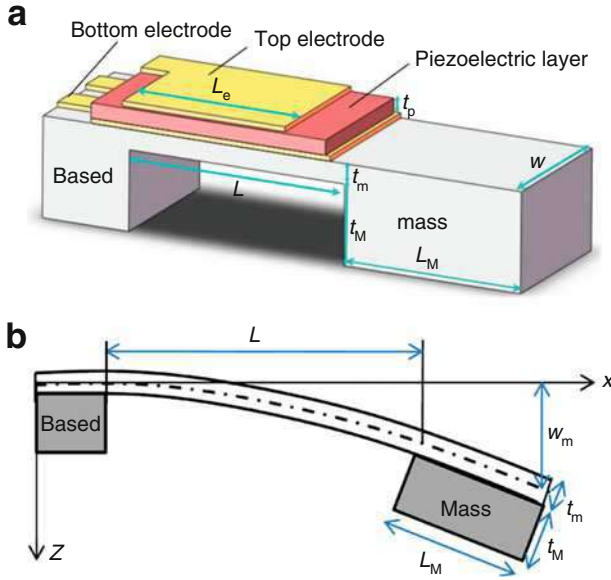
AlN and ZnO are nonferroelectric materials and can be deposited onto silicon substrate by sputtering. The preparation process is compatible with low temperature MEMS processes. PZT is a ferroelectric material that requires polarization to have piezoelectric properties. At present, PZT thin film materials are usually prepared by sol-gel method and sputtering process, which need high-temperature annealing at above 600 °C. However, the piezoelectric stress coefficient and electromechanical coupling coefficient of the PZT is higher than both AlN and ZnO.

## Modeling of PVEH

The typical cantilevered PVEH is shown in Fig. 4. Piezoelectric layer is sandwiched between two electrodes which are constructed on a silicon cantilever beam. At the free end of the cantilever, a proof mass is attached whose center of gravity does not coincide with the point of attachment. The PVEH operates in  $d_{31}$  mode. Electromechanical coupling model can be obtained by Hamilton's principle. The variational indicator (Preumont 2006) is as follows:

$$V.I. = \int_{t_1}^{t_2} (\delta(T_m^* + T_M^* + W_e^* - V) + \delta W_{nc}) dt = 0 \quad (6)$$

where superscript “\*” denotes coenergy.  $T_m$ ,  $T_M$ ,  $W_e$ ,  $V$ , and  $W_{nc}$  denote kinetic energy of the cantilever, kinetic energy of the proof mass, electrical energy, potential energy, and external work, respectively. The constitutive equations of the small signal assumed for a general piezoelectric material are (Meitzler et al. 1988):



**Fig. 4** The schematic of the PVEH (a) and the cross-section of the PVEH with base excitation (b)

$$\begin{bmatrix} T \\ D \end{bmatrix} = \begin{pmatrix} c^E & -e^t \\ e & \epsilon^S \end{pmatrix} \begin{bmatrix} S \\ E \end{bmatrix}. \tag{7}$$

Where  $S$ ,  $T$ ,  $E$ , and  $D$  matrices are defined as applied strain, developed stress, applied electric field, and developed electric displacement, respectively.  $c$ ,  $\epsilon$ , and  $e$  matrices denote stiffness, permittivity, and piezoelectric constant of the piezoelectric material. And these physical parameters are obtained differently depending on whether the structure is a beam or a plate. Superscript  $t$  indicates the transpose of the matrix. Superscripts  $E$  and  $S$  indicate parameters at constant electric field and constant strain. The coenergy term of the cantilever is:

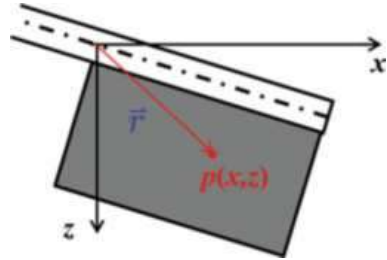
$$T_m^* = \frac{1}{2} \int \rho_s \dot{w}^t \dot{w} dV_s + \frac{1}{2} \int \rho_p \dot{w}^t \dot{w} dV_p. \tag{8}$$

where  $w(x, t) = w_B(t) + w_{rel}(x, t)$  is the absolute displacement,  $w_{rel}(x, t)$  is the relative displacement,  $w_B(t)$  is the base displacement, and  $\rho$  is the density. Subscript  $s$  and  $p$  denote silicon beam and piezoelectric layer, respectively. The velocity of proof mass is the sum of the  $Z$ -direction and  $X$ -direction velocity, as shown in Fig. 5.

$$\vec{v}_M = \dot{\vec{w}}(L, t) + \vec{r} \dot{\vec{w}}'(L, t) = (\dot{w}(L, t) + x\dot{w}'w'(L, t))\vec{e}_z + z\dot{w}'w'(L, t)\vec{e}_x \tag{9}$$

where the  $\vec{r}$  is distance between coordinate origin and any point  $p(x, z)$  on the proof mass. So the coenergy terms of the proof mass is:

**Fig. 5** The coordinate of the proof mass



$$T_M^* = \frac{1}{2} \int \rho \left( \dot{w}(L, t) + x\dot{w}'(L, t) \right)^2 dV_M$$

$$+ \frac{1}{2} \int \rho \left( z\dot{w}'(L, t) \right)^2 dV_M = \frac{1}{2} M_0 \dot{w}^2(L, t) + S_0 \dot{w}'(L, t) \dot{w}(L, t) + \frac{1}{2} J_0 \dot{w}' w'^2_{rel}(L, t). \quad (10)$$

where  $M_0 = m_0 L_M + mL_M$ ,  $S_0 = \frac{M_0 L_M}{2}$ , and  $J_0 = \frac{m_0 L_M}{3} (L_M^2 + l_M^2) + \frac{m L_M}{3} (L_M^2 + l_m^2)$ .  $m$  and  $m_0$  are the linear density of beam and proof mass, respectively. The electrical coenergy and potential energy are:

$$W_e^* - V = \frac{1}{2} \int (E^t \epsilon E + 2S^t e E - S^t c_p S) dV_p - \frac{1}{2} \int S^t c_s S dV_s. \quad (11)$$

Based on Raleigh-Ritz approach, the relative displacement and scalar electrical potential can be expressed as:

$$w_{rel}(x, t) = \sum_{i=1}^{nr} \psi_{ri}(x) \eta_i(t) = \psi_r(x) \eta(t). \quad (12)$$

$$\varphi(x, t) = \sum_{i=1}^{nr} \phi_{vi}(x) v_i(t) = \phi_v(x) v(t). \quad (13)$$

where  $\psi_r(x)$  and  $\psi_v(x)$  represent mechanical mode shape vector and electric mode shape vector, respectively.  $\eta(t)$  and  $v(t)$  are the modal mechanical response and modal electric response, respectively. The nonconservative virtual work contributed from external forces and load resistance for discrete systems is given by

$$\delta W_{nc} = - \sum_i c_m \dot{\eta} \delta w - \sum_i q_i \delta \varphi \quad (14)$$

where  $c_m$  is the coefficient of damping force, and  $q$  is the charge. The strain of cantilever can be obtained by Euler-Bernoulli beam theory



**Table 4** The equivalent terms

Equivalent terms	Expression
$M$	$\int m\psi_r^t \psi_r dx + M_0 \psi_r^t(L) \psi(L) + 2\psi_r^t(L) S_0 \psi_r^t(L) + J_0 \psi_r^t(L) \psi_r^t(L)$
$C$	$\sum_{i=1}^n \psi_r^t c_a$
$\Theta$	$\int (-z_t \psi_r^t) e^t (-\nabla \phi_v) dV_p$
$K$	$\int (-z_t \psi_r^t) c_s (-z_t \psi_r^t) dV_s + \int (-z_t \psi_r^t) c_p (-z_t \psi_r^t) dV_p$
$B_f$	$\int m \psi_r^t dx + M_0 \psi_r^t(L) + \psi_r^t(L) S_0$
$C_p$	$\int (-\nabla \phi_v)^t e^t (-\nabla \phi_v) dV_p$
$q$	$\sum_{j=1}^n \phi_{vj} q_j$

$$S = -z_t \frac{\partial^2 w(x, t)}{\partial x^2} = -z_t \psi_r^t(x) \eta(t). \tag{15}$$

Electric field can be defined as gradient of electric potential:

$$E = -\nabla \varphi(x, t) = -\nabla \phi_v(x) v(t). \tag{16}$$

Substituting Eqs. 7, 8, 9, 10, 11, 12, 13, 14, 15, and 16 to Eq. 6 and noted that  $\delta w_B = \delta \dot{w}_B = 0$

$$M \ddot{\eta} + C \dot{\eta} + K \eta - \Theta v = -B_f \ddot{w}_B. \tag{17}$$

$$\Theta \eta + C_p v + q = 0. \tag{18}$$

The equivalent terms, including the mass ( $M$ ), damping coefficient ( $C$ ), coupling ( $\Theta$ ), stiffness ( $K$ ), inertial force ( $B_f$ ), and capacitive ( $C_p$ ), are listed in Table 4.  $q$  denotes the total charge generated by piezoelectric film.

The mode shape  $\psi_{rN}$  can be determined by Euler-Bernoulli beam theory

$$EI \psi_{rN}^{(4)} - m \omega_N^2 \psi_{rN} = 0. \tag{19}$$

$$\psi_{rN} = a \cos \frac{\lambda_N}{L} x + b \cosh \frac{\lambda_N}{L} x + c \sin \frac{\lambda_N}{L} x + d \sinh \frac{\lambda_N}{L} x. \tag{20}$$

where  $\omega_N = \lambda_N^2 \sqrt{EI/(mL^4)}$  is the natural frequency of the PVEH.  $EI$  is equivalent stiffness. The coefficients  $a, b, c,$  and  $d$  can be determined by boundary conditions and orthogonality conditions as shown in Table 5.

In most cases, the first mode shape is the dominant mode shape for PVEH. So the first order vibration is considered and the Eqs. 17 and 18 simplified as

**Table 5** The boundary conditions and orthogonality conditions

Boundary conditions	Orthogonality conditions
$\psi_{rN}(0) = 0, \psi'_{rN}(0) = 0$	$M_{ij} = \delta_{ij}$
$EI\psi''_{rN}(L) = \omega_N^2 J_0 \psi'_{rN}(L) + \omega_N^2 S_0 \psi_{rN}(L)$	$K_{ij} = \delta_{ij} \omega_i^2$
$EI\psi'''_{rN}(L) = -\omega_N^2 M_0 \psi_{rN}(L) - \omega_N^2 S_0 \psi'_{rN}(L)$	

$$\ddot{\eta} + 2\zeta\omega_1\dot{\eta} + \omega_1^2\eta - \theta v = -B_f\ddot{w}_B. \quad (21)$$

$$\theta\dot{\eta} + C_p\dot{v} + \frac{1}{R_l}v = 0 \quad (22)$$

where  $\zeta_m = C/2\omega_1$  is the damping ratio. The potential for  $d_{31}$  mode varies from 0 at the top electrode to +1 at the bottom electrode (positive polar) expressed as:

$$\phi_v = \frac{t_{pu} - z_t}{t_p}, t_{su} \leq z_t \leq t_{pu}. \quad (23)$$

where  $t_{su}, t_{pu}$  is the distance from the bottom and top of the piezoelectric layer to the neutral plane respectively,  $t_p$  is thickness of the piezoelectric layer. So,

$$\theta = e_{31} \int (-z_t \psi'' )^t (-\nabla \phi_v) dV_p = -\frac{e_{31} W}{2} [t_{pu} + t_{su}] \psi'_{r1}(L_e). \quad (24)$$

$$C_p = \int (-\nabla \phi_v)^t \varepsilon^t (-\nabla \phi_v) dV_p = \frac{\varepsilon_{33} W L_e}{t_p}. \quad (25)$$

If the base displacement is  $w_B = Y_0 \cos \omega t$ , where  $\omega$  and  $Y_0$  are the frequency and amplitude of the base motion. We assume that the frequency of the mechanical response and the electric response are the same as the frequency of base motion, i.e.,  $\eta(t) = H \cos(\omega t + \varphi_1)$  and  $v(t) = V \cos(\omega t + \varphi_2)$ , respectively.  $H$  and  $V$  are the amplitude of the mechanical response and the electric response, respectively. Therefore,

$$\eta(t) = \frac{F_r \cos(\omega t + \varphi_1)}{\omega_1^2 \sqrt{(1 - \Omega^2 + 2\zeta_e r \Omega^2)^2 + (2\zeta_m \Omega + 2\zeta_e \Omega)^2}}. \quad (26)$$

$$v(t) = \frac{-\omega_0 R_l \theta F_r \cos(\omega t + \varphi_2)}{\omega_1^2 \sqrt{(1 - \Omega^2 - 2\zeta_m r \Omega^2)^2 + \Omega^2 (2\zeta_m + r(1 - \Omega^2) + rK_{ef})^2}}. \quad (27)$$

$$w_{rel} = \frac{F_r \cos(\omega t + \varphi_1) \varphi_{r1}(x)}{\omega_1^2 \sqrt{(1 - \Omega^2 + 2\zeta_e r \Omega^2)^2 + (2\zeta_m \Omega + 2\zeta_e \Omega)^2}}. \quad (28)$$

**Table 6** The variables and the corresponding expressions

Variables	Expressions
Normalized frequency $\Omega$	$\omega/\omega_1$
Normalized load resistance $r$	$\omega_1 R_l C_p$
Electromechanical coefficient $K_{ef}$	$\frac{\theta^2}{\omega_1^2 C_p}$
Equivalent electric damping ratio $\zeta_e$	$\frac{rK_{ef}}{2(1+\Omega^2 r^2)}$
Equivalent exciting force $F_r$	$-B_f \ddot{w}_B$

So the maximum displacement of the cantilever is

$$\begin{aligned}
 w_m &= w(L, t) + \partial_x w(L, t) \times L_M \\
 &= [\varphi_{r1}(x) + \partial_x \varphi_{r1}(x) \times L_M]_{x=L} \times H.
 \end{aligned}
 \tag{29}$$

The power output of the PVEH can be expressed as:

$$\begin{aligned}
 P_{Rave} &= \frac{1}{T} \int_0^T \frac{v^2}{R_l} dt \\
 &= \frac{\Omega^2 r K_{ef} F_r^2}{2\omega_1 \left( (1 - \Omega^2 - 2\zeta_m r \Omega^2)^2 + \Omega^2 (2\zeta_m + r(1 - \Omega^2) + rK_{ef})^2 \right)}.
 \end{aligned}
 \tag{30}$$

Where  $\varphi_1 = \arctan \left[ \frac{2\zeta_m \Omega + 2\zeta_e \Omega}{1 - \Omega^2 + 2\zeta_e r \Omega^2} \right]$ ,  $\varphi_2 = \arctan \left[ \frac{1 - \Omega^2 - 2\zeta_m r \Omega^2}{-\Omega(2\zeta_m + r(1 - \Omega^2) + rK_{ef})} \right]$ . The normalized variables and the corresponding expressions are listed in Table 6.

The above method is called directly solving differential equations method. The power output of the PVEH can also be obtained through the method of conservation of energy, which is called conservation of energy method. Let Eq. 21 be multiplied by  $\dot{\eta}$  and Eq. 22 be multiplied by  $v$ .

$$\frac{d}{dt} \left[ \left( \frac{1}{2} \dot{\eta}^2 \right) + \frac{1}{2} \omega_1^2 \eta^2 \right] + 2\zeta \omega_1 \dot{\eta}^2 - \theta v \dot{\eta} = -B_f \ddot{w}_B \dot{\eta}.
 \tag{31}$$

$$\theta \dot{\eta} v + \frac{1}{R_l} v^2 + \frac{d}{dt} \left( \frac{1}{2} C_p v^2 \right) = 0.
 \tag{32}$$

Integrating them over a cycle gives:

$$\frac{1}{T} \int_0^T 2\zeta \omega_1 \dot{\eta}^2 dt - \frac{1}{T} \int_0^T \theta v \dot{\eta} dt = -\frac{1}{T} \int_0^T B_f \ddot{w}_B \dot{\eta} dt.
 \tag{33}$$

$$\frac{1}{T} \int_0^T \theta \dot{\eta} v dt + \frac{1}{T} \int_0^T \frac{v^2}{R_l} dt = 0.
 \tag{34}$$

Therefore,

$$\begin{aligned} P_{Rave} &= \frac{1}{T} \int_0^T \frac{v^2}{R_l} dt = -\frac{1}{T} \int_0^T \theta \dot{\eta} v dt \\ &= 2\zeta_e \omega_0^2 \omega_1 \times H^2 \times \frac{(\sin \varphi_1 + r\Omega \cos \varphi_1)}{\cos \varphi_2} \times \frac{\sin(\varphi_1 - \varphi_2)}{2}. \end{aligned} \quad (35)$$

$$P_{\zeta_m} = \frac{1}{T} \int_0^T 2\zeta_m \omega_1 \dot{\eta}^2 dt = \zeta_m \times \omega_1 (\omega_0 H)^2 \quad (36)$$

$$P_{in} = -\frac{1}{T} \int_0^T B_f \ddot{w}_B \dot{\eta} dt = \frac{\sin \varphi_1}{2} F_r \times \omega_0 H. \quad (37)$$

$$P_{\zeta_m} + P_{Rave} = P_{in}. \quad (38)$$

where  $P_{in}$  is the input power from environment,  $P_{\zeta_m}$  is the damping loss.

## Optimization of PVEH

For a given PVEH and the PVEH operating in a given excitation acceleration, both  $\Omega$  and  $r$  determine the power output of the PVEH. The optimal  $\Omega$  and  $r$  can be found by differentiating Eq. 30 with respect to  $\Omega$  and  $r$ , setting the result equal to zero, and solving for  $\Omega$  and  $r$ , respectively. When the damping ratio is small, i.e.,  $K_{ef} \geq 4\zeta_m(1 + \zeta_m)$ , there are two optimal solutions:

$$\Omega_{opt1} = \frac{\sqrt{2 + K_{ef} - 4\zeta_m^2 - (A^2 - 16 - 16K_{ef})^{1/2}}}{\sqrt{2}}. \quad (39)$$

$$r_{opt1} = \frac{K_{ef} + 4\zeta_m^2 - (A^2 - 16 - 16K_{ef})^{1/2}}{4(\zeta_m + K_{ef}\zeta_m)}. \quad (40)$$

and

$$\Omega_{opt2} = \frac{\sqrt{2 + K_{ef} - 4\zeta_m^2 + (A^2 - 16 - 16K_{ef})^{1/2}}}{\sqrt{2}}. \quad (41)$$

$$r_{opt2} = \frac{K_{ef} + 4\zeta_m^2 + (A^2 - 16 - 16K_{ef})^{1/2}}{4(\zeta_m + K_{ef}\zeta_m)}. \quad (42)$$

When the damping ratio is large, i.e.,  $K_{ef} < 4\zeta_m(1 + \zeta_m)$ , there is only one solution:

$$\Omega_{opt} = \sqrt{\frac{2 + K_{ef} - 4\zeta_m^2 + A}{6}}. \quad (43)$$

$$r_{opt} = \frac{\sqrt{8 + 8\zeta_m^2 + 4K_{ef}(3 + K_{ef} + 2\zeta_m^2) + A(4 + 3K_{ef} - 4\zeta_m^2 - A)}}{2\sqrt{2}\sqrt{1 + 3K_{ef} + 3K_{ef}^2 + K_{ef}^3 - \zeta_m^2 - 2K_{ef}\zeta_m^2 - K_{ef}^2\zeta_m^2}}. \quad (44)$$

where  $A^2 = 16 + 16K_{ef} + K_{ef}^2 - 16\zeta_m^2 - 8K_{ef}\zeta_m^2 + 16\zeta_m^4$ . The optimal power output can be obtained by substituting Eqs. 39, 40, 41, 42, 43, and 44 to Eq. 30. The electromechanical conversion efficiency of the PVEH is defined as:

$$CE = \frac{P_{Rave}}{P_{in}}. \quad (45)$$

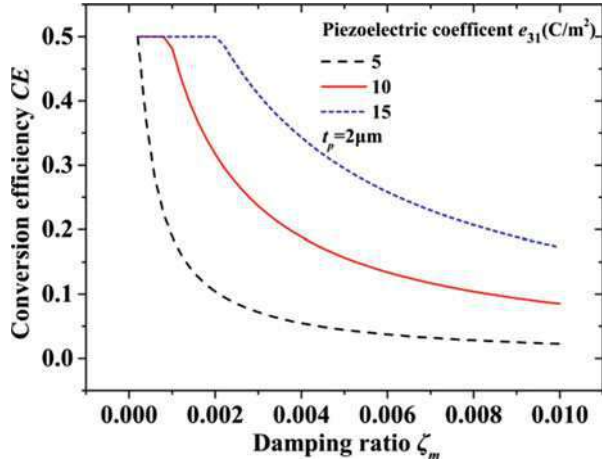
Equation 45 shows that the  $CE$  is only related to  $K_{ef}$  and  $\zeta_m$ , i.e.,  $CE = CE(K_{ef}, \zeta_m)$  and is unrelated to the strain and the displacement of the cantilever. Therefore, as long as the  $K_{ef}$  and  $\zeta_m$  are determined, the  $CE$  is determined for a given PVEH. And  $K_{ef} \geq 4\zeta_m(1 + \zeta_m)$  leads to  $CE = 50\%$ . When  $K_{ef} < 4\zeta_m(1 + \zeta_m)$ ,  $CE < 50\%$  and small  $\zeta_m$  or large  $K_{ef}$  leads to large  $CE$ . Figures 6 and 7 show the  $CE$  changes with damping ratio and electromechanical coefficient. Table 7 lists the variables and corresponding values for the whole numerical calculation.

In practical applications, ambient vibration source determines the excitation frequency and acceleration. At the same time, the allowable stress of the cantilever limits the maximum displacement of the cantilever. Therefore, the constraint conditions of excitation frequency, acceleration, and maximum displacement of cantilever should be taken into account when the PVEH is designed. So in the following numerical optimization, the constraint conditions of the given acceleration amplitude is 0.5 g ( $g = 9.8 \text{ m/s}^2$ ), and the given amplitude of the PVEH is 800  $\mu\text{m}$ . In this chapter, we mainly focus on the optimization of damping ratio  $\zeta_m$  and electromechanical coupling coefficient  $K_{ef}$ .

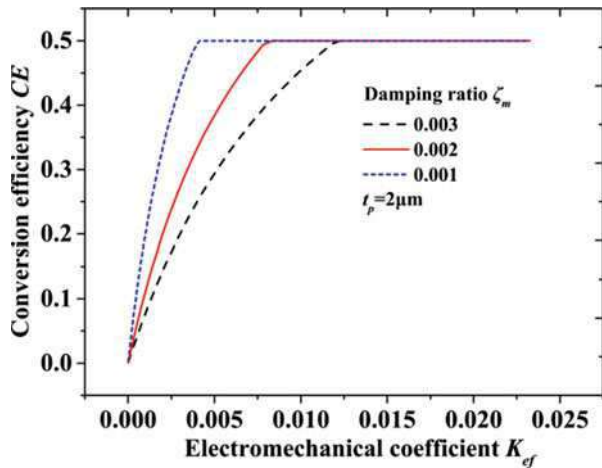
## Effect of Damping Ratio $\zeta_m$

Figure 8 shows that the power output of the PVEH decreases as damping ratio increases at given acceleration for different piezoelectric coefficient, and the power output will be infinite when the damping ratio is zero. Figure 9 shows that the maximum displacement of the cantilever is also large at small damping ratio. Figure 10 shows that the power output increases firstly with increasing damping ratio and then remains the same for given amplitude of PVEH with different piezoelectric coefficient. When the damping ratio is small, the maximum displacement of the cantilever is easy to reach its limit at ambient vibration excitation. In this case, to ensure that the PVEH operates in security, the overload protection structure should be designed. Therefore, the performance of the PVEH should be analyzed at given

**Fig. 6** The electromechanical conversion efficiency with damping ratio



**Fig. 7** The electromechanical conversion efficiency with electromechanical coefficient

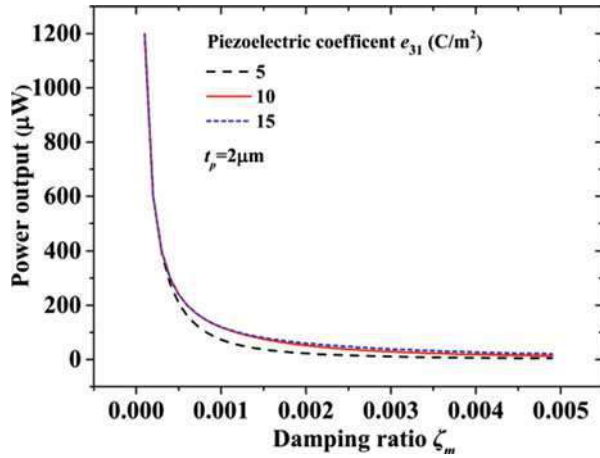


amplitude of cantilever. Figure 11 shows that the required acceleration increases as damping ratio increases for different piezoelectric coefficient. From Eq. 37, the input power also increases with increasing damping ratio, and  $CE = 50\%$  at  $\zeta_m < \frac{1}{2}(-1 + \sqrt{1 + K_{ef}})$ . So the power output increases as damping ratio increases. When  $\zeta_m \geq \frac{1}{2}(-1 + \sqrt{1 + K_{ef}})$ ,  $CE < 50\%$  and the  $CE$  decreases with increasing  $\zeta_m$ , so the power output of PVEH is basically unchanged as shown in Fig. 10. Figures 10 and 11 also show that both the power output and required acceleration increase with increasing piezoelectric coefficient. When the damping ratio is large, the maximum displacement of cantilever is difficult to reach its limited at ambient vibration excitation. The performance of the PVEH should be analyzed at given acceleration. Figure 9 shows that the maximum displacement of the cantilever decreases with increasing damping ratio for different piezoelectric coefficient, so

**Table 7** The variables and corresponding values for numerical calculation

Variables	Description	Values
$\rho_s$	Density of silicon	2.33 g/cm <sup>3</sup>
$c_{s11}$	Stiffness of silicon	170 GPa
$\nu_s$	Poisson's ratio of silicon	0.278
$\rho_p$	Density of piezoelectric layer	7.5 g/cm <sup>3</sup>
$c_{p11}$	Stiffness of piezoelectric layer	60 GPa
$\nu_p$	Poisson's ratio of piezoelectric layer	0.29
$L$	Cantilever length	3 mm
$W$	Cantilever width	14 mm
$t_s$	Thickness of silicon cantilever	50 $\mu$ m
$L_M$	Mass length	8 mm
$t_M$	Mass thickness	0.5 mm

**Fig. 8** The power output with damping ratio at given acceleration



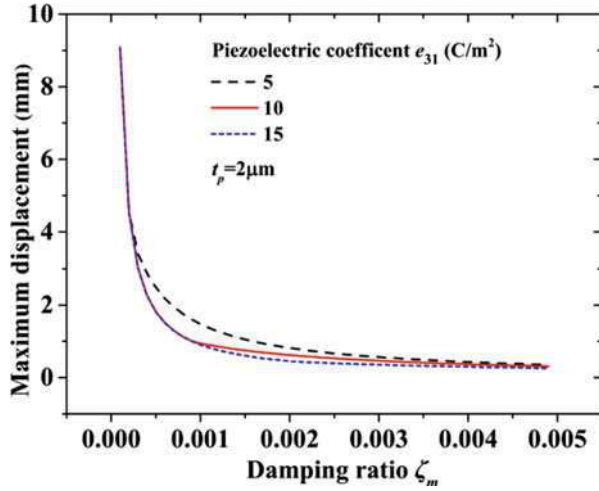
the input power also decreases with increasing damping ratio. Figure 6 shows that the *CE* remains the same firstly and then decreases with increasing damping ratio. So the power output decreases with increasing damping ratio as shown in Fig. 8. Therefore, small damping ratio is not always high power output, which challenges previous literature suggestion that lower damping ratio tends to higher power output.

Typically, the damping of the cantilever configuration consists of four parts: airflow force, squeeze force, support losses, and structural damping. These four components can be modeled as adding linearly and are defined as:

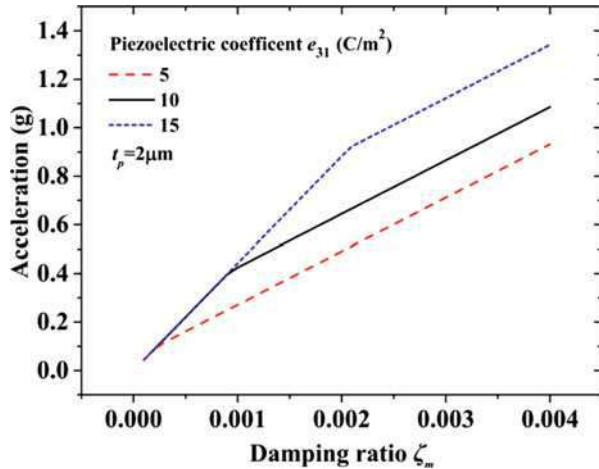
$$\zeta_m = \zeta_{air} + \zeta_{squ} + \zeta_{sup} + \zeta_{stru}. \tag{46}$$

where  $\zeta_{air} = \frac{3\pi\mu W + \frac{3}{4}\pi W^2 \sqrt{2\rho_{air}\mu\omega}}{2\rho_{beam}WtL\omega}$ ,  $\zeta_{squ} = \frac{\mu W^2}{2\rho_{beam}g_0^3 h\omega}$ ,  $\zeta_{sup} = 0.23h^3/L^3$ , and  $\zeta_{stru} = \beta/2$ .  $\mu$  is the viscosity of air,  $\beta$  is the structural damping factor of the beam

**Fig. 9** The maximum displacement with damping ratio at given acceleration



**Fig. 10** The power output with damping ratio at given amplitude of PVEH



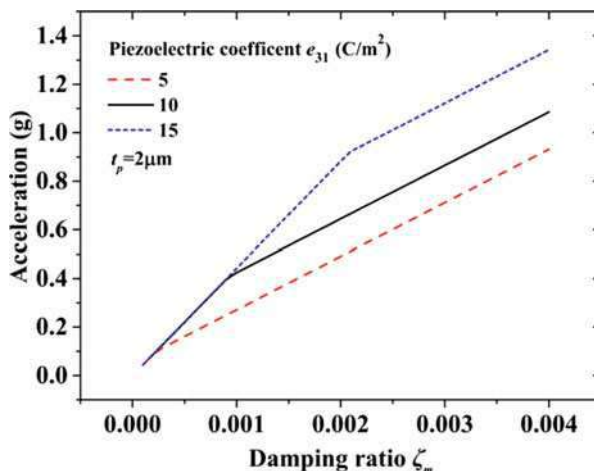
material and for silicon is  $\beta = 5 \times 10^{-6}$ .  $\rho_{beam}$ ,  $\rho_{air}$  are the density of cantilever and air.  $t$  is the thickness of cantilever.  $g_0$  is the gap between the bottom surface of the beam and a fixed floor. The operating frequency  $\omega$  coincides with the first-order natural frequency. So the damping ratio can be controlled by the structure of the PVEH and the viscosity of air.

### Effect of Electromechanical Coefficient $K_{ef}$

Electromechanical coefficient  $K_{ef}$  is mainly related with piezoelectric coefficient and the thickness of piezoelectric layer. The  $K_{ef}$  increases as the piezoelectric coefficient increases. When the thickness of the piezoelectric layer is far less than the thickness



**Fig. 11** The required acceleration with damping ratio at given amplitude of PVEH

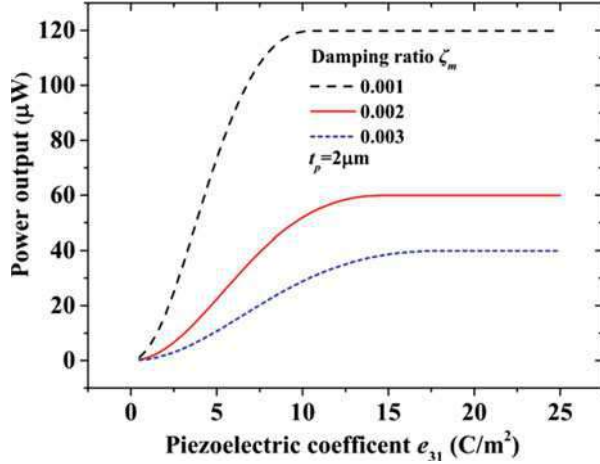


of the substrate, such as the piezoelectric thin film layer of the MEMS PVEH, the  $K_{ef}$  increases as the thickness of the piezoelectric layer increases.

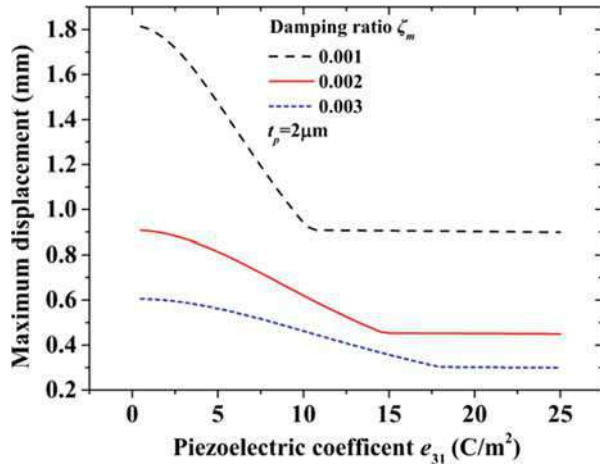
Figure 12 shows that the power output of the PVEH increases as piezoelectric coefficient increases firstly and then remains the same at given acceleration for different damping ratio. Figure 13 shows that the maximum displacement of the cantilever decreases as the piezoelectric coefficient increases firstly and then is about the same at given acceleration for different damping ratio. And the more the damping ratio is, the less the power output and maximum displacement of cantilever are. Figures 14 and 15 show that both the power output and required acceleration increase with the increasing piezoelectric coefficient firstly and then remain the same at given amplitude of PVEH for different damping ratio. And the more the damping ratio is, the more the power output and maximum displacement of cantilever are. Figure 16 shows that the equivalent electrical damping  $\zeta_e$  increases with the increasing piezoelectric coefficient firstly and then remains the same. Figure 17 shows that both the input power and damping loss decrease with increasing piezoelectric coefficient, and the damping loss decreases more quickly.

When the damping ratio is small, as mentioned in section “[Effect of Damping Ratio  \$\zeta\_m\$](#) ,” the performance of the PVEH should be analyzed at given amplitude of PVEH. The  $CE$  is less than 50% at  $K_{ef} < 4\zeta_m(1 + \zeta_m)$  and increases with  $K_{ef}$  as shown in Fig. 7. In this case, the equivalent electrical damping ratio  $\zeta_e$  increases as piezoelectric coefficient increases. So the total equivalent damping ratio (sum of the  $\zeta_m$  and  $\zeta_e$ ) increases as piezoelectric coefficient increases. Therefore, the required acceleration as well as the input power increases as piezoelectric coefficient increases. So the power output increases as piezoelectric coefficient increases. When  $K_{ef} \geq 4\zeta_m(1 + \zeta_m)$ ,  $CE = 50\%$ . The equivalent electrical damping ratio  $\zeta_e$  remains the same in this case. So the total equivalent damping ratio as well as the required acceleration and input power remain the same as the piezoelectric coefficient increases. Therefore, the power output is basically unchanged in this case. For large damping ratio, as mentioned in section “[Effect of Damping Ratio  \$\zeta\_m\$](#) ,” the

**Fig. 12** The power output with piezoelectric coefficient at given acceleration

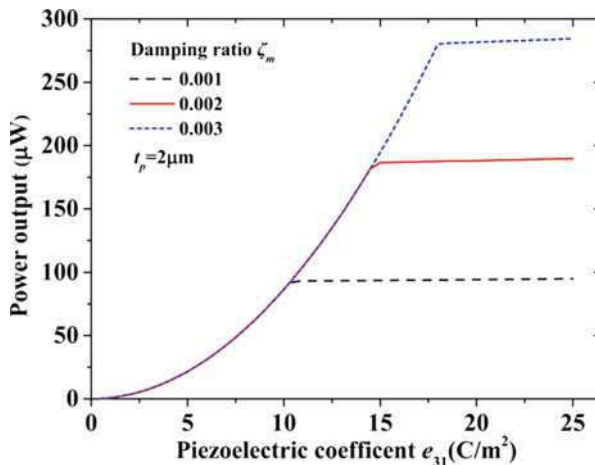


**Fig. 13** The maximum displacement with piezoelectric coefficient at given acceleration

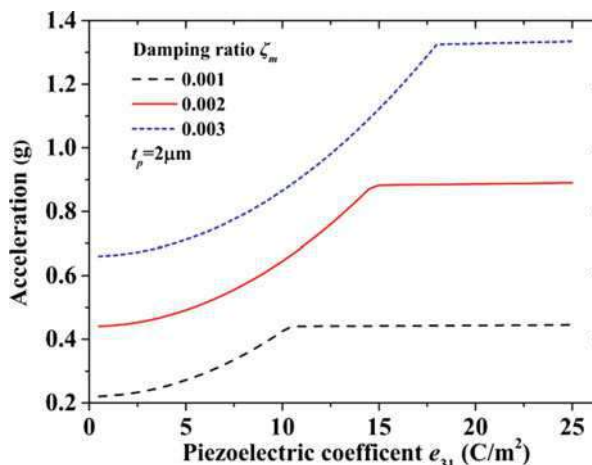


performance of the PVEH should be analyzed at given acceleration. When  $K_{ef} < 4\zeta_m(1 + \zeta_m)$ , the total equivalent damping ratio increases as piezoelectric coefficient increases. So the maximum displacement of the PVEH decreases as the piezoelectric coefficient increases as shown in Fig. 13. Therefore, both the input power and damping loss decrease as piezoelectric coefficient increases as shown in Fig. 17. However, the damping loss decreases more quickly. So the power output increases as piezoelectric coefficient increases. When  $K_{ef} \geq 4\zeta_m(1 + \zeta_m)$ , the  $\zeta_e$  is basically unchanged as shown in Fig. 16. So the total equivalent damping ratio is also unchanged. Therefore, both the maximum displacement and the power output remain the same with the piezoelectric coefficient as shown in Figs. 12 and 13. Thus, higher piezoelectric coefficient is not always better for PVEH, which is consistent with Renno et al. (2009) conclusion. The conclusion has important significance for MEMS PVEH due to the difficult preparation of piezoelectric thin film.

**Fig. 14** The power output with piezoelectric coefficient at given amplitude of PVEH

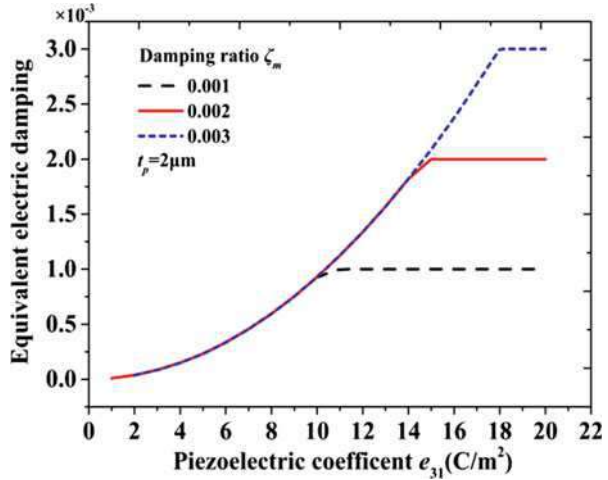


**Fig. 15** The required acceleration with piezoelectric coefficient at given amplitude of PVEH

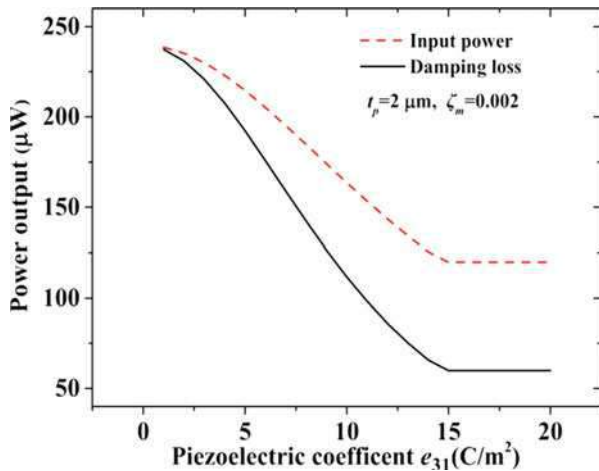


Figures 18 and 19 show that the power output and maximum displacement of PVEH change with the thickness of piezoelectric layer at given acceleration. Figures 20 and 21 show that the power output and required acceleration change with the thickness of piezoelectric layer at given amplitude of PVEH. It can be found that both the thickness of piezoelectric layer and the piezoelectric coefficient have similar effects on the performance of the PVEH. And the interpretation of the changing trends about the performance versus the thickness of piezoelectric layer curves is also the same as the interpretation of the changing trends about the performance versus the thickness of piezoelectric layer curves. Figures 18, 19, 20, and 21 show that optimal thickness of the piezoelectric layer is related to damping ratio. The smaller the damping ratio is, the thinner the piezoelectric layer will be.

**Fig. 16** Equivalent electric damping with piezoelectric coefficient



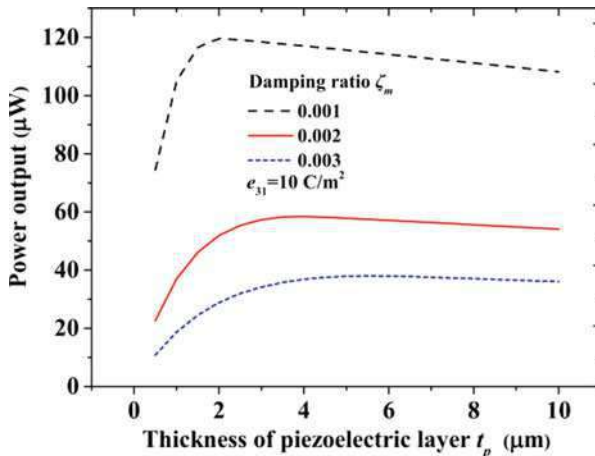
**Fig. 17** The input power and damping loss with piezoelectric coefficient at given acceleration



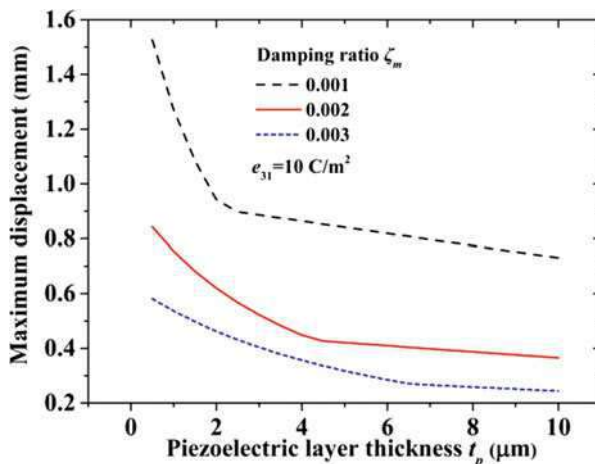
## Fabrication of PVEH

Both AlN and PZT are commonly used in MEMS PVEH, so in this section, we intended to give the readers an understanding of the full microfabrication process flow of PVEHs based on AlN and PZT, including AlN thin films based PVEH which operate in  $d_{31}$  mode, PZT thin films based PVEH which operate in  $d_{31}$  mode and in  $d_{33}$  mode, and bulk PZT thick films based PVEH which operate in  $d_{31}$  mode. Each of the main aspects covers the detailed fabrication descriptions of several highlighted energy harvesters.

**Fig. 18** The power output with the thickness of piezoelectric layer at given acceleration



**Fig. 19** The maximum displacement with the thickness of piezoelectric layer at given acceleration

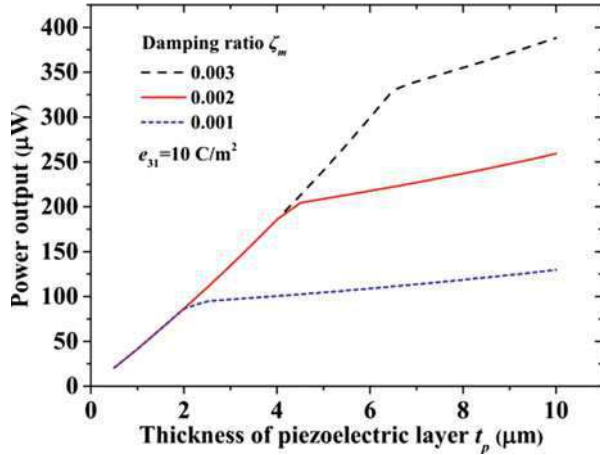


### The Fabrication of the AlN Thin Films Based PVEH

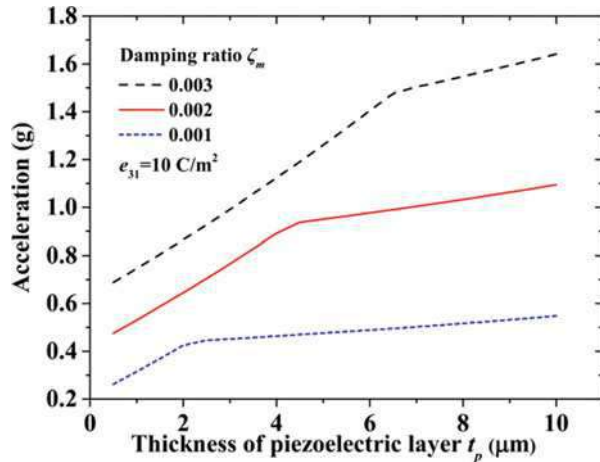
Among the piezoelectric thin film materials, such as PZT, ZnO, and AlN, AlN is a promising energy harvesting material for its interesting material properties, including good electromechanical coupling coefficients, low permittivity, and high Young’s modulus. Furthermore, although AlN does not have very high piezoelectric coefficients compared to PZT, it is a nonferroelectric material which does not require poling. Additionally, AlN can be grown by low temperature techniques, such as sputtering deposition, making it fully compatible with standard CMOS technology (Elfrink et al. 2009).

A few groups have proposed AlN-based MEMS energy harvester. Zhang et al. (2011) prepared low residual stress AlN thin films on Si substrates and fabricated MEMS PVEH with the AlN thin films. The device can generate 1.42  $\mu\text{W}$  from 1 g

**Fig. 20** The power output with the thickness of piezoelectric layer at given amplitude of PVEH



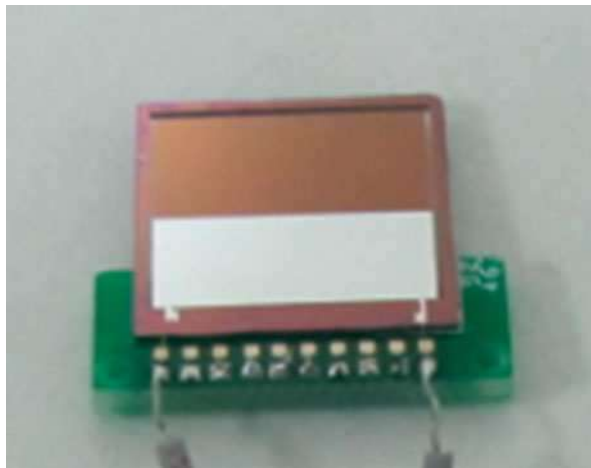
**Fig. 21** The required acceleration with the thickness of piezoelectric layer at given amplitude of PVEH



acceleration at its resonant frequency of 1042 Hz. Defosseux et al. (2012) developed an AlN VEH for low frequency and acceleration vibration environments. About  $0.62 \mu\text{W}$  average power can be obtained at  $0.25 \text{ g}$  acceleration and  $214 \text{ Hz}$  frequency. Elfrink et al. (2009, 2010, 2011) have carried out a series of researches on the AlN PVEH. The latest results show that a harvester produced about  $6 \text{ V}$  peak load voltage and  $489 \mu\text{W}$  average power at  $4.5 \text{ g}$  acceleration and  $1011 \text{ Hz}$  frequency, which is the highest power output so far for AlN thin film PVEH. However, the resonant frequency and acceleration amplitude are mismatch with ambient vibration sources. Bertacchini et al. (2011) have demonstrated a novel MEMS AlN PVEH which can harvest energy from different vibration orientation, additionally the device have a wide frequency range from  $500 \text{ Hz}$  to  $1.5 \text{ kHz}$ . However, the output power is only  $0.4 \text{ pW}$ .

The research group of Zhiyu Wen (He and Gao 2013; He et al. 2013; Shang et al. 2013; Zhao et al. 2015; Zhengguo et al. 2013) from Chongqing University

**Fig. 22** The AlN-based PVEH

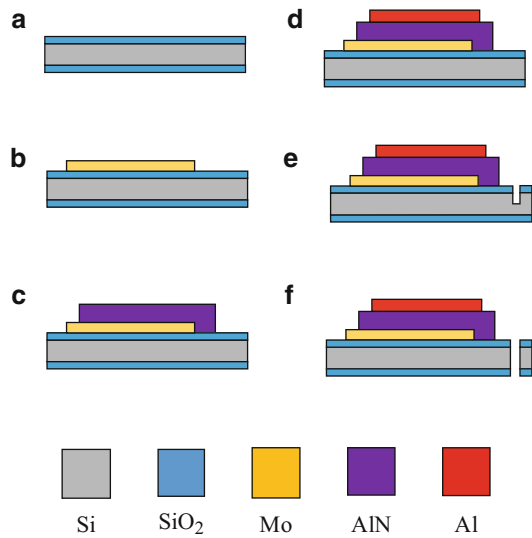


systematically studied the AlN based PVEH, and one of the fabricated AlN based PVEH can generate  $81.32 \mu\text{W}$  from 1 g acceleration at its resonant frequency of 241.2 Hz, as shown in Fig. 22. The dimensions of the cantilever is  $5 \text{ mm} \times 12.4 \text{ mm} \times 0.05 \text{ mm}$  and the dimensions of the proof mass is  $6 \text{ mm} \times 12.4 \text{ mm} \times 0.05 \text{ mm}$ . The microfabrication process flow is shown in Fig. 23. The fabrication starts from a double-sided-polished, n-type 4-in. (100) silicon wafer which is  $500.0 \pm 10.0 \mu\text{m}$  thick. Before depositing multilayer structures, the wafer is cleaned in RCA1 and RCA2 at  $80 \text{ }^\circ\text{C}$  for 10 min. Then silicon oxygen isolation layer (300 nm thick  $\text{SiO}_2$ ) is deposited at  $1050 \text{ }^\circ\text{C}$  (Fig. 23a). The next step is series of deposition and patterning: 300 nm thick molybdenum (Mo) film as bottom electrode. At the same time, selecting Mo as bottom electrode can obtain better piezoelectric performance of AlN. 1.0  $\mu\text{m}$  thick piezoelectric material AlN layer and 400 nm thick Al film as top electrode (Fig. 23b–d). The AlN layer was deposited by pulsed-DC magnetron sputtering and the details has been presented in reference (Zhengguo et al. 2013). Next step is etching a 50  $\mu\text{m}$  depth trench from the frontside to define the cantilevers by deep-reactive ion etching (Fig. 23e). The last step is backside etching to form the proof mass and release the structures (Fig. 23f).

### The Fabrication of the PZT Thin Films Based PVEH

PZT is the most popular piezoelectric material utilized for energy harvesters due to its high piezoelectric coefficient, and high quality PZT thin films have been prepared in PVEH fabrication through sol-gel method and pulsed laser deposition (PLD) technique and so on. There are two configurations used for PVEH,  $d_{31}$  mode and  $d_{33}$  mode. A  $d_{31}$  mode piezoelectric cantilever has a PZT layer in between the top and bottom electrodes, while a  $d_{33}$  mode piezoelectric cantilever eliminates a bottom electrode by utilizing IDEs on the top of a PZT layer.

**Fig. 23** The process flow of AlN PVEH



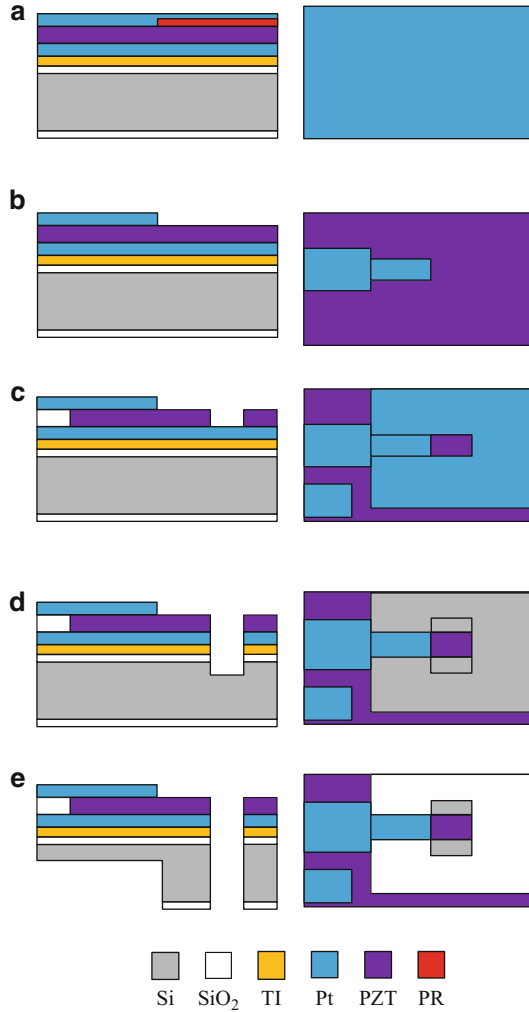
### The Fabrication of the $d_{31}$ Mode PZT Thin Films PVEH

The D. J. Kim group (Shen et al. 2008, 2009; Kim et al. 2013) has made a great effort on PZT thin film PVEH with sol-gel method. Recent studies show that the PVEH generated 0.46 V peak load voltage and 2.15  $\mu\text{W}$  average power at 0.5 g acceleration and 243 Hz frequency. Fang et al. (2006) have fabricated a PZT thin film VEH which generated 2.16  $\mu\text{W}$  average power at 1 g acceleration and 609 Hz frequency. However, the proof mass was not integrated with the cantilever, which would be an additional difficulty in production. Jambunathan et al. (2012) fabricated a PZT thin film PVEH, and the PZT thin film was prepared by pulsed laser deposition (PLD) technique. The PVEH generated 51  $\mu\text{W}$  average power at 0.9 g acceleration and 615 Hz frequency. Liu et al. (2011) developed a PZT thin film PVEH with low resonance frequency and wide operation bandwidth. Unlike the traditional single cantilever PVEH, the PVEH consists of a silicon proof mass and a silicon supporting beam which is integrated with ten parallel-arrayed PZT energy harvesting elements. Experimental results show that PZT elements connected in series or in parallel produce similar power into matched load and the results also show that the operation frequency bandwidth can be widened to 17 Hz (30–47 Hz). However, the maximum power output is only 51.3 nW.

In this case study, a unimorph 31-mode PZT PVEH reported by D. J. Kim group (Shen et al. 2008) was selected. The microfabrication process of the PZT PVEH was shown in Fig. 24. The fabrication process began with a 100 mm low-cost bare silicon wafer. SiO<sub>2</sub> was grown on both sides of the silicon wafer by a wet O<sub>2</sub> method. Interlayer Ti and electrode Pt were deposited one after another by magnetron sputtering. PZT was deposited layer by layer using the sol-gel method to reach a thickness of 1  $\mu\text{m}$ . The top electrode was patterned by the first mask and obtained by a lift-off process after Pt deposition on a layer of patterned photoresist (PR). The

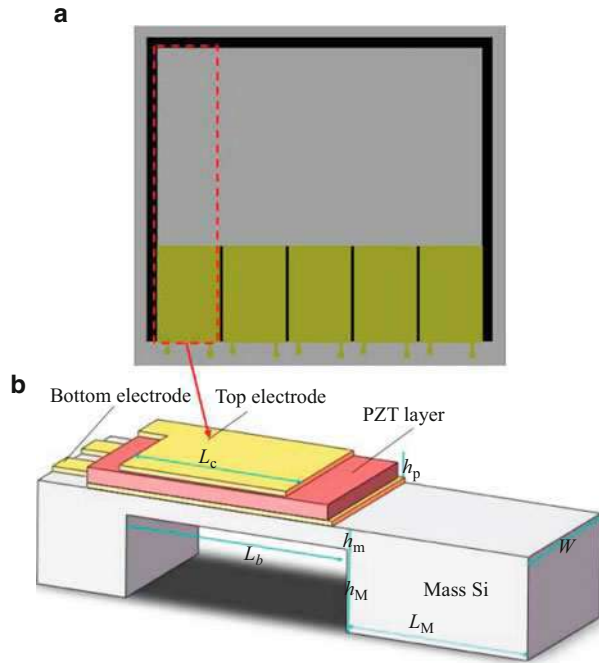


**Fig. 24** The microfabrication process of PZT PVEH



bottom electrode together with the cantilever structure was patterned by the second mask and etched by inductively coupled plasma (ICP) reactive ion etching (RIE) to create a bare window for easy access during wire bonding. The cantilever structure without the bottom electrode was then patterned by the third mask and etched by RIE. The backside proof mass was finally patterned by the fourth mask and the cantilever structure was finally released after backside Si RIE. The fabricated device was shown in Fig. 25. The device has beam dimensions of about  $4.800 \text{ mm} \times 0.400 \text{ mm} \times 0.036 \text{ mm}$  with an integrated Si proof mass with dimension of about  $1.360 \text{ mm} \times 0.940 \text{ mm} \times 0.456 \text{ mm}$ . Experiment results show that the device produced  $160 \text{ mV}_{\text{pk}}$ ,  $2.15 \text{ }\mu\text{W}$  from  $2 \text{ g}$  acceleration at its resonant frequency of  $461.15 \text{ Hz}$ .

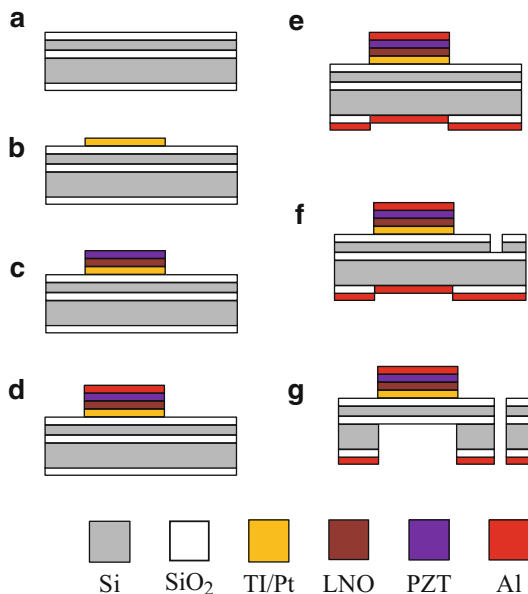
**Fig. 25** Schematics of PZT thin film MEMS PVEH, the overall structure (a) and the details structure of single cantilever (b)



For the  $d_{31}$  mode PVEH, the voltage output is proportional to the thickness of the piezoelectric, and the thickness of the PZT thin film is usually 1–4  $\mu\text{m}$ . so the voltage output of the  $d_{31}$  mode PZT PVEH is less than 1 V. In practical applications, diode bridges rectifier circuit are often utilized after the PVEH. The typical voltage drop of diode is 0.3–0.8 V. So the low voltage output of PVEH will cause most of power loss in rectifier circuit, even cause the rectifier circuit to turn off and no power transferred. To solve these problems, the Zhiyu Wen group from Chongqing University proposed a piezoelectric cantilevers array PVEH (Deng et al. 2014) in which the piezoelectric cantilevers array are connected with one silicon proof mass as shown in Fig. 25. The piezoelectric cantilevers array connected with one silicon proof mass configuration can achieve series connecting and thereby improving voltage output, and the configuration can also ensure high power output. The microfabrication process of the proposed device is shown in Fig. 26.

The fabrication process starts with a 4-in. silicon-on-insulator (SOI) wafer consisted of a 47  $\mu\text{m}$  thick device layer, a 1  $\mu\text{m}$  buried oxide layer, and a 453  $\mu\text{m}$  thick silicon substrate. A  $\text{SiO}_2$  (0.3  $\mu\text{m}$ ) layer was deposited on the wafer using Plasma Enhanced Chemical Vapor Deposition (PECVD) technique as shown in Fig. 26a. Interlayer Ti (50 nm) and bottom electrode Pt (150 nm) were obtained by lift-off photolithography technique after Ti and Pt were deposited on a layer of photo resist which was patterned by the first mask as shown in Fig. 26b. Ti was used for better adhesion between the  $\text{SiO}_2$  layer and the bottom electrode Pt. Then  $\text{LaNiO}_3$  (LNO) (60 nm) and PZT (3  $\mu\text{m}$ ) layer were coated sequentially by sol-gel method followed by

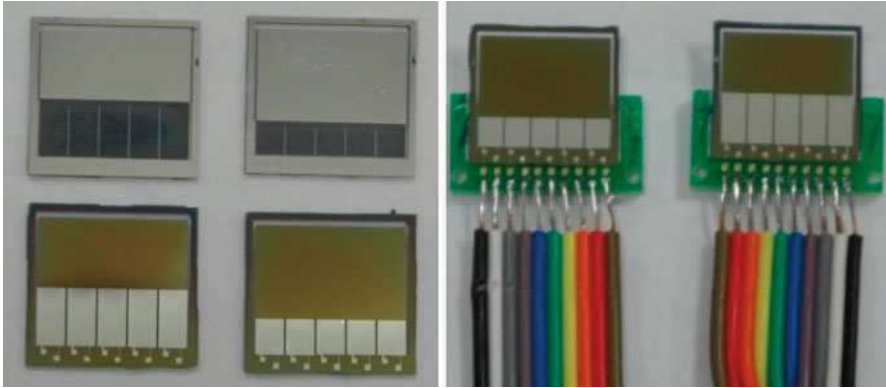
**Fig. 26** The microfabrication flow of the PVEH



a series of heat treatment respectively. LNO thin films easily form (100)/(001) oriented PZT thin films, which are suitable for piezoelectric MEMS devices. Furthermore, the LNO will reduce the degradation in the ferroelectric and piezoelectric properties of PZT thin films derived from a MEMS microfabrication process (Kobayashi et al. 2007, 2008). Both PZT and LNO layer were patterned by the second mask; see Fig. 26c. Thereafter, Al (120 nm) was deposited and patterned as the top and bottom electrode by the third mask, as shown in Fig. 26d. For deep silicon etched of backside, an Al layer was sputtered and patterned by the fourth mask after the SiO<sub>2</sub> of backside was etched, as shown in Fig. 26e. Then the beams structure were patterned by the fifth mask and etched by inductively coupled plasma (ICP) after the 0.3 μm SiO<sub>2</sub> layer was etched, as shown in Fig. 26f. In the backside, the Al layer was etched and used as mask to define beams shape through ICP process; the beams shape can be patterned until the buried SiO<sub>2</sub> layer of the SOI wafer was removed completely through reactive ion etching (RIE) process, as shown in Fig. 26g.

All the fabricated devices were poled at 150 °C with applying a field of 12.5 V/μm for 30 min. To evaluate, the fabricated PVEHs were glued to printed circuit boards (PCB) and the bonding pads were connected to the PCB pads by gold wires through wires bonding technique, as shown in Fig. 27.

The fabricated devices have the dimensions of each cantilever as 3 mm × 2.4 mm × 0.052 mm and the proof mass dimensions as 8 mm × 12.4 mm × 0.5 mm. The harvester produces 5.876 V RMS open circuit voltage with 204.51 μW power for 1 g acceleration at the frequency 233.5 Hz.



**Fig. 27** The fabricated PVEH

### The Fabrication of the $d_{33}$ Mode PZT Thin Films PVEH

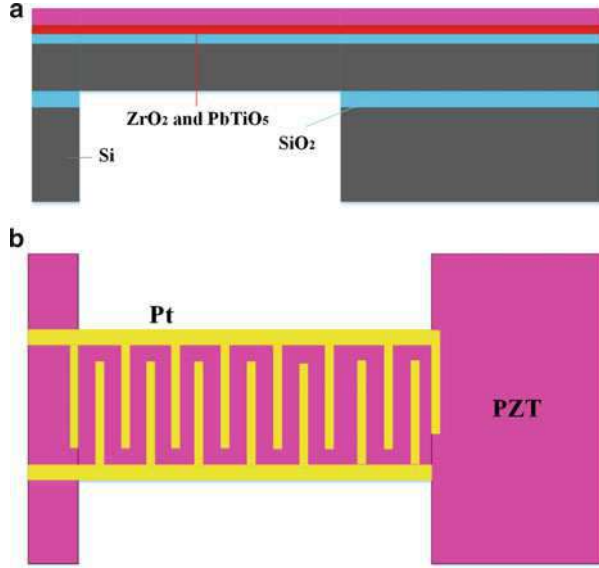
Normally, the  $d_{33}$  constant is large than the  $d_{31}$  constant. For PZT, the  $d_{33}$  constant is commonly known to be approximately twice as large as the  $d_{31}$  constant. Therefore, designing a  $d_{33}$  structure properly could produce more energy at  $d_{31}$  structure. The open-circuit voltage of the PZT thin film PVEH can be expressed as:

$$V_{3i} = \sigma_{xx} g_{3i} L_i \quad (47)$$

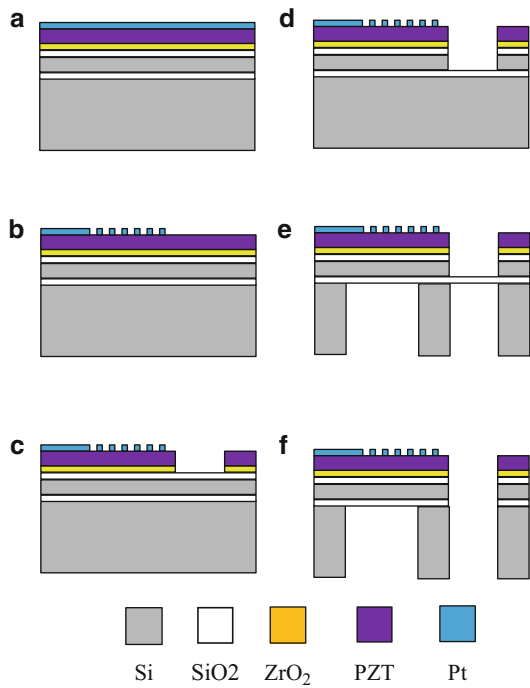
Where  $\sigma_{xx}$  is the stress,  $g_{3i}$  (Vm/N) the piezoelectric constants, and  $L_i$  is the distance between electrodes which could be either  $t_{\text{pzt}}$  (for  $d_{31}$  mode) or  $L_3$  (for  $d_{33}$  mode). The  $d_{31}$  coefficient is directly proportional to  $g_{31}$  via the dielectric coefficient ( $\epsilon_{\text{pzt}}$ ) of the piezoelectric. In the  $d_{31}$  mode case,  $t_{\text{pzt}}$  from Eq. 47 is limited to the thickness of the PZT layer, while  $L_3$  can be independently long in the  $d_{33}$  case. In addition, the magnitudes of the  $d_{33}$  and  $g_{33}$  coefficients are 2–2.5 times larger than their  $d_{31}$  and  $g_{31}$  counterparts. So the  $d_{33}$  mode device can be easy to obtain high open-circuit voltage, and the open-circuit voltage is easy to adjust by changing the electrode spacing  $L_3$ . This then enables the activation of the rectifying circuit.

AlN and ZnO are nonferroelectric materials, so the polarization direction of the prepared thin film is fixed in one direction. When this thin film was used to fabricate  $d_{33}$  mode PVEH, the PVEH will exist charge offset. So the  $d_{33}$  mode PVEH always achieve with PZT thin film. Jong Cheol Park et al. (2010) modeled and fabricated a  $d_{33}$ -mode MEMS PVEH. The schematics of the  $d_{33}$  mode MEMS PVEH was shown in Fig. 28. Figure 29 shows the fabrication process. The fabrication of the  $d_{33}$  mode MEMS PVEH started with a silicon-on-insulator (SOI) wafer with a 10  $\mu\text{m}$  thick device layer and a 1- $\mu\text{m}$ -thick buried silicon oxide layer. First, thermal oxidation was performed to obtain a silicon oxide layer with a thickness of 1500  $\text{\AA}$  for the electric diffusion barrier. Then, a ZrO<sub>2</sub> buffer layer with a thickness of 2000  $\text{\AA}$  was deposited, and a PbTiO<sub>3</sub> seed layer was applied as an interlayer between the ZrO<sub>2</sub> and PZT thin films. A PZT thin film with a thickness of 1  $\mu\text{m}$  was deposited with a Zr/Ti ratio of 52/48. These thin films were processed using the sol-gel method. A Pt

**Fig. 28** Schematics of the  $d_{33}$  mode MEMS PVEH, (a) sectional view, (b) vertical view



**Fig. 29** The fabrication flows of the  $d_{33}$  MEMS PVEH



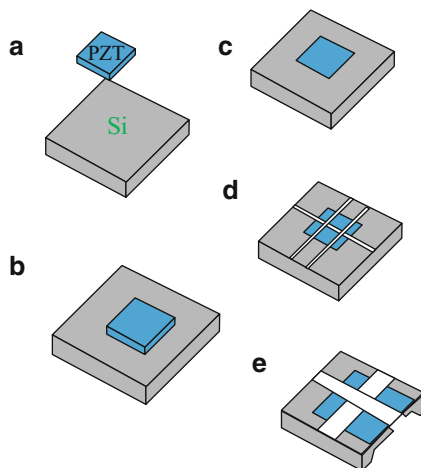
layer was sputtered to form interdigital electrodes with a width and spacing of 5  $\mu\text{m}$  (Fig. 29a). The Pt and PZT layers were patterned by the inductively coupled plasma (ICP) dry etching technique (Fig. 29b, c). A silicon cantilever beam and an inertial mass were defined by front and backside silicon deep reactive-ion etching (RIE), respectively (Fig. 29d, e). In particular, the buried silicon oxide was utilized as an etch stop for deep RIE to control the thickness of the cantilever beam accurately. Finally, the device was released by etching the buried oxide (Fig. 29f). Experimental results show that the fabricated device generated an electrical power of 1.1  $\mu\text{W}$  for a load of 2.2  $\text{M}\Omega$  with 4.4 VP-P from a vibration with an acceleration of 0.39 g at its resonant frequency of 528 Hz. The corresponding power density was  $7.3 \text{ mW}\cdot\text{cm}^{-3}\cdot\text{g}^{-2}$ .

Based on aerosol deposition PZT thin film, Lee et al. (2009) fabricated two piezoelectric MEMS generators,  $d_{31}$  mode PVEH and  $d_{33}$  mode PVEH. The fabricated  $d_{33}$  mode device possessed a maximum open circuit output voltage of 4.127 VP-P and a maximum output power of 1.288  $\mu\text{W}$  with 2.292 VP-P output voltage at its resonant frequency of 214 Hz at a 2 g acceleration. While the fabricated  $d_{31}$  mode device possessed a maximum open circuit output voltage of 2.415 VP-P and a maximum output power of 2.099  $\mu\text{W}$  with 1.587 VP-P output voltage at its resonant frequency of 255.9 Hz at a 2 g acceleration. In order to investigate the power output of the  $d_{31}$  and  $d_{33}$  mode PVEH,  $d_{31}$  and  $d_{33}$  mode PZT thin film PVEHs were fabricated by Kim et al. (2013). Both types of devices have the same dimensions in a cantilever structure. The resonant frequencies of the cantilevers are 243 Hz. Experimental results show that the output power of the  $d_{31}$  mode PVEH was 2.15  $\mu\text{W}$ . The  $d_{33}$  mode PVEHs generated output power ranging between 0.62 and 1.71  $\mu\text{W}$  when the width of the interdigital electrode (IDE) is ranging from 8 to 16  $\mu\text{m}$  and finger spacing is varied from 4 to 16  $\mu\text{m}$ . The above two results show that the power output for  $d_{33}$  mode PVEH is not higher than the  $d_{31}$  mode PVEH. It may be attributed to the easier poling of  $d_{31}$  mode PZT thin film (Wu and Lee 2012; Knight et al. 2011).

## The Fabrication of the PZT Thick Films PVEH

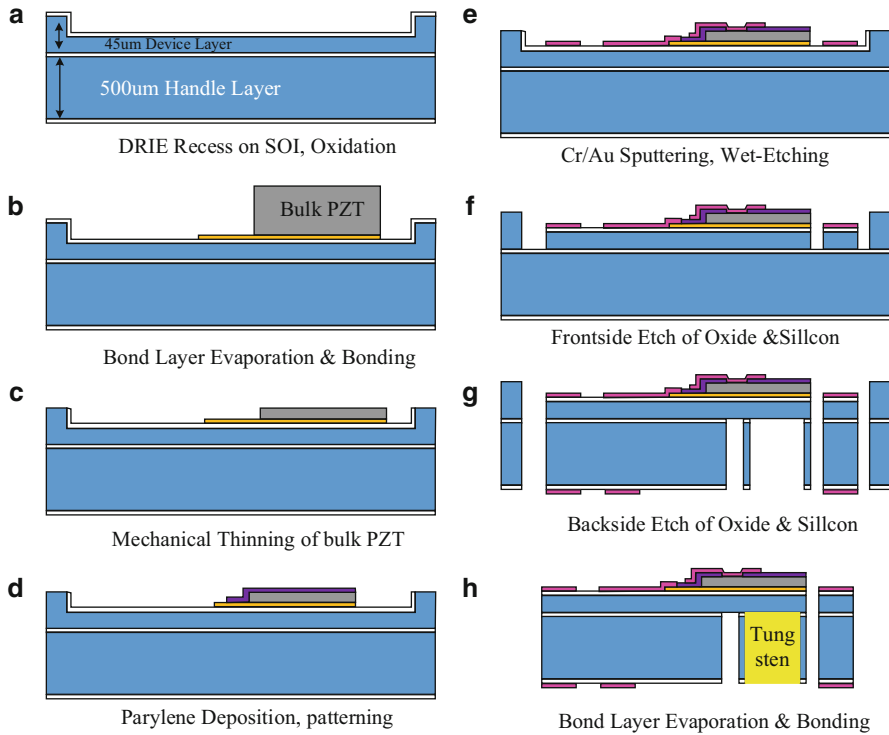
The common deposition processes of PZT and AlN using sol-gel and sputtering methods limit the thin film thickness within 1~4  $\mu\text{m}$ . Literature (Wang et al. 1999) shows that the optimization electromechanical coupling coefficient appears in that situation when the thickness ratio of the substrate to piezoelectric layer is less than 2. However, the thickness ratio for the MEMS based PVEH is always more than 10, so increasing the piezoelectric layer thickness can effectively increase the electromechanical coupling coefficient and then the electromechanical conversion efficiency improves. Recently, the preparation of piezoelectric thick film has drawn wide attention. Some research institutions (such as University of Michigan (Aktakka et al. 2011, 2013), École Polytechnique Fédérale de Lausanne (EPFL) (Janphuang et al. 2012, 2014), and Shanghai Jiao tong University (Tang et al. 2011, 2012) have devoted to prepare piezoelectric thick films utilizing bonding process and thinning process. The bulk piezoelectric material is integrated on a silicon wafer by a bonding

**Fig. 30** The basic fabrication process of cantilever beam structures with bulk PZT



process firstly and then the piezoelectric layer with a desired thickness through thinning process. Fig. 30 shows the basic fabrication process (Tang et al. 2012). Firstly, the surfaces of bulk PZT and silicon wafers are polished, as shown in Fig. 30a. Then, the two wafers are bonded together, as shown in Fig. 30b. Following that, the bulk PZT wafer is thinned to a suitable thickness and the thinned PZT wafer is micromachined into the desired pattern by the dicing method. Finally, the back surface of the silicon wafer is structured with a wet or dry etching process. Screen printing is also used for preparing thick film. Technical University of Denmark (Lei et al. 2011; Xu et al. 2012) has developed the preparation of PZT thick film based on screen printing technology and has fabricated PZT/PZT thick film bimorph vibration energy harvester.

Figure 31 shows the fabrication process of the PZT thick film harvester from University of Michigan (Aktakka et al. 2011). The PZT thick film was prepared through bonding process and thinning process, as shown in Fig. 31b, c. In order to better control the PZT thickness and wafer-level uniformity during mechanical thinning of the PZT, a deep groove was etched firstly through DRIE (Fig. 31a), the silicon wafer surface is used as a stop/masking layer. Then the bulk PZT was bond on the groove by low-temperature (200 °C) AuIn transient liquid phase bonding (Fig. 31b). The AuIn layer was also used for bottom electrode. And then the bulk PZT was thinned by mechanical thinning process, the thickness of the PZT layer is 20  $\mu\text{m}$  (Fig. 31c). The next step is to deposit an insulation layer between the top and bottom electrodes (Fig. 31d). The Parylene is selected for insulation layer due to its stress-free conformal evaporation at room temperature and very low Young's modulus. After that patterning the Parylene insulation layer, sputtering Cr/Au and wet-etching Cr/Au to achieve the preparation of the top and bottom electrodes (Fig. 31e). The oxide and silicon of the foreside and backside were etched for releasing the device structures (Fig. 31f, g). Finally, a tungsten was bond by low-temperature (200 °C) AuIn transient liquid phase bonding (Fig. 31h). The fabricated devices with the dimensions of  $7 \times 7 \times 0.55 \text{ mm}^3$  can produce 205  $\mu\text{W}$  at 1.5 g (154 Hz) at resonance.



**Fig. 31** Fabrication flows of the thinned-PZT PVEH

## Summary and Challenges

This chapter provided an introduction to the fundamental theory and the typical microfabrication of MEMS piezoelectric vibration energy harvester. Firstly, the application background of the vibration energy harvester was introduced, and the advantages of piezoelectric vibration energy harvesters were illustrated. Then the working principle of the piezoelectric vibration energy harvester was introduced, including piezoelectric transducing mechanism, piezoelectric equation, and the piezoelectric materials commonly used for piezoelectric vibration energy harvester. After that, the theoretical model was deduced through Hamilton's principle and Euler-Bernoulli beam theory. Based on the theoretical model, the optimization of the piezoelectric vibration energy harvester was performed, and the optimization results show that the electromechanical conversion efficiency is not more than 50% and the PVEH with low damping ratio does not always have high power output, which challenges previous literature suggestion that lower damping ratio tends to higher power output. Finally, the microfabrication of the typical MEMS-based piezoelectric energy harvesters was described.



Recently, in order to improve the performance of the MEMS PVEH, researchers have focused on improving both energy conversion and energy harvesting, respectively. To improving energy conversion, researchers focus on the preparation of high performance piezoelectric film, including the preparation of high performance PZT (Isarakorn et al. 2011) and AlN films (Jackson et al. 2013), doping modified PZT film (Van Minh et al. 2016) and AlN film (Barth et al. 2016; Nguyen et al. 2016), and developing new high performance piezoelectric materials (Won et al. 2016; Gupta et al. 2016; Lee et al. 2016). Jackson et al. (2013) studied the effect of AlN film orientation on AlN properties. Experiment results show that the presence of non-(002) AlN peaks significantly decreases the piezoelectric constant  $d_{31}$ , in a material, and this therefore affects the amount of energy that can be harvested from a MEMS PVEH. Barth et al. (2016) focus on the influence of Sc doping on piezoelectric, mechanical, and energy harvesting properties for AlN films deposited by reactive pulse magnetron sputtering. The prepared films show the expected increase of piezoelectric properties as well as the softening of the material with higher Sc concentrations. The fabricated PVEH based on AlScN films can generate 350  $\mu$ W power as well as 70  $\mu$ W for pure AlN. In order to improve energy harvesting, broadband PVEHs were studied (Leadenham and Erturk 2015; Harne and Wang 2013). However, limited to the MEMS process technology, the broadband MEMS PVEHs are hard to achieve, and new structure and new mechanism based broadband MEMS PVEH which can be fabricated through MEMS process technology should be explored.

---

## References

- Aktakka EE, Peterson RL, Najafi K (2011) Thinned-PZT on SOI process and design optimization for piezoelectric inertial energy harvesting. In: Solid-state sensors, actuators and microsystems conference (TRANSDUCERS), 2011 16th international: IEEE, pp 1649–1652
- Aktakka EE, Peterson RL, Najafi K (2013) Wafer-level integration of high-quality bulk piezoelectric ceramics on silicon. *IEEE Trans Electron Devices* 60:2022–2030
- Anton SR, Sodano HA (2007) A review of power harvesting using piezoelectric materials (2003–2006). *Smart Mater Struct* 16:R1
- Barth S, Bartzsch H, Glöß D, Frach P, Modes T, Zywitzki O, Suchanek G, Gerlach G (2016) Magnetron sputtering of piezoelectric AlN and AlScN thin films and their use in energy harvesting applications. *Microsyst Technol* 22:1–5
- Bertacchini A, Scorcioni S, Dondi D, Larcher L, Pavan P, Todaro M T, Campa A, Caretto G, Petroni S, Passaseo A (2011) AlN-based MEMS devices for vibrational energy harvesting applications. In: 2011 proceedings of the European solid-state device research conference (ESSDERC): IEEE, pp 119–122
- Defosseux M, Allain M, Defay E, Basrou S (2012) Highly efficient piezoelectric micro harvester for low level of acceleration fabricated with a CMOS compatible process. *Sensors Actuators A Phys* 188:489–494
- Deng L, Wen Z, Zhao X, Yuan C, Luo G, Mo J (2014) High voltage output MEMS vibration energy harvester in  $d_{31}$  mode with PZT thin film. *J Microelectromech Syst* 23:855–861
- Dubois-Ferrière H, Fabre L, Meier R, Metrailler P (2006) TinyNode: a comprehensive platform for wireless sensor network applications. In: Proceedings of the 5th international conference on Information processing in sensor networks: ACM, pp 358–365

- Elfrink R, Kamel TM, Goedbloed M, Matova S, Hohlfeld D, Van Anandel Y, Van Schaijk R (2009) Vibration energy harvesting with aluminum nitride-based piezoelectric devices. *J Micromech Microeng* 19:94005
- Elfrink R, Renaud M, Kamel TM, De Nooijer C, Jambunathan M, Goedbloed M, Hohlfeld D, Matova S, Pop V, Caballero L (2010) Vacuum-packaged piezoelectric vibration energy harvesters: damping contributions and autonomy for a wireless sensor system. *J Micromech Microeng* 20:104001
- Elfrink R, Matova S, de Nooijer C, Jambunathan M, Goedbloed M, van de Molengraft J, Pop V, Vullers R, Renaud M, van Schaijk R (2011) Shock induced energy harvesting with a MEMS harvester for automotive applications. In: *Electron devices meeting (IEDM), 2011 I.E. international: IEEE*, pp 25–29
- Erturk A, Inman DJ (2011) Appendix A: piezoelectric constitutive equations. In: *Piezoelectric energy harvesting*, Wiley, New York, pp 343–348
- Fan F, Tian Z, Wang ZL (2012) Flexible triboelectric generator. *Nano Energy* 1:328–334
- Fang H, Liu J, Xu Z, Dong L, Wang L, Chen D, Cai B, Liu Y (2006) Fabrication and performance of MEMS-based piezoelectric power generator for vibration energy harvesting. *Microelectron J* 37:1280–1284
- Gupta MK, Kim S, Kumar B (2016) Flexible high-performance lead-free Na<sub>0.47</sub>K<sub>0.47</sub>Li<sub>0.06</sub>NbO<sub>3</sub> microcube-structure-based piezoelectric energy harvester. *Acs Appl Mater Inter* 8:1766–1773
- Hande A, Polk T, Walker W, Bhatia D (2007) Indoor solar energy harvesting for sensor network router nodes. *Microprocess Microsyst* 31:420–432
- Harne RL, Wang KW (2013) A review of the recent research on vibration energy harvesting via bistable systems. *Smart Mater Struct* 22:23001
- He X, Gao J (2013) Wind energy harvesting based on flow-induced-vibration and impact. *Microelectron Eng* 111:82–86
- He X, Shang Z, Cheng Y, Zhu Y (2013) A micromachined low-frequency piezoelectric harvester for vibration and wind energy scavenging. *J Micromech Microeng* 23:125009
- Isarakorn D, Briand D, Janphuang P, Sambri A, Gariglio S, Triscone JM, Guy F, Reiner JW, Ahn CH, de Rooij NF (2011) The realization and performance of vibration energy harvesting MEMS devices based on an epitaxial piezoelectric thin film. *Smart Mater Struct* 20:25015
- Jackson N, O Keeffe R, Waldron F, O'Neill M, Mathewson A (2013) Influence of aluminum nitride crystal orientation on MEMS energy harvesting device performance. *J Micromech Microeng* 23:75014
- Jambunathan M, Elfrink R, Vullers R, van Schaijk R, Dekkers M, Broekmaat J (2012) Pulsed laser deposited-PZT based MEMS energy harvesting devices. In: *Proceedings of ISAF-ECAPD-PFM 2012: IEEE*, pp 1–4
- Janphuang P, Lockhart R, Briand D, de Rooij NF (2012) Wafer level fabrication of vibrational energy harvesters using bulk PZT sheets. *Procedia Engineering* 47:1041–1044
- Janphuang P, Lockhart R, Uffer N, Briand D, de Rooij NF (2014) Vibrational piezoelectric energy harvesters based on thinned bulk PZT sheets fabricated at the wafer level. *Sensors Actuators A Phys* 210:1–9
- Kim S, Park H, Kim S, Wikle HC, Park J, Kim D (2013) Comparison of MEMS PZT cantilevers based on and modes for vibration energy harvesting. *J Microelectromech Syst* 22:26–33
- Knight RR, Mo C, Clark WW (2011) MEMS interdigitated electrode pattern optimization for a unimorph piezoelectric beam. *J Electroceram* 26:14–22
- Kobayashi T, Ichiki M, Kondou R, Nakamura K, Maeda R (2007) Degradation in the ferroelectric and piezoelectric properties of Pb (Zr, Ti) O<sub>3</sub> thin films derived from a MEMS microfabrication process. *J Micromech Microeng* 17:1238
- Kobayashi T, Ichiki M, Kondou R, Nakamura K, Maeda R (2008) Fabrication of piezoelectric microcantilevers using LaNiO<sub>3</sub> buffered Pb (Zr, Ti) O<sub>3</sub> thin film. *J Micromech Microeng* 18:35007
- Leadnham S, Erturk A (2015) Nonlinear M-shaped broadband piezoelectric energy harvester for very low base accelerations: primary and secondary resonances. *Smart Mater Struct* 24:55021

- Ledermann N, Murali P, Baborowski J, Gentil S, Mukati K, Cantoni M, Seifert A, Setter N (2003) {1 0 0}-textured, piezoelectric Pb (Zr x, Ti 1– x) O 3 thin films for MEMS: integration, deposition and properties. *Sensors Actuators A Phys* 105:162–170
- Lee BS, Lin SC, Wu WJ, Wang XY, Chang PZ, Lee CK (2009) Piezoelectric MEMS generators fabricated with an aerosol deposition PZT thin film. *J Micromech Microeng* 19:65014
- Lee G, Shin D, Kwon Y, Jeong S, Koh J (2016) Optimized piezoelectric and structural properties of (Bi, Na) TiO 3–(Bi, K) TiO 3 ceramics for energy harvester applications. *Ceram Int* 42 (13):14355–14363
- Lei A, Xu R, Thyssen A, Stoot AC, Christiansen TL, Hansen K, Lou-Moeller R, Thomsen EV, Birkelund K (2011) MEMS-based thick film PZT vibrational energy harvester. In: *Micro electro mechanical systems (MEMS), 2011 I.E. 24th international conference on IEEE*, pp 125–128
- Liu H, Tay CJ, Quan C, Kobayashi T, Lee C (2011) Piezoelectric MEMS energy harvester for low-frequency vibrations with wideband operation range and steadily increased output power. *J Microelectromech Syst* 20:1131–1142
- Meitzler AH, Tiersten HF, Warner AW, Berlincourt D, Couquin GA, Welsh III FS (1988) *IEEE standard on piezoelectricity*. IEEE, New York
- Nguyen, H H, Oguchi H, Kuwano H (2016). Combinatorial approach to MgHf co-doped AlN thin films for Vibrational Energy Harvesters. *Journal of Physics Conference Series* 773:012075
- Park JC, Park JY, Lee Y (2010) Modeling and characterization of piezoelectric-mode MEMS energy harvester. *J Microelectromech Syst* 19:1215–1222
- Preumont A (2006) *Mechatronics: dynamics of electromechanical and piezoelectric systems*. Kluwer Academic Publishers, Springer, Dordrecht, Netherlands
- Renno JM, Daqaq MF, Inman DJ (2009) On the optimal energy harvesting from a vibration source. *J Sound Vib* 320:386–405
- Saadon S, Sidek O (2011) A review of vibration-based MEMS piezoelectric energy harvesters. *Energy Convers Manage* 52:500–504
- Shang Z, Li D, Wen Z, Zhao X (2013) Flow-induced-vibration MEMS energy harvesting unit based on AlN film in silicon. *Opt Precis Eng* 12:11
- Shen D, Park J, Ajitsaria J, Choe S, Wickle HC III, Kim D (2008) The design, fabrication and evaluation of a MEMS PZT cantilever with an integrated Si proof mass for vibration energy harvesting. *J Micromech Microeng* 18:55017
- Shen D, Park J, Noh JH, Choe S, Kim S, Wickle HC, Kim D (2009) Micromachined PZT cantilever based on SOI structure for low frequency vibration energy harvesting. *Sensors Actuators A Phys* 154:103–108
- Suzuki Y (2011) Recent progress in MEMS electret generator for energy harvesting. *IEEJ T Electr Electr* 6:101–111
- Tang G, Liu JQ, Liu HS, Li YG, Yang CS, He DN, DzungDao V, Tanaka K, Sugiyama S (2011) Piezoelectric MEMS generator based on the bulk PZT/silicon wafer bonding technique. *Phys Status Solidi (a)* 208:2913–2919
- Tang G, Liu J, Yang B, Luo J, Liu H, Li Y, Yang C, He D, Dao VD, Tanaka K (2012) Fabrication and analysis of high-performance piezoelectric MEMS generators. *J Micromech Microeng* 22:65017
- Ujihara M, Carman GP, Lee DG (2007) Thermal energy harvesting device using ferromagnetic materials. *Appl Phys Lett* 91:93508
- van Minh L, Sano T, Fujii T, Kuwano H (2016) Vibrational micro-energy harvesters utilizing Nb-doped Pb(Zr,Ti)O<sub>3</sub> films on stainless steel substrates. *Journal of Physics Conference Series* 773:012003
- Wang Q, Du X, Xu B, Eric Cross L (1999) Electromechanical coupling and output efficiency of piezoelectric bending actuators. *IEEE Trans Ultrason Ferroelectr Freq Control* 46:638–646
- Won SS, Lee J, Venugopal V, Kim D, Lee J, Kim IW, Kingon AI, Kim S (2016) Lead-free Mn-doped (K<sub>0.5</sub>, Na<sub>0.5</sub>) NbO<sub>3</sub> piezoelectric thin films for MEMS-based vibrational energy harvester applications. *Appl Phys Lett* 108:232908

- Wu WJ, Lee BS (2012) Piezoelectric MEMS power generators for vibration energy harvesting. In: Lallart M (ed) *Small-scale energy harvesting*. InTech, pp 156–157
- Xu R, Lei A, Dahl-Petersen C, Hansen K, Guizzetti M, Birkelund K, Thomsen EV, Hansen O (2012) Fabrication and characterization of MEMS-based PZT/PZT bimorph thick film vibration energy harvesters. *J Micromech Microeng* 22:94007
- Zhang F, Wang L (2002). *Modern piezoelectricity*. Science Press, Beijing.
- Zhang J, Cao Z, Kuwano H (2011) Fabrication of low-residual-stress aln thin films and their application to microgenerators for vibration energy harvesting. *Jpn J Appl Phys* 50:9N–18N
- Zhao X, Shang Z, Luo G, Deng L (2015) A vibration energy harvester using AlN piezoelectric cantilever array. *Microelectron Eng* 142:47–51
- Zhengguo S, Dongling L, Zhiyu W, Xingqiang Z (2013) The fabrication of vibration energy harvester arrays based on AlN piezoelectric film. *J Semicond* 34:114013
- Zorlu Ö, Topal ET, Kulah H (2011) A vibration-based electromagnetic energy harvester using mechanical frequency up-conversion method. *IEEE Sensors J* 11:481–488



# Triboelectric Nanogenerators

Chi Zhang and Zhong Lin Wang

## Contents

Overview .....	1336
Fundamental Principle Modes of Triboelectric Nanogenerators .....	1337
Vertical Contact-Separation Mode .....	1337
Lateral Sliding Mode .....	1338
Single-Electrode Mode .....	1338
Freestanding Triboelectric-Layer Mode .....	1339
TENG as Microscale Power Source .....	1339
Woven-Structured TENG for Wearable Devices .....	1339
Multilayered Electret Film-Based TENG .....	1341
Cover-Sheet-Based TENG for Charging Mobile Electronics .....	1341
Liquid-Metal-Electrode-Based High-Performance TENG .....	1344
Rotary Disk-Based Direct-Current TENG .....	1345
Multilayered Cylindrical TENG for Enhancing Output Current .....	1346
Radial Grating-Based High-Performance TENG .....	1348
Contact- and Noncontact-Mode TENG for Harvesting Energy from Automobile Brake .....	1349
TENG with Pulsed Output Current for Efficient Charging of Li-ion Batteries .....	1351
TENG as Mega-scale Power Source .....	1351
Networks of TENGs .....	1353
Multilayer Wavy-Structured Robust TENG .....	1354
Structural Optimization of TENG for Harvesting Water Wave Energy .....	1355
Air-Driven Membrane-Structured TENG Array .....	1355

---

C. Zhang (✉)

Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China

e-mail: [czhang@binn.cas.cn](mailto:czhang@binn.cas.cn)

Z. L. Wang

Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China

School of Material Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA

© Springer Nature Singapore Pte Ltd. 2018

1335

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,

[https://doi.org/10.1007/978-981-10-5945-2\\_38](https://doi.org/10.1007/978-981-10-5945-2_38)

TENG as Self-Powered Active Sensors .....	1357
Self-Powered Wireless Sensing Node .....	1357
Transparent Paper-Based Page Mark and Anti-theft Sensor .....	1361
Self-Powered Velocity and Trajectory Tracking Sensor Array .....	1362
Self-Powered Vehicle Collision Acceleration Sensor .....	1362
Self-Powered Crack Detector and Rotating Speed Sensor .....	1365
TENG for Tribo-Controlled Devices .....	1365
Comparison of EMIG and TENG .....	1366
TENG as a Control Source for MEMS Actuators .....	1368
Triboelectrification-Induced Ultraviolet Emission from Plasmon Discharge .....	1370
Triboelectrification-Induced Trace Memorization Device .....	1371
Summary and Perspectives .....	1371
Acknowledgments .....	1373
References .....	1373

---

### Abstract

Triboelectric nanogenerator (TENG) is a new energy technology for converting human kinetic and ambient mechanical energy into electricity. The principle of the TENG is based on triboelectrification and electrostatic induction, in which the induced triboelectric charges can generate a potential drop and drive electron flow by a mechanical force. Since invented in 2012, the TENG has made rapid research progress and demonstrated various potential applications.

This chapter will first introduce the working principle and four fundamental modes of the TENG and emphatically present the different applications of the TENG, such as microscale power source for wearable and portable electronics, mega-scale power source by harvesting water wave energy, self-powered active sensors for internet of things, and triboelectric-voltage-controlled source for capacitive devices.

The goal of this chapter is to provide the reader with general concept, principle and modes of the TENG as a new energy technology, and a broader overview of potential applications that can be utilized to harvest human kinetic and ambient mechanical energy.

---

### Keywords

Triboelectric nanogenerator · Triboelectrification · Vertical contact-separation · Lateral sliding · Single-electrode · Freestanding triboelectric-layer · Microscale power source · Mega-scale power source · Self-powered active sensors · Tribo-controlled devices

---

## Overview

There are all kinds of mechanical energy in our living environment, such as deformation, musculation, vibration, wind, and ocean wave. By using nanotechnology, they can be effectively collected and converted into electricity for self-powered micro-/nano-system. Because mechanical energy from human and environment is

everywhere all the time, it is an optimal solution for the requirements of sustainable power supply (Wang 2008; Wang and Wu 2012).

Trieboelectrification is a very common phenomenon in various mechanical movements, which is always neglected for the difficulty in collection or considered as a negative effect for decreasing the mechanical efficiency. If the trieboelectricity could be collected from human kinetic and ambient mechanical energy, it will produce important applications in potable electronics and sensor networks. Up to now, electrostatic micro-generator has been successfully developed by MEMS technology (Beeby et al. 2007), but the silicon-based device has special production conditions, complicated process, and high cost. In 2012, trieboelectric nano-generator (TENG) as a new energy technology has been invented by Prof. Zhong Lin Wang's group (Fan et al. 2012; Wang 2014). The TENG is based on organic and metal films without expensive materials and equipment. It has easy processing, long life, and convenient integration for potential large-scale industrial production and practical application.

In this chapter, four fundamental principle modes of the TENG are introduced, and various applications of the TENG are summarized in details, such as micro-scale power source, mega-scale power source, self-powered active sensors, and tribo-controlled devices. Most of these research works are achieved by Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences.

---

## Fundamental Principle Modes of Trieboelectric Nanogenerators

The trieboelectric effect is the phenomenon that a material becomes electrically charged after it contacts a different material through friction, which has been known for thousands of years (Henniker 1962). When two different materials come into contact, trieboelectric charges are produced on the surfaces. While when separated, electrons are driven in the electrode to flow in order to balance the electric potential drop created. Based on such a principle, the four different modes of TENGs are invented, as elaborated in the following.

### Vertical Contact-Separation Mode

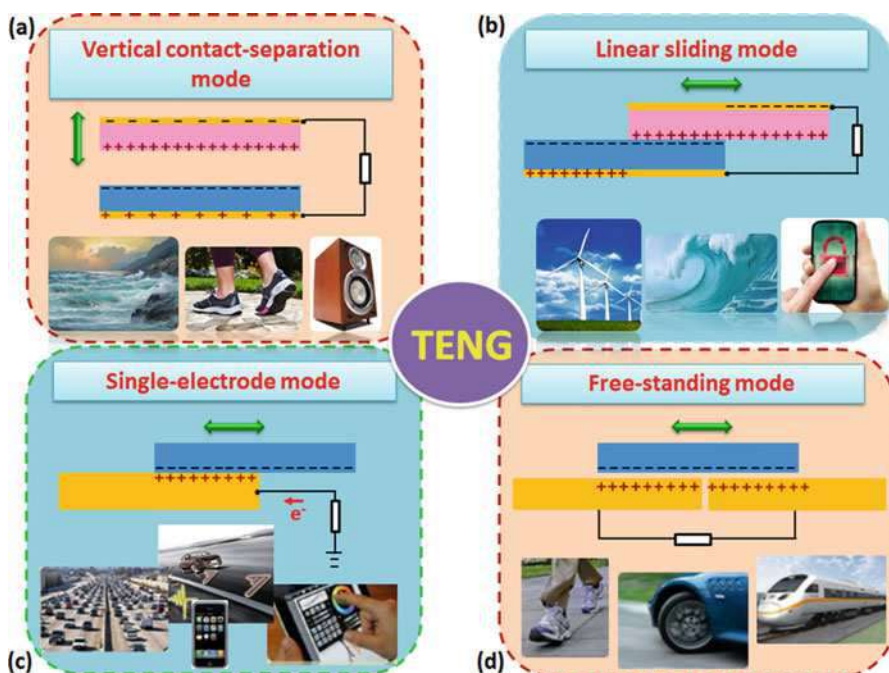
The vertical contact-separation mode as the first mode is shown in Fig. 1a (Zhu et al. 2012; Wang et al. 2012). Two dissimilar dielectric films face with each other, and there are electrodes deposited on the top and the bottom surfaces of the stacked structure. A physical contact between the two dielectric films creates oppositely charged surfaces. Once the two surfaces are separated by a small gap by the lifting of an external force, a potential drop is created. If the two electrodes are electrically connected by a load, free electrons in one electrode would flow to the other electrode to build an opposite potential in order to balance the electrostatic field. Once the gap is closed, the trieboelectric charge-created potential disappears, and the electrons flow back (Niu et al. 2013).

## Lateral Sliding Mode

The structure in lateral sliding mode has the same initial status as that of the vertical contact-separation. When two dielectric films are in contact, a relative sliding in parallel to the surface also creates triboelectric charges on the two surfaces, as shown in Fig. 1b (Wang et al. 2013; Zhu et al. 2013a). A lateral polarization is thus introduced along the sliding direction, which drives the electrons on the top and bottom electrodes to flow in order to fully balance the field created by the triboelectric charges. A periodic sliding apart and closing generates an AC output. This is the sliding mode TENG. The sliding can be a planar motion, a cylindrical rotation, or disk rotation (Jing et al. 2014; Lin et al. 2013).

## Single-Electrode Mode

The above two modes have two electrodes interconnected by a load. Such TENGs can freely move so that they can work for mobile cases. In some cases, the object that is part of the TENG cannot be electrically connected to the load because it is a



**Fig. 1** The four fundamental modes of triboelectric nanogenerators (TENGs). (a) Vertical contact-separation mode, (b) in-plane contact-sliding mode, (c) single-electrode mode, and (d) freestanding triboelectric-layer mode (Reprinted with permission from Wang 2014. Copyright 2014 Royal Society of Chemistry)



mobile object, such as a human walking on a floor. In order to harvest energy from such a case, a single-electrode TENG is introduced, in which the electrode on the bottom part of the TENG is grounded as shown in Fig. 1c. If the size of the TENG is finite, an approaching or departing of the top object from the bottom one would change the local electrical field distribution, so that there are electron exchanges between the bottom electrode and the ground to maintain the potential change of the electrode. This energy harvesting strategy can be in both contact-separation mode and sliding mode (Yang et al. 2013; Zhang et al. 2013; Niu et al. 2014a).

### **Freestanding Triboelectric-Layer Mode**

The freestanding triboelectric-layer mode is shown in Fig. 1d (Wang et al. 2014a). A moving object is naturally charged due to its contact with air or other object, such as our shoes walking on floors that are usually charged. The charges remain on the surface for hours, and contact or friction is unnecessary within this period of time because the charge density reaches a maximum. If we make a pair of symmetric electrodes underneath a dielectric layer, and the size of the electrodes and the gap distance between the two are of the same order as the size of the moving object, the object's approach to and/or departure from the electrodes creates an asymmetric charge distribution in the media, which causes the electrons to flow between the two electrodes to balance the local potential distribution. The oscillation of the electrons between the paired electrodes produces power. The moving object does not have to touch directly the top dielectric layer of the electrodes so that, in rotation mode, free rotation is possible without direct mechanical contact; wear of the surfaces can be drastically reduced. This is a good approach for extending the durability of the TENGs. Such a design has shown the potential for harvesting energy from a freely moving object without an electric connection.

---

### **TENG as Microscale Power Source**

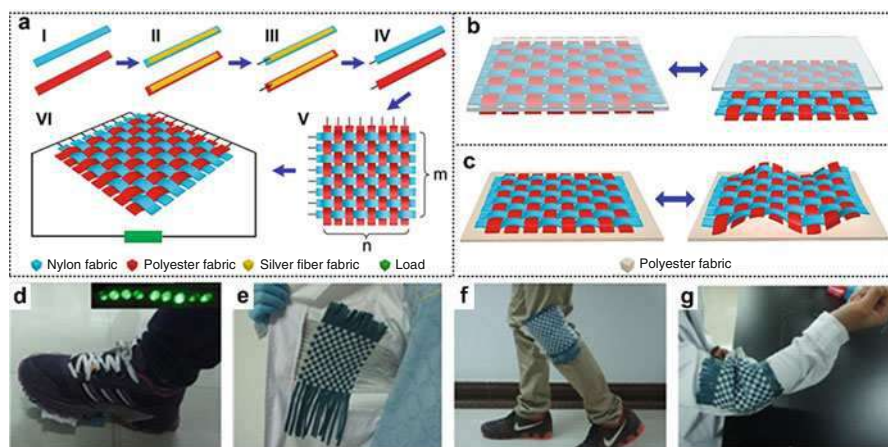
The first and direct goal of developing TENGs is to power small electronics. Based on the four fundamental principle modes illustrated above, we have developed various TENGs depending on specific applications. By harvesting various types of mechanical energy, these TENGs can be provided as microscale power source for small electronics.

### **Woven-Structured TENG for Wearable Devices**

To date quite a few wearable electronics have entered the market, which are changing the life pattern of consumers. However, the limited lifetime and energy storage capacity have made rechargeable batteries the bottleneck in wearable technology especially with the increase of wearable devices and their large distribution.

To solve this problem, a woven-structured triboelectric nanogenerator (W-TENG) using commodity nylon fabric, polyester fabric, and conductive silver fiber fabric was demonstrated (Zhou et al. 2014). The silver fiber fabric that served as the electrodes is a kind of mixed textile made of Ag fibers and cotton fibers. Since it is washable, flexible, and highly conductive ( $0.3 \Omega/\text{cm}^2$ ), silver fiber fabric is more suitable for wearable TENGs than conventional metal electrodes or carbon electrodes. The structure and fabrication process of a  $M \times N$  W-TENG were shown in Fig. 2a, where the  $M$  nylon electrodes and the  $N$  polyester electrodes were connected in parallel, respectively. In general, the W-TENG has two basic working modes: deformation and non-deformation mode. In the non-deformation mode (Fig. 2b), a freestanding acrylic triboelectric layer is driven by a vertical force to contact with and separate from the W-TENG periodically. In the deformation mode (Fig. 2c), the W-TENG is driven by a lateral force to stretch and compress periodically on the surface of the polyester fabric which acts as the freestanding triboelectric layer. Different from other TENG, electricity is generated by a freestanding triboelectric layer that moves relatively on the surface of the W-TENG rather than the relative movement between the positive and negative electrodes.

With the advantage of being flexible, washable, breathable, wearable, and able to be triggered by a freestanding triboelectric layer, this W-TENG can move freely without any constraint and is suitable for wearable electronics. To demonstrate the potential applications of the W-TENG, the W-TENG is integrated into shoes, coats,



**Fig. 2** Woven-structured triboelectric nanogenerator (W-TENG). (a) Schematic diagram illustrating the fabricating process of a W-TENG. (b) Structure and working process of the W-TENG under non-deformation mode, in which acrylic works as the freestanding triboelectric layer. (c) Structure and working process of the W-TENG under deformation mode, where polyester fabric works as the freestanding triboelectric layer. (d) A power generating shoe (P-shoe) which could harvest energy from footsteps and light up about nine LEDs. (e) The W-TENG integrated in a coat harvesting energy from the shaking of clothes. (f) The W-TENG harvesting energy from leg joints bent. (g) The W-TENG harvesting energy from arm joints (Reprinted with permission from Zhou et al. 2014. Copyright 2014 American Chemical Society)

and trousers to harvest different kinds of mechanical energy from human motion (walking, movement of joints, shaking or deformation of clothes, etc.), Fig. 2d–g. This work has presented a new approach in applying triboelectric nanogenerator to wearable devices.

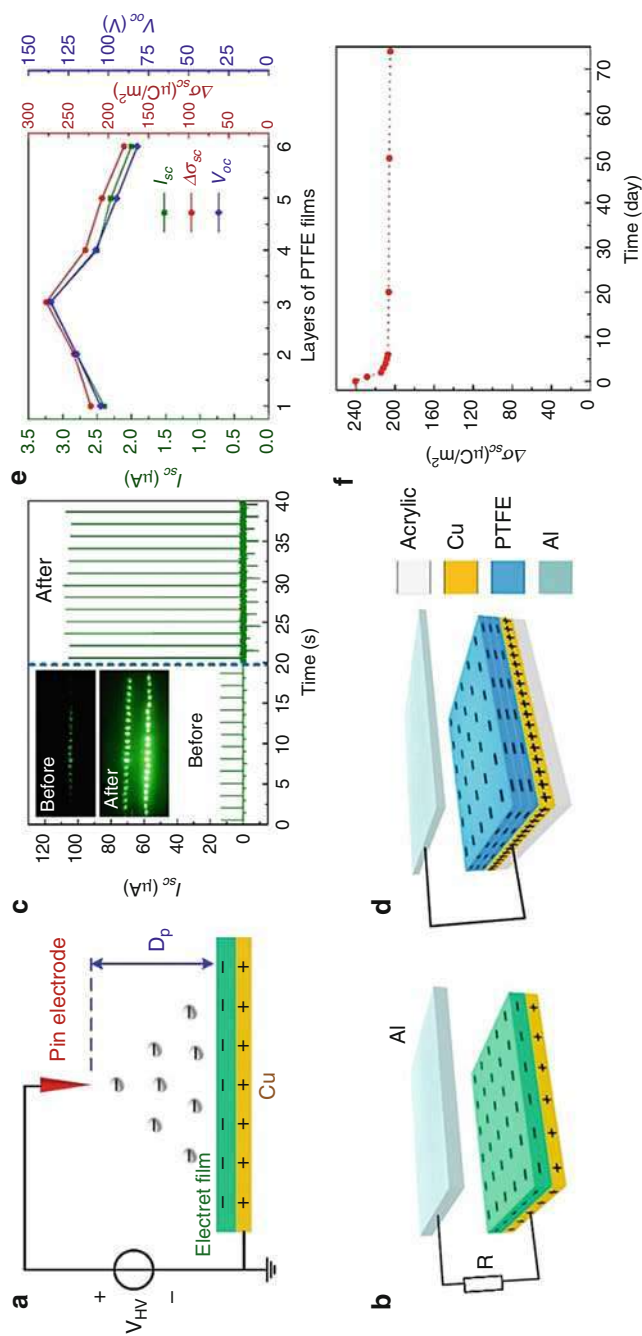
## Multilayered Electret Film-Based TENG

To enhance the effective surface charge density as well as the output performances of TENG, a novel electret film-based TENG (E-TENG) fabricated by corona charging has been first proposed (Zhou et al. 2016). Electret is a kind of dielectrics that can quasi-permanently preserve large amount of charges. By corona charging, plenty of negative charges can be injected onto the electret film (Fig. 3a) resulting in greatly enhanced surface charge density which can be several times higher than that of conventional thin films. Figure 3b shows the structure of the E-TENG in vertical contact-separation mode, where many materials can serve as the electret film such as polytetrafluoroethylene (PTFE), fluorinated ethylene propylene (FEP), amorphous fluoropolymer CYTOP™ (Asahi Glass Co., Ltd.), and perfluorinated amorphous polymer Teflon® AF (DuPont). After charge injection at the optimal corona charging conditions of voltage, pinpoint distance, and time ( $V_{HV} = -7$  kV,  $D_P = 1$  cm and  $T_C = 2$  min), a PTFE film ( $7 \times 7$  cm  $\times$  50  $\mu$ m in size)-based E-TENG can produce the short-circuit current  $I_{SC}$  of 108  $\mu$ A (Fig. 3c), transferred electric charge density  $\Delta\sigma_{SC}$  of 240  $\mu$ C/m<sup>2</sup>, and open-circuit voltage  $V_{OC}$  of 2450 V, which are all about seven times larger than that of the ordinary PTFE film-based TENG.

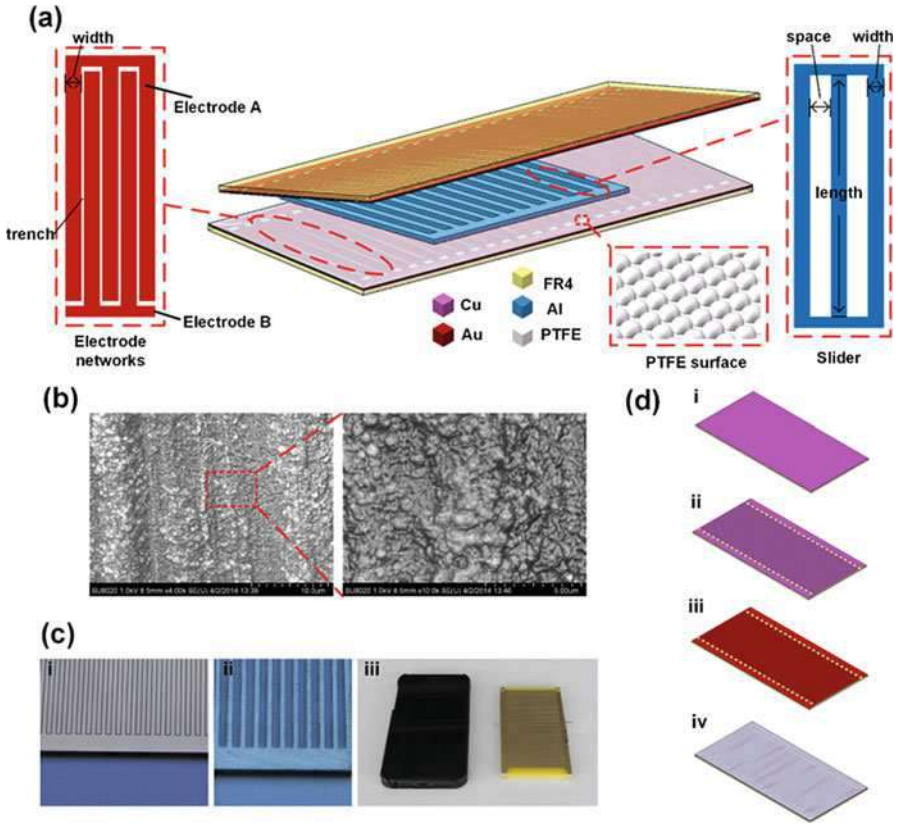
Based on the corona charging, several multilayered E-TENGs were fabricated to further improve the output performances, in which the three-layered E-TENG (Fig. 3d) with the transferred electric charge density of 278  $\mu$ C/m<sup>2</sup> was proven to be the most effective (Fig. 3e). The E-TENGs can successfully light up LEDs (Fig. 3c) and a self-powered LCD display system with the enhanced output performances. Therefore, it has been proven that the injection of charges into electret films by corona charging is a simple, effective, and cost-efficient approach to enhance the effective surface charge density and thereby the output performances of the TENG with excellent durability and good stability (Fig. 3f), which is quite beneficial to the applications of the TENG in self-powered portable electronics and sensor networks.

## Cover-Sheet-Based TENG for Charging Mobile Electronics

As for mobile electronics, it is generally desired that adding a power harvester should not add much weight, volume, or additional components to the electronics, and the general requirement is to build the energy harvester using the available infrastructure as the base. In this work, the authors invented a cover-sheet-based TENG (CS-TENG) in sliding mode by using the protection structure of electronics such as a cover of a smartphone (Tang et al. 2014a). As shown in Fig. 4, it consists of two epoxy resin sheets, on which two complementary micro-grating electrodes were



**Fig. 3** Multilayered electret film-based triboelectric nanogenerator (TEN-G). **(a)** Charge injection principle of the electret film by corona charging. **(b)** Structure of the electret film-based TEN-G (E-TENG) in vertical contact-separation mode. **(c)** Comparisons of short-circuit current  $I_{sc}$  before and after the corona charging process. Insets in **(c)** show the LEDs lit up by the TEN-G before and after corona charging ( $V_{HV} = -7$  kV,  $D_p = 1$  cm, and  $T_c = 2$  min). **(d)** Structure of the three-layered E-TENG. **(e)**  $I_{sc}$ ,  $\Delta\sigma_{sc}$ , and  $V_{oc}$  of the multilayered E-TENGs with PTFE films of  $1.3 \times 1.3$  cm  $\times$  80  $\mu$ m by corona charging ( $V_{HV} = -7$  kV,  $D_p = 1$  cm, and  $T_c = 2$  min). **(f)** Stability of the surface charges on PTFE film-based E-TENG fabricated by corona charging in conditions of  $V_{HV} = -7$  kV,  $D_p = 1$  cm, and  $T_c = 2$  min (Reprinted with permission from Zhou et al. 2016. Copyright 2016 Springer)

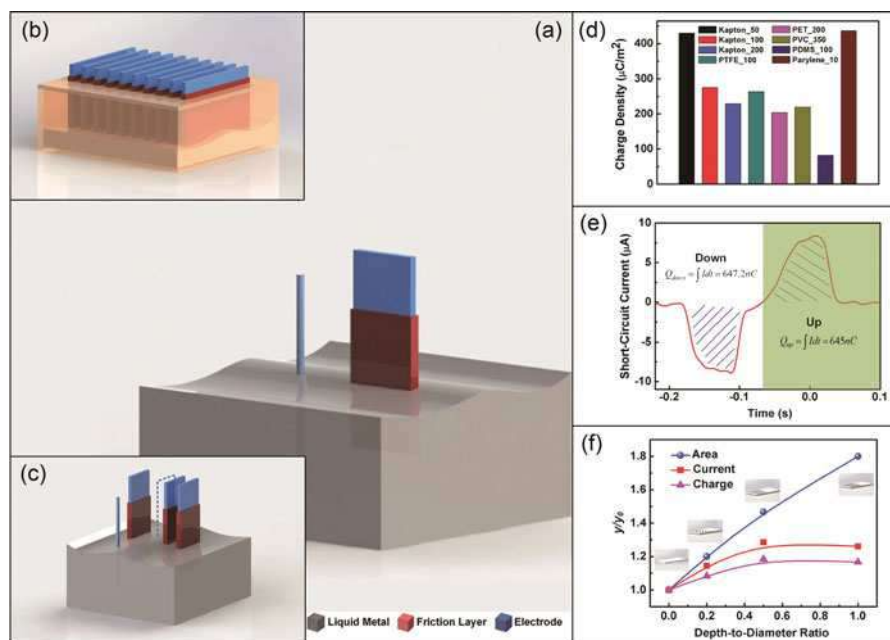


**Fig. 4** Cover-sheet based triboelectric nanogenerator (CS-TENG). (a) Schematic illustrations of the CS-TENG. (b) Scanning electron microscopy (SEM) images of the pre-treated PTFE’s surface. (c) Photographs of the stator, the slider and the packaged devices (compared to the size of an iPhone5 at left). (d) Fabrication sketches of the stator by printing circuit technology (Reprinted with permission from Tang et al. 2014a. Copyright 2016 Elsevier Ltd)

placed. Operating at an in-plane sliding velocity of 1 m/s, the CS-TENG has a contacting area of 18 cm<sup>2</sup> (the cover’s total area is 50 cm<sup>2</sup>, smaller than the normal size of a smartphone) and a thickness of 3 mm and could generate an alternative current (AC) around 88.8 μA, with an average effective power density of 29.3 W/m<sup>2</sup> and 9.77 kW/m<sup>3</sup>. It successfully powered up household light bulbs, demonstrating the capability of the CS-TENG as a power supply for regular electronics. Due to the cover-shape design, the CS-TENG was integrated with a portable temperature meter, which can be driven by hand motions without any battery, opening a range of applications of TENG in conventional sensor systems. As an approach that is cost-effective, simple implementing, scalable, and industrially available, the CS-TENG can be widely applied for a variety of self-powered portable electronics.

## Liquid-Metal-Electrode-Based High-Performance TENG

All of the TENGs reported today are normally based on solid materials, so that the effectiveness of contact, especially to the nanometer level, can be largely affected by the roughness of the two surfaces and the match between the two. According to the literature, the previous TENG's charge density is about  $100 \mu\text{C}/\text{m}^2$  (Lin et al. 2013b; Du et al. 2014). Meanwhile, the solid-solid friction will result in heat generation and dissipation. Those two largely limit the efficiency of the TENG. In this work, the first liquid-metal-based triboelectric nanogenerator (LM-TENG) is developed, as shown in Fig. 5 (Tang et al. 2015). Operating at a separating velocity of  $0.25 \text{ m/s}$ , the LM-TENG having a contact area of  $15 \text{ cm}^2$  could generate a voltage of  $679 \text{ V}$  and a current of  $9 \mu\text{A}$ . More importantly, its output charge density reaches  $430 \mu\text{C}/\text{m}^2$ , which is over four times higher than that of the solid-solid-contact TENG, and its power density reaches  $6.7 \text{ W}/\text{m}^2$  and  $133 \text{ KW}/\text{m}^3$ . Owing to the shape-adaptive advantage, with a 20-layer configuration, the LM-TENG achieves a high current output above  $130 \mu\text{A}$ . In addition, the liquid-metal-based TENG's energy conversion



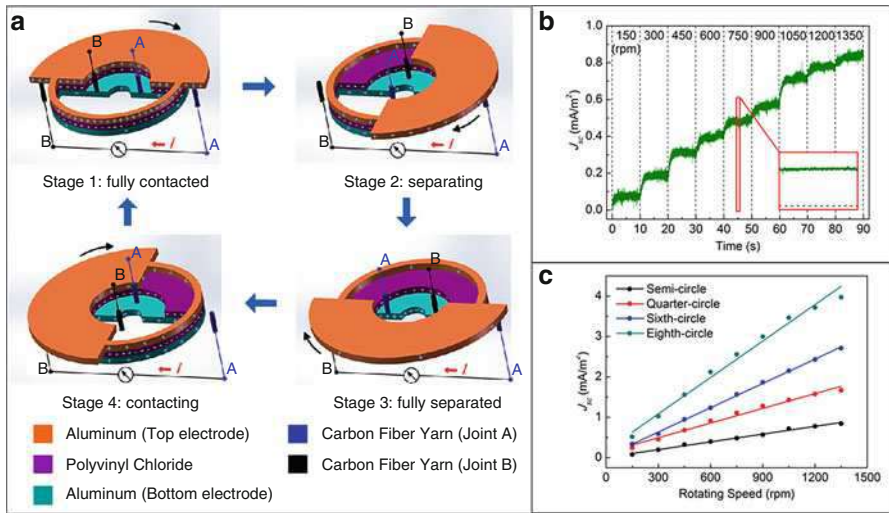
**Fig. 5** The liquid-solid contact triboelectric nanogenerator based on liquid metal (LM-TENG). (a, b, c) Schematic illustration showing the device configuration of the single, multilayered and packaged LM-TENGs. (d) Dependences of the transferred charge density of the LM-TENG on various friction materials. (e) Measurement of the short-circuit current of the LM-TENG. (f) Varying tendencies of the surface area, current, and transferred charges for the four kinds of surface morphologies (here  $y$  represents the value measured at each depth-to-diameter ratio and  $y_0$  represents the value of the plane structure) (Reprinted with permission from Tang et al. 2015. Copyright 2015 John Wiley & Sons, Inc)



efficiency is demonstrated to be 70.6%, even higher than the previous reported 55%. Furthermore, the LM-TENG is applied for vibration energy harvesting. At amplitude of 1.2 mm and a frequency of 10 Hz, the LM-TENG outputs a voltage of 145 V and a current of 5.8  $\mu\text{A}$  and charges a 100  $\mu\text{F}$  capacitor to 1 V in 65 s. This work presents a new approach for enhancing the TENG's output efficiency and further broadening its applications.

### Rotary Disk-Based Direct-Current TENG

An innovative design of a direct-current triboelectric nanogenerator (DC-TENG) based on a rotating disk design for harvesting rotational mechanical energy is reported (Zhang et al. 2014a). The DC-TENG consists of two disks and two pairs of flexible electric brushes that are made of carbon fiber and contact two electrodes, respectively (Fig. 6a). This design not only facilitates a rotation-induced periodic, in-plane charge separation for electricity generation but also introduces a direct-current generation method with the flexible electric brushes. During the rotation, two disks have distinct triboelectric polarities for a cyclic in-plane charge separation between them, and an alternating current is generated between the two electrodes. In the first half cycle, the top electrode contacts with the electric brush of joint A, and the bottom electrode contacts with the electric brush of joint B, forming a current



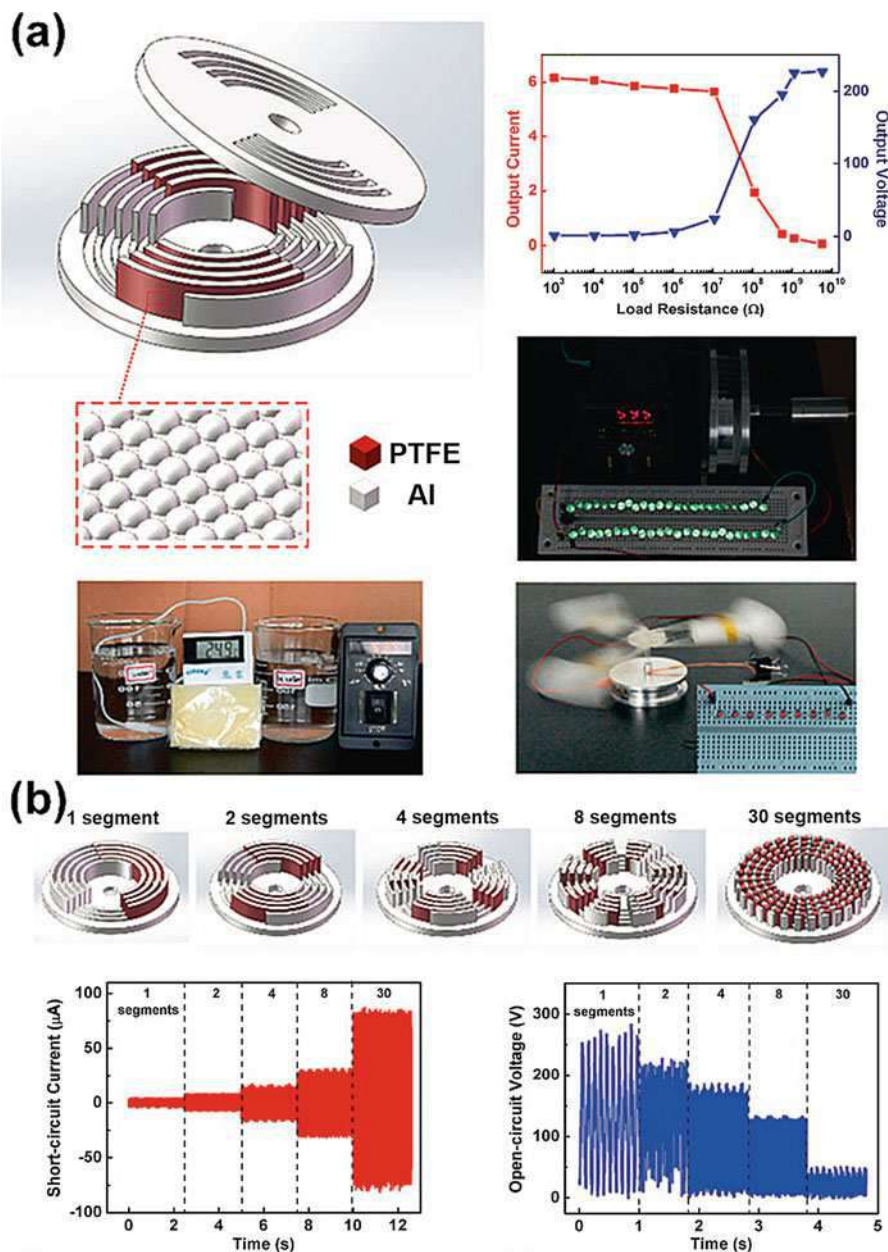
**Fig. 6** Rotary disk based direct-current TENG (DC-TENG). (a) Schematic illustrations showing the proposed working principle of the rotating disk based DC-TENG with the electrons flow diagram in four consecutive stages within a full cycle of direct current output. (b) The measured short-circuit current density with different rotating speeds from 150 to 1350 rpm. (c) The summarized relationship between the short-circuit current density and the rotating speed of the various structures (Reprinted with permission from Zhang et al. 2014a. Copyright 2014 John Wiley & Sons, Inc)

flow in the direction from joint A to B. In the second half cycle, because of the sliding contact and automatic switch between the electric brushes and the two electrodes, the top electrode contacts with the electric brush of joint B, and the bottom electrode contacts with the electric brush of joint A, also forming a current flow in the direction from joint A to B. Therefore, the current is reversed in the second half cycle, and a direct current is generated in the full cycle. The role played by the rotating speed and the segmentation number is thoroughly investigated. The short-circuit current density is measured with different rotating speeds from 150 to 1350 rpm (Fig. 6b), and the output performances are also compared and analyzed with different structures, which have one sector of half circle, two sectors of quarter circle, three sectors of sixth circle, and four sectors of eighth circle, respectively (Fig. 6c). The experimental results show the direct-current enhancement not only at higher speed but also with more segments. Owing to the great advantages of its direct and continuous current output characteristics over previous TENG, multiple light-emitting diodes (LEDs) were lit up directly without a rectifier bridge, and an energy storage unit such as a capacitor was quickly and directly charged. This work presents a novel DC-TENG technology and opens up more potential applications for harvesting rotational mechanical energy and powering portable electronics.

## Multilayered Cylindrical TENG for Enhancing Output Current

Unlike the traditional electromagnetic induction generator, the TENG is equivalent to a current source where the current depends significantly on the induced charge amount and the charge transferring's frequency. Considering that, the cylindrical TENG was optimized in this work by employing multilayer integration and segmentation (Tang et al. 2014b). It was found that when the layer number is four, the segmental number is eight, and the rotating speed is 600 rpm, the TENG's output current reached 30  $\mu\text{A}$ , with the output voltage around 220 V. Multiple light-emitting diodes can be powered up, and a temperature sensor can sustainably work without any other supply (Fig. 7a). More importantly, we found that segmentation is a perfect self-power-managing technique for TENG. Figure 7b shows the measured short-circuit current and open-circuit voltage with different segmentation number. As expected, the  $I_{sc}$  was largely enhanced from 3.7 to 86  $\mu\text{A}$  when the segmentation number increased from 1 to 30. On the other hand, the open-circuit voltage decreased with the increasing segmentation number. The decay of  $V_{oc}$  might be attributing to a significant increase in the side capacitance between the two cylinder petal series, which was predicted in the previous work (Niu et al. 2014b). FEA analysis also shows this descending trend with respect to the increasing segmentation number. As the positive and negative charges distribute more dispersedly, the electric potential difference gets lower. This demonstration implies a self-power-managing triboelectric nanogenerator, which means the segmentation configuration will lower the TENG's voltage and increase its output current without an external transformer.

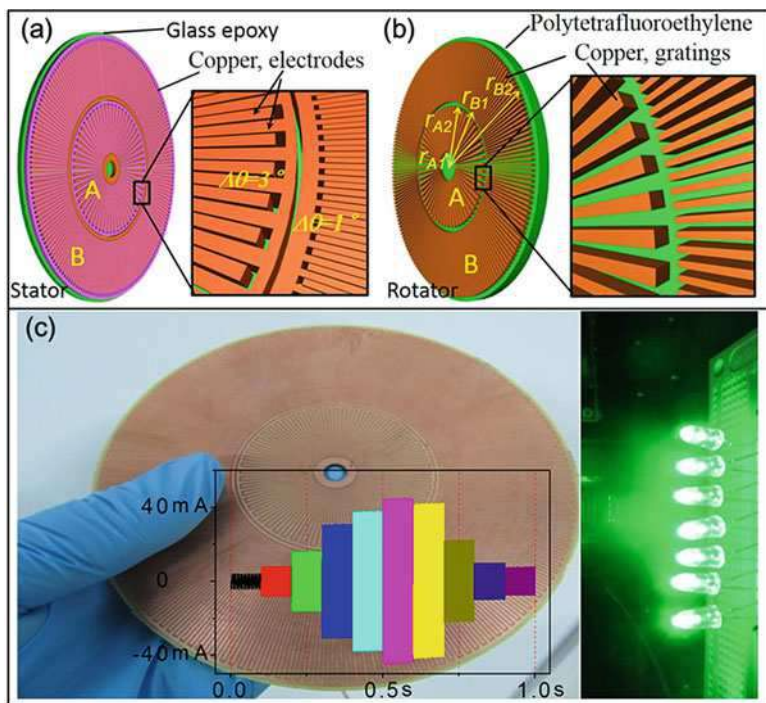




**Fig. 7** Cylindrical TENGs with segmentation design and multilayer integration. **(a)** Basic structure, output characteristics, and applications of the cylindrical TENG with created PTFE nanohemisphere array on the surface area. **(b)** Configurations and characteristics of various segmentation numbers of TENGs (Reprinted with permission from Tang et al. 2014b. Copyright 2014 John Wiley & Sons, Inc)

## Radial Grating-Based High-Performance TENG

The sliding mode, designed with micro-sized grating structure, offers a unique and straightforward solution in harvesting energy from the relative sliding between two surfaces. Here, an industrialization technology of printed circuit board (PCB) was introduced to prepare TENG with composite disk structure (Han et al. 2015a). Two pieces of PCB-based disks were fitted coaxially. One disk is a rotator and the other is a stator. The surface of PCB-based disk has a copper pattern – grating embedded into PCB base material. Two central angles of  $3^\circ$  and  $1^\circ$  for the gratings were integrated on a disk TENG to improve the space utilization and output power of the device (Fig. 8a, b). By means of PCB technology, TENG can be manufactured in quantity with high yield and lifetime. Figure 8c is a photograph of TENG based on PCB technology. When it is operated by a motor at a speed of 600 rpm, several green LEDs were lighted up, and the luminous intensity of single green LED is beyond



**Fig. 8** Radial grating based high-performance TENG. (a, b) Schematic diagram of the TENG with a composite-disk structure: the inner TENG A and the outer TENG B. The center angles of the TENG A and TENG B are  $3^\circ$  and  $1^\circ$ , respectively. (c) Photograph of the TENG based on PCB. The photograph of green LEDs powered by TENG at a rotation rate of 600 rpm. The luminous intensity of single green LED is beyond 1000 mcd. The inset of (c) is the short-circuit current at different frequency (Reprinted with permission from Han et al. 2015a. Copyright 2015 Springer)

1000 mcd, which is equal to several candles and is sufficient to light a room. The AC signal generated directly from TENG can also power some luminescent devices, such as electroluminescent paper, and the brightness is beyond  $40 \text{ cd/m}^2$ , which is near its nominal brightness. To improve the applicability, a commercial transformer was used in the output port of TENG to lower the output voltage. When the device was worked at different frequencies, the different output currents can be obtained, and the maximum current reaches 45 mA at the frequency of  $\sim 2.1 \text{ kHz}$ , which is 15 times of the original output current.

## Contact- and Noncontact-Mode TENG for Harvesting Energy from Automobile Brake

A common characteristic of these modes requires a direct contact of the two materials in order to harvest energy, which may raise the issue such as durability, lifetime, and stability for the TENG. For example, the continuous friction in lateral sliding mode disk TENG will produce a large amount of heat and accelerate the material depletion, thereby restricting its wide applications. Now let us consider a practical situation such as a disk-based braking system in an automobile. During braking, the brake pads are forced to in contact with the rotating disk, but in the non-braking situation for the most of time, the pad may be a bit off the disk. Although the demonstrated TENG can harvest energy from a disk brake system during tight contact, an approach is missing for the case when the pad is off the disk with minimized friction.

The working mechanism of the noncontact-mode TENG is illustrated in Fig. 9a (Han et al. 2014a). At the initial position, we assumed that the PTFE and Al had been in contact so that the two surfaces were charged in negative and positive, respectively, due to triboelectrification process. Then, the two surfaces are separated for a distance  $d$ , and the Al film is right above the PTFE layer (Fig. 9a-i). When the PTFE layer slides in parallel to Al layer for a small angle (Fig. 9a-ii), the free electrons in the Cu electrode beneath the PTFE will flow to the Al layer to balance/screen the nonmobile triboelectric charges in the PTFE layer. This process continues until the two films are completely mismatched (Fig. 9a-iii). When the Al layer continues to rotate and starts to overlap with the next PTFE layer on the other segment of the disk (Fig. 9a-iv), the free electrons in the Al layer will flow back to the corresponding Cu electrode beneath the PTFE layer. This is a cycle of electricity generation. A simulative brake pad TENG was fabricated, and the electric generation mechanism during braking process is illuminated (Fig. 9b). Figure 9c presents the measured  $J_{SC}$  after rectification for multiple of cycles of braking (contact mode) and non-braking (non-contacting mode): from stop status to rotating and then to periodic braking. When the disks rotate at a speed of 900 r/min with  $d = 3 \text{ mm}$ , the induced  $J_{SC}$  reaches  $1.5 \text{ mA/m}^2$ , but a significantly enhanced  $J_{SC}$  up to  $43 \text{ mA/m}^2$  was achieved after braking. The corresponding peak  $V_{OC}$  is also beyond 500 V when braking (shown in the inset of Fig. 9c). The output signals can also be confirmed by the brightness of the red LEDs (Fig. 9d).

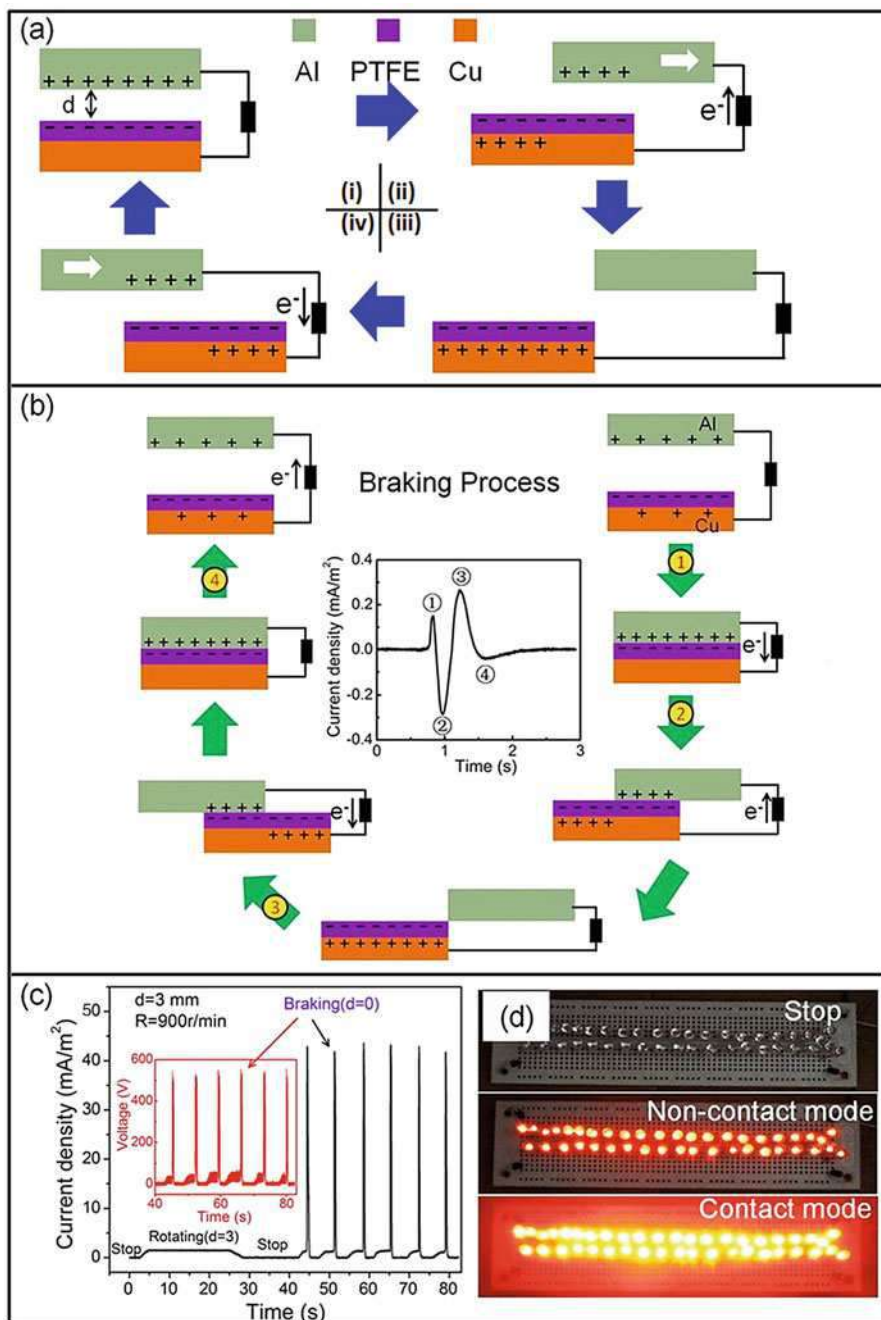


Fig. 9 (continued)

## TENG with Pulsed Output Current for Efficient Charging of Li-ion Batteries

The energy harvested by the TENG from the environment is usually time dependent or climate dependent. Thus, energy storage by batteries or capacitors and associated power management circuits are required in order to obtain constant or controllable power supplies. In this work, a feasible and efficient charging of lithium-ion batteries with pulsed output current generated by a rotating TENG was demonstrated (Pu et al. 2016). The TENG has a 2D planar structure, comprising of a bottom stator and a top rotor, both of which were fabricated based on the printed circuit board (PCB) technology (Fig. 10a). Fast Li-ion extractions of typical electrode materials, i.e.,  $\text{LiFePO}_4$  and  $\text{Li}_4\text{Ti}_5\text{O}_{12}$ , were achieved by the TENG at 250 rpm rotating speed. The estimated coulombic efficiency of the TENG charging and the following 0.5 C discharging can be higher than 90%, comparable with that of constant current charging (Fig. 10b, c).

The power utilization efficiency should be optimized when storing the energy of the TENG by a battery. A transformer was found to be effective to bridge the gap between the battery impedance and matched impedance of the TENG (Fig. 10d). A commercial Li-ion battery ( $\text{LiCoO}_2$  cathode and graphite anode), possessing a capacity of 200 mAh when being galvanostatically cycled at 50 mA, was utilized to store the energy generated by the TENG with the transformer (coil ratio of 36.7). After charging for 1 h, the voltage of the cell increased to  $\approx 4.2$  V, and the subsequent discharging at 50 mA delivered a capacity of 130 mAh, equivalent to 65% of complete charge (Fig. 10e). This capacity is enough to completely drive many small personal electronics, such as smart bracelets and watches. Considering the readily scale-up capability of the TENG, the energy stored by the TENG is promising to power larger-scale personal electronics, like smartphone and pad, which typically require battery capacity of  $\approx 10^3$  mAh.

---

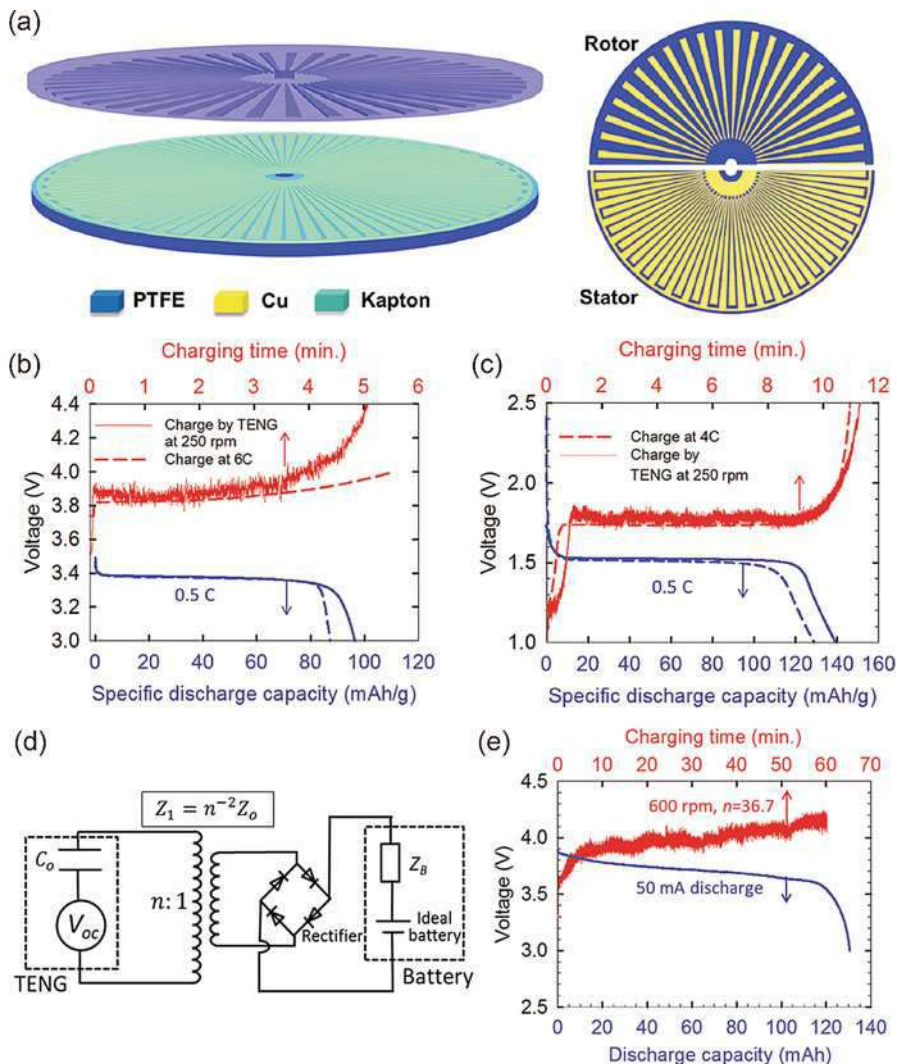
## TENG as Mega-scale Power Source

With a rapid consumption of the world's reserved fossil fuels, seeking new energy sources is vital for the sustainable development of humankind. Simultaneously, the climate change and global warming caused by the overexploitation of fossil fuels



**Fig. 9** Contact and non-contact mode TENGs for harvesting energy from automobile brake. (a) Sketches for the electricity generation process in a full cycle for double-electrode and single-electrode TENG at the non-contact mode. (b) The basic mechanism of electricity generation process in a braking cycle. The inset is the measured  $J_{SC}$  change versus times, and the four current peaks are corresponding to four electricity generation process in (b). (c) The measured  $J_{SC}$  curve at different situation: stop, rotating and braking. The inset is the corresponding  $V_{OC}$  curve. (d) Photographs of a continuous luminescence process for red LEDs when rotating (non-contact mode) and a flashlight when braking (contact mode) (Reprinted with permission from Han et al. 2014a. Copyright 2014 Elsevier Ltd)





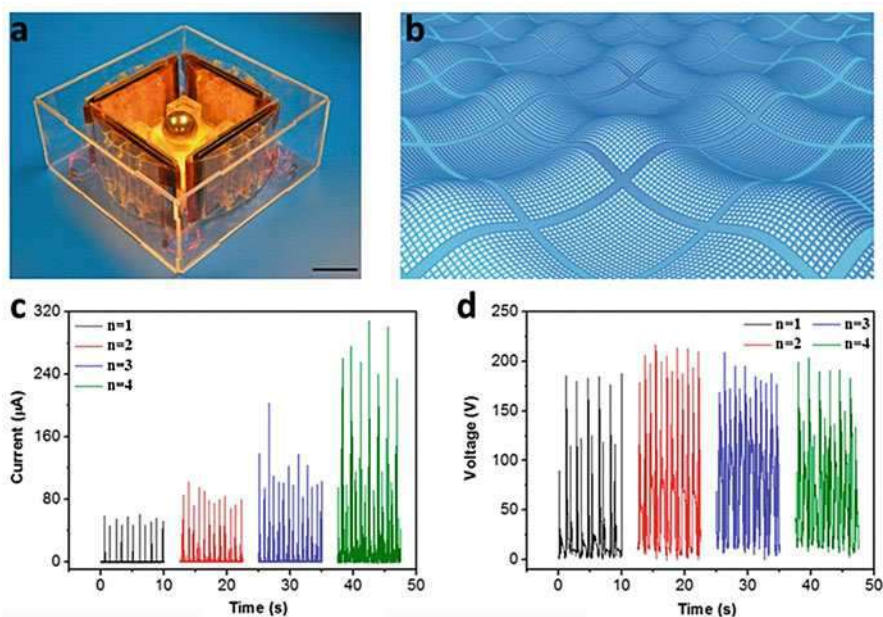
**Fig. 10** Efficient charging of Li-ion batteries with pulsed output current of TENGs. (a) Schematic diagram showing the structural design of the TENG. (b, c) Voltage profiles of LiFePO<sub>4</sub> and Li<sub>4</sub>Ti<sub>5</sub>O<sub>12</sub> half cells charged by the TENG (solid lines) and constant current (dotted lines). (d) The equivalent circuit of battery charging by the TENG with the aid of a transformer and a rectifier. (e) Voltage profiles of a commercial Li-ion battery charged by the TENG (Reprinted with permission from Pu et al. 2016. Copyright 2016 John Wiley & Sons, Inc)

have been the most important environmental concerns (Brown 1999). By examining the world resources, we detected that energy available in ocean is almost unexplored; the water wave energy, for example, has not been utilized owing to bulky instruments, low conversion efficiency, and engineering difficulty and complexity based

on the electromagnetic generators. The TENG is developed as microscale power source, and their assembly and integration can be the basis for harvesting mega-scale energy, which can be a new and efficient approach for harvesting the kinetic energy from water wave as a renewable and sustainable power source.

## Networks of TENGs

A network design made of TENGs is reported for large-scale harvesting of kinetic water energy (Fig. 11a, b) (Chen et al. 2015a). Relying on surface charging effect between the conventional polymers and very thin layer of metal as electrodes for each TENG, the TENG networks (TENG-NW) that naturally float on the water surface convert the slow, random, and high-force oscillatory wave energy into electricity. On the basis of the measured output of a single TENG, the TENG-NW is expected to give an average power output of 1.15 MW from 1 km<sup>2</sup> surface area (Fig. 11c, d). Given the compelling features, such as being lightweight, extremely cost-effective, environmentally friendly, easily implemented, and capable of floating

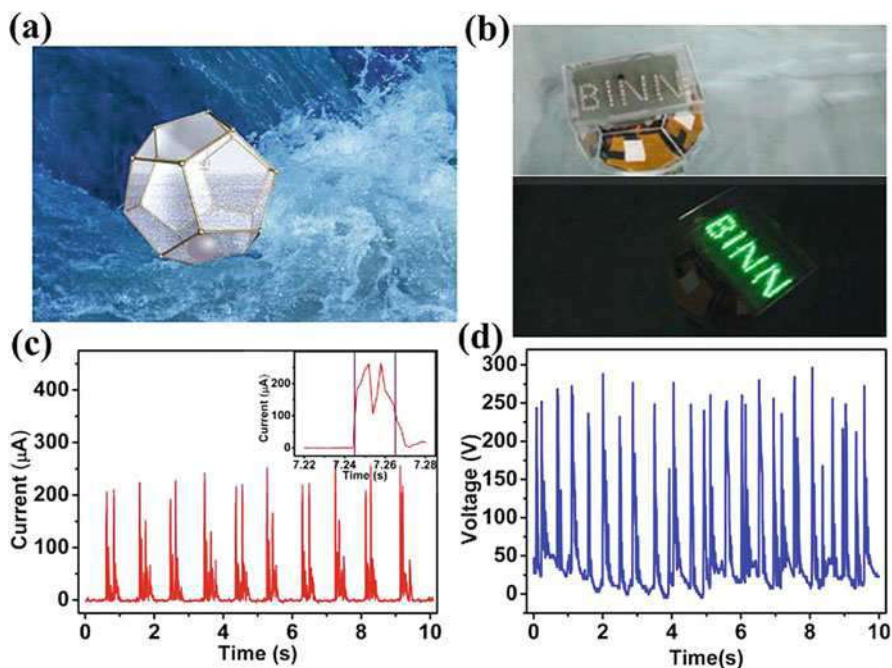


**Fig. 11** TENG network (TENG-NW) and its electrical output characterization for water wave energy harvesting. (a) Photograph of an as fabricated single unit of the TENG-NW. The scale bar is 5 cm. (b) Schematic illustration of the TENG-NW that consists of thousands of single units. (c) Rectified short circuit current of the TENG-NWs with unit number  $n = 1, 2, 3, 4$ . (d) Open circuit voltage of the TENG-NWs with unit number  $n = 1, 2, 3, 4$  (Reprinted with permission from Chen et al. 2015a. Copyright 2015 American Chemical Society)

on the water surface, the TENG-NW renders an innovative and effective approach toward large-scale blue energy harvesting from the ocean.

### Multilayer Wavy-Structured Robust TENG

A regular dodecahedron device integrated with 12 sets of multilayer wavy-structured robust triboelectric nanogenerators (WS-TENGs) is reported for harvesting water wave energy (Fig. 12a, b) (Zhang et al. 2016). Each WS-TENG is composed of a wavy-structured Cu-Kapton-Cu film and two fluorinated ethylene propylene (FEP) thin films sputtered with metal electrodes as a sandwich structure. A hard ball is enclosed inside a polyhedron made by WS-TENGs as the walls; a collision of the ball with the WS-TENG in responding to the kinetic motion of water wave converts mechanical energy into electricity. A high output voltage and current of about 250 V and 150  $\mu\text{A}$ , respectively (Fig. 12c, d), are measured by a single unit of WS-TENGs in water. Considering the units can be connected into a net structure, the average



**Fig. 12** Schematic illustration, photographs and output characteristics of the wavy-structured TENGs (WS-TENGs). (a) The schematic illustration of WS-TENGs in water. The twelve multilayer WS-TENGs are located to each internal surface of the dodecahedron device. The insets is the schematic of the work about A network design made of triboelectric nanogenerators (TENGs) for large-scale harvesting of kinetic water energy. (b) Photographs of the device for powering LEDs by water wave. (c) The short-circuit current and (d) the open-circuit voltage of the device in water (Reprinted with permission from Zhang et al. 2016. Copyright 2016 Elsevier Ltd)



output power is expected to be 0.64 MW from 1 km<sup>2</sup> surface area in a depth of 5 m. By the virtues of cost-effective, low carbon, and environmentally friendly, the development of WS-TENGs can be a significant step toward the large-scale water wave energy harvesting and has the potential to be the dominant generator for blue energy.

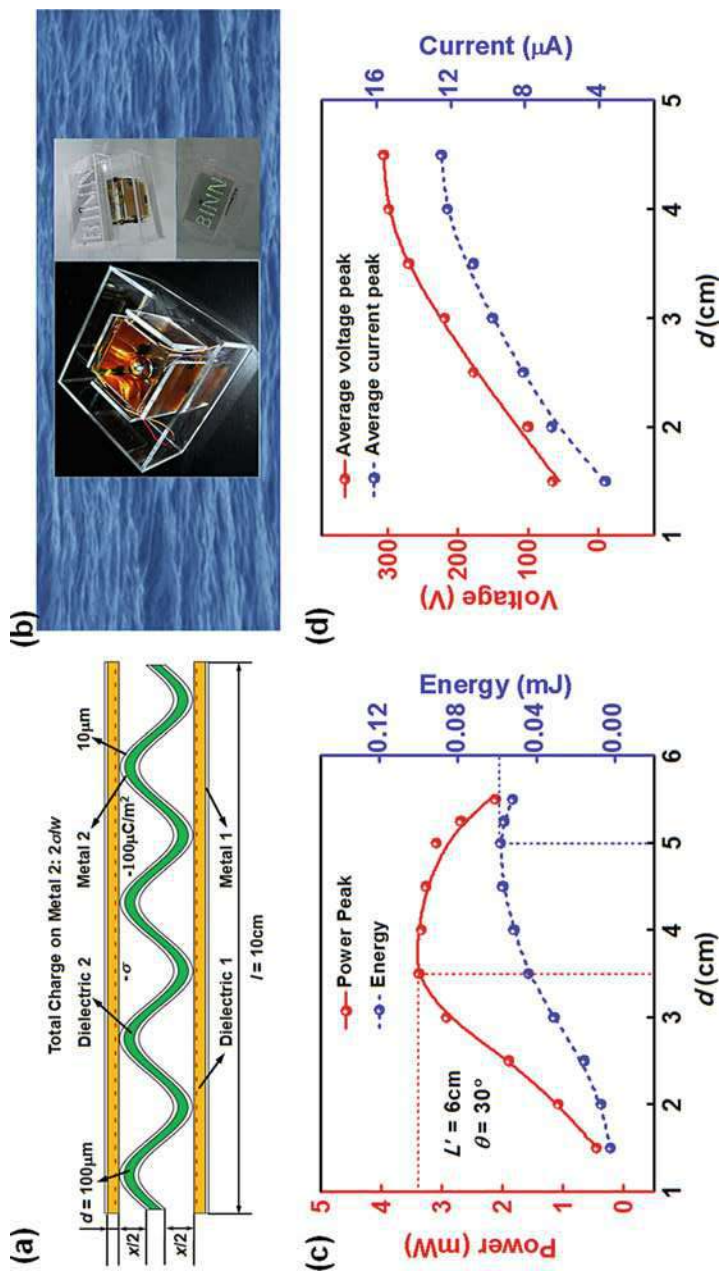
## Structural Optimization of TENG for Harvesting Water Wave Energy

Networks formed by triboelectric nanogenerators have recently been proposed as a promising technology for harvesting water wave energy. Based on the network design, a basic unit has been studied and structurally optimized by combination of the theoretical calculations and experimental studies (Jiang et al. 2015). Here, we provide an example on the approach of structural optimization for the TENG units toward water wave energy harvesting. The finite element model for a wavy-structured TENG is established (Fig. 13a), in which a wavy structure with both sides of middle dielectric layer coated by metal is sandwiched between two flat dielectric layers bonded with electrode. The device with four basic wavy-structured TENG units as standing anchored walls enclosed by a metal ball has been fabricated to harvest the water wave energy (Fig. 13b). From the viewpoint of theory, the influences of ball size on the maximum output power and generated electric energy have been examined for constant ball moving length and sloping angle (Fig. 13c). Under the currently adopted conditions, there exists an optimum ball size for the device to reach the maximum power and energy, which is contributed to the competition between ball mass and allowable moving space. Similarly, the measured output voltage and current indicate that an optimum ball diameter exists to provide the maximized output, verifying the above theoretical results (Fig. 13d).

After the TENG units with the optimized performance are integrated into the TENG network, the output performance of this network will be improved evidently. The TENG network exhibits more obvious advantages over traditional electromagnetic generator, such as small size, lightweight, cost-effective, and high efficiency at low frequency. The TENG technology will play a more important role in the current development of blue energy harvesting.

## Air-Driven Membrane-Structured TENG Array

The TENG shows particular advantages in transforming low-frequency mechanical energy into electricity (Zi et al. 2016) and is likely to be a brand new solution for utilizing water wave energy in the ocean. An integrated TENG array device based on air-driven membrane structures has been developed (Xu et al. 2017). The detailed structure of the device is shown in Fig. 14a, b, as is mainly composed of two parts, the inner oscillator and the outer shell, connected with elastic bands to form a spring-levitated oscillator structure. Such structure would oscillate in response to water waves when the whole device is floating on water (Fig. 14c), even in a resonant state.



**Fig. 13** Theoretical calculation and experimental measurement on the performance optimization of TENG for harvesting water wave energy. **(a)** Finite element model of the way-structured TENG. **(b)** Photograph of an as-fabricated TENG device and optical images of dozens of LEDs driven by the water wave. **(c)** Maximum output power and generated electric energy of the TENG calculated from the theoretical model as functions of the enclosed ball size. The moving length of the ball and sloping angle are kept constant. **(d)** Measured relationship between average voltage peak and current peak values and ball size for the fabricated devices (Reprinted with permission from Jiang et al. 2015. Copyright 2015 ACS publications)

The two deformable air chambers would be pressed alternatively by the shell, causing pressure difference inside the chambers, which would drive the TENG units to work in contact-separation mode by deforming the soft membranes where TENG electrodes are attached to (Fig. 14b). With the design of using air pressure to transfer and distribute harvested water wave energy, the device can drive a series of integrated TENG units effectively and simultaneously.

As a demonstration, a device integrating 38 TENG units shows high output of transferred charges per cycle of 15  $\mu\text{C}$ , short-circuit current of 187  $\mu\text{A}$ , and optimized peak power density of 13.23  $\text{W}/\text{m}^3$  (Fig. 14d). It can easily light up 600 LEDs with harvested water wave energy (Fig. 14e). With excellent scalability of the array, the device can easily integrate large-scale high-density TENG units in one package, as would greatly augment the output. While possessing merits of light weight, low cost, etc., it indicates exciting perspective for using TENG technology as a novel technical route to exploit ocean energy, for powering ocean instruments and supplying electricity to the grid.

---

## TENG as Self-Powered Active Sensors

The TENG can be a power source for sensors as a self-powered sensing system. On the other hand, the TENG is also a device that converts a mechanical trigger into an electric output signal, which means that it can be used to direct sensing a dynamic mechanical action without applying a power unit to the device. Several different self-powered active sensors have been developed as below.

### Self-Powered Wireless Sensing Node

Typically, TENG's output is an AC signal that responses to the frequency at which the mechanical triggering is applied; thus, the output is required to be converted from AC into DC and stored before driving a conventional electronics. More importantly, the output of a TENG has a common characteristic of high voltage but low current and total transported charges. Therefore, it needs transformation before applying to drive conventional electronics. But such a power-transformed-and-managed method is different from the traditional method of using a transformer for sinusoidal type AC signal, because the output of a TENG can be a short pulse at variable frequency. In 2014, a power-transformed-and-managed TENG (PTM-TENG) is developed by integrating a contact-separation-mode TENG with a self-connection-switching capacitors array that are connected in serial when being charged and then in parallel during discharging (Fig. 15) (Tang et al. 2014c). It is found that the PTM-TENG's output voltage can be tunably decreased, and its output current and charges per applied load (impact) are increased. As a comparison, the time for charging a 10  $\mu\text{F}$  capacitance to 5 V by the conventional TENG was 380 s, while by using a PTM-TENG (8-capacitor array), it was only 47 s, which means that the output charges were enhanced by eight times. Moreover, it is reported that conventional

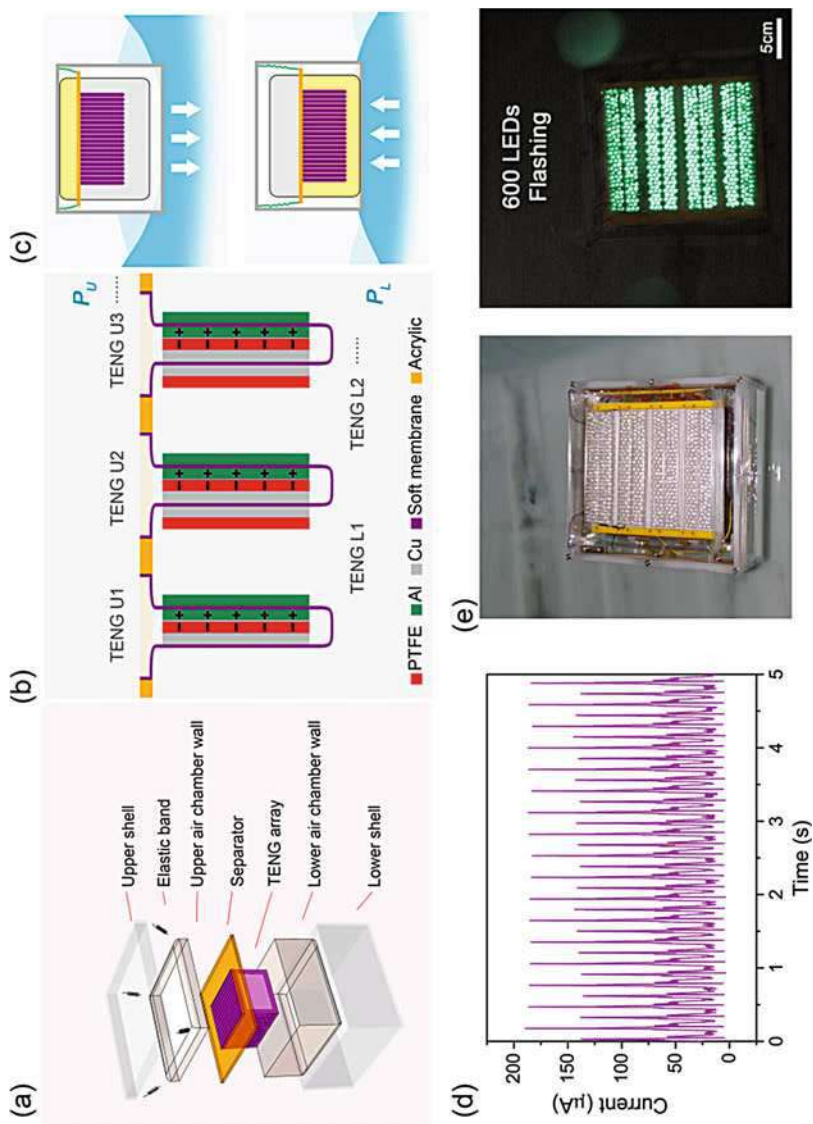
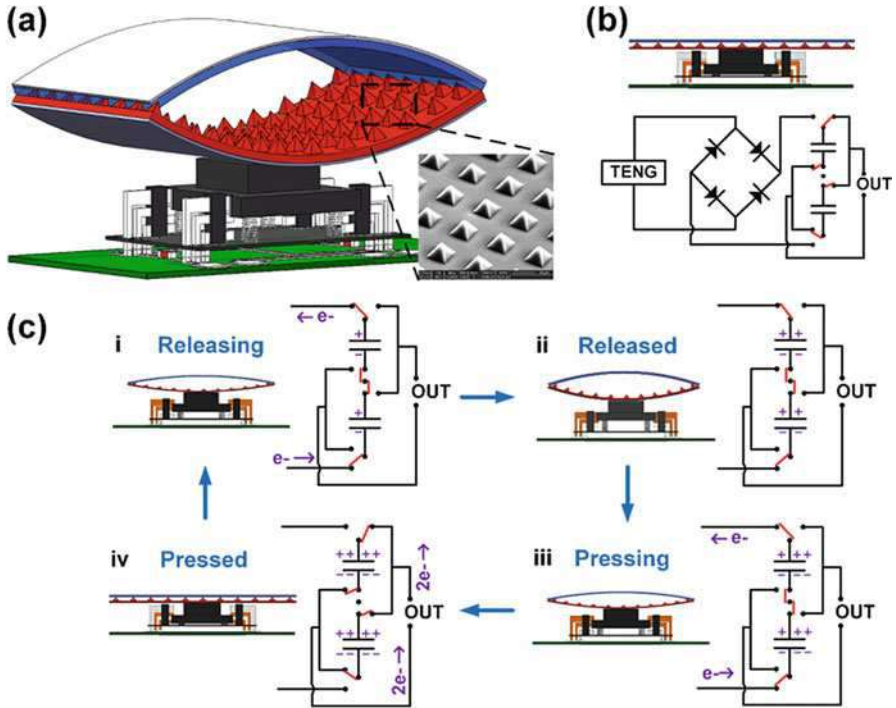


Fig. 14 (continued)



**Fig. 15** Schematic and working principle of the power-managed triboelectric nanogenerator (PTM-TENG). (a) 3D structure of the PTM-TENG. (b) The pressed PTM-TENG’s cross sectional view and the equivalent circuit diagram. (c) A full working cycle of the PTM-TENG (Reprinted with permission from Tang et al. 2014c. Copyright 2014 IOP Publishing Ltd)

TENG’s output voltage at the external load decreases with the lowering of the load resistance and the working frequency (Cheng et al. 2013), which means a decrease of the output energy at a smaller load resistance or a slower impact speed. However, the PTM-TENG shows independence on these two factors. As a result, when the load resistance was 10 kΩ, and the external impact speed was 0.1 m/s, the output energy was enhanced for 2200 times by the PTM-TENG. Furthermore, the PTM-TENG’s charging/discharging mode was changed from continuous to instantaneous by the capacitors array, which can gigantically enhance the instantaneous output current

**Fig. 14** Integrated TENG array based on air-driven membrane structures for water wave energy harvesting. (a) Exploded diagram of the structural design of the TENG array device. (b) Sectional view and detailed structure of the TENG array.  $P_U$  and  $P_L$  represent the pressure in the upper and lower air chambers, respectively. (c) Driving principle of the device by water waves. The inner oscillator would oscillate corresponding to water waves. (d) Rectified short-circuit current of the whole device. (e) Photographs of the device with 600 LEDs floating on water before and after being lighted up (Reprinted with permission from Xu et al. 2017. Copyright 2017 Elsevier Ltd)

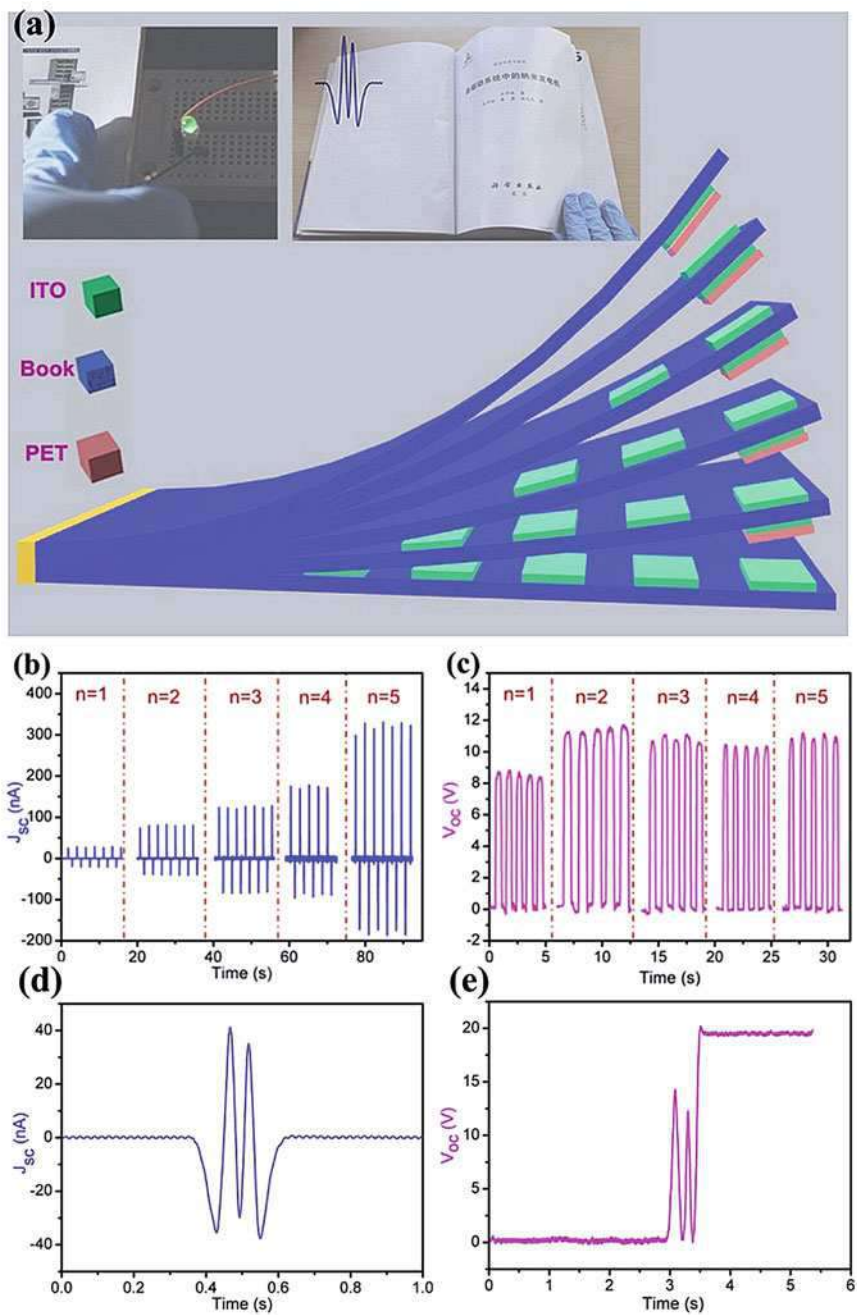


Fig. 16 (continued)



and power. Finally, the PTM-TENG was successfully applied in a wireless touch sensor. Without any power supply, the sensing node will not only detect the touch stimulation but also convert the mechanical energy caused by the stimulation into electric power for the infrared communication.

## Transparent Paper-Based Page Mark and Anti-theft Sensor

Paper has been widely used for thousands of years in human civilization, for the outstanding advantages of lightweight, cheap, flexible, and environment friendly. Over the past decades, paper-based functional electronic devices have opened up a new era of applications in circuits, chips, and sensors. Although the development of integrated technology has enlarged the range of applications of paper-based electronic devices, the fatal weakness for the paper-based systems is too much dependence on external power supply. Hence making the paper-based systems work independently and sustainably has profound significance. In this work, a new type of paper-based triboelectric nanogenerator (PTENG) is demonstrated using indium tin oxide (ITO) film and polyethylene terephthalate (PET) film as triboelectric surfaces (Zhang et al. 2014b). It can be made in a book or any other paper products to harvest kinetic energy when the papers contact, bend, or relatively slide with the advantages of both transparent and flexible. In addition, grating-structured PTENGs are also integrated into a book as a self-powered anti-theft sensor (Fig. 16a). To demonstrate the functionality of the fabricated sensor, the action of stealing the book is simulated. When a book equipped with our self-powered anti-theft sensor is touched, it will work in contact-separation mode with multilayer structure. With the number of layers increased, the current output would continue to improve (Fig. 16b), but the output voltage is almost kept in constant in the whole process (Fig. 16c). The mechanical triggering agitated during handling the book pages can be effectively converted into an electric output to either drive a commercial electronic device or trigger a warning buzzer. Furthermore, different grating structures on each page produce different number of output peaks, which can accurately position the turned pages and record the pages flipped over. When one grating structure on the even page slides past the multi-grating structure on the odd page, the measured output signals were depicted by two kinds of features, which is short-circuit current (Fig. 16d) and open-circuit voltage (Fig. 16e). This work provides a potential method to develop the self-powered, durable, cost-effective, anti-theft system for books, paintings, and any other flexible materials in the future.



**Fig. 16** Multilayer grating structure of the sensor integrated in a book and its function. (a) Multilayer grating structure of the sensor integrated in a book, the *insets* are the sensor driving a LED and positioning the book. (b) The short-circuit currents with different layers. (c) The open-circuit voltage with different layers. (d) The JSC corresponds to the grating number of 3. (e) The  $V_{OC}$  corresponds to the grating number of 3 (Reprinted with permission from Zhang et al. 2014b. Copyright 2014 Springer)

## Self-Powered Velocity and Trajectory Tracking Sensor Array

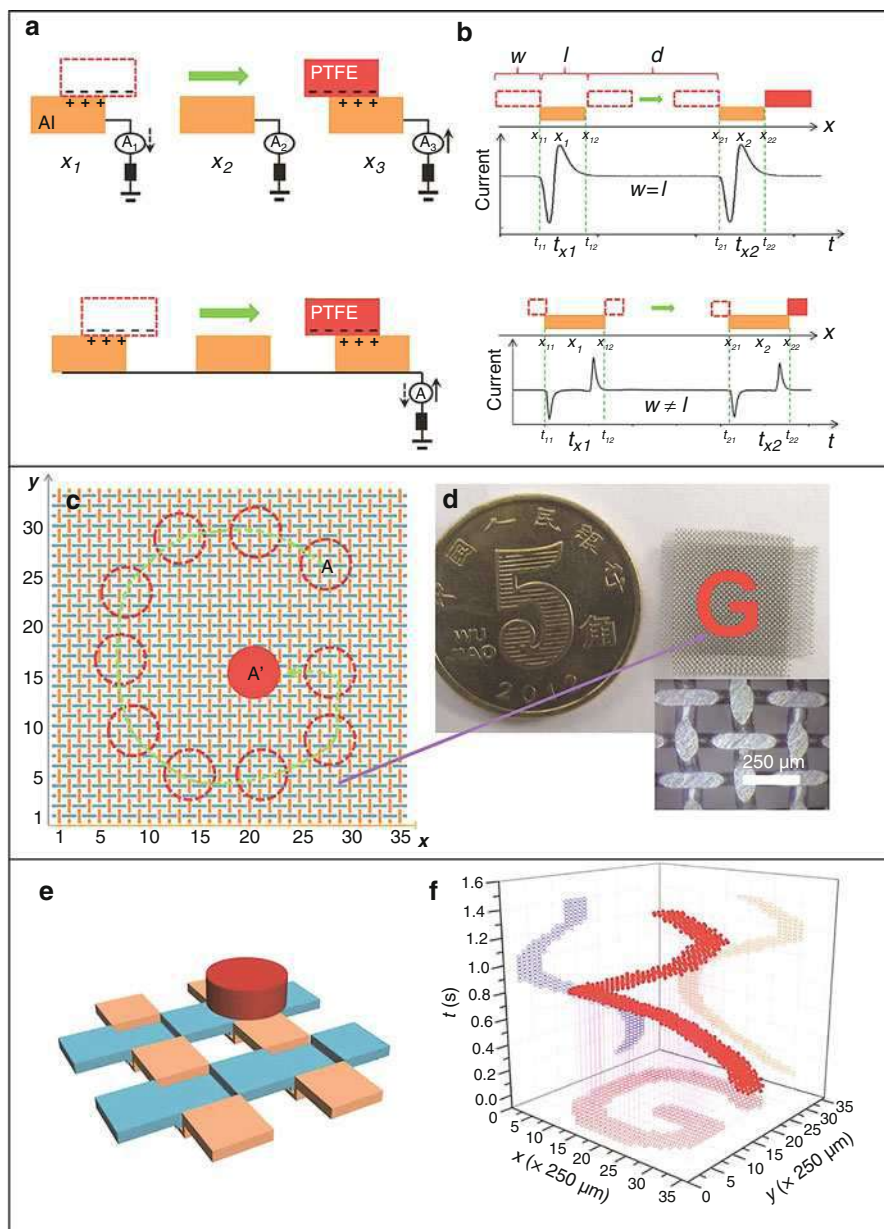
Displacement tracking sensor is frequently used in today's smartphone and human tracking systems. However, a general design for this type of sensors is based on capacitive, optical, or magnetic effects; a common characteristic of all these sensors is that they do need an external power source. Here, a 2D velocity and trajectory tracking sensor (VTTS) was fabricated by an ingenious electrode design based on an array of single-electrode TENGs. When an object moves on a 2D plane, it can be regarded as a plane rectangular coordinate system and whatever it locates in, as characterized by corresponding coordinates  $(x, y)$ . Conversely, provided that we had detected the coordinates of the object in real time, its velocity and trajectory were able to be obtained by X-Y data at the corresponding time. Based on the single-electrode TENG, the working principle of the 1D VTTS for location tracking is shown in Fig. 17a, b (Han et al. 2014b). The generation of electric signals from every channel is based on the charge transfer between the metal strips or wires and the ground when an electrostatically charged object slides on the local surface of the device.

Furthermore, knitting structure VTTS is sketched in Fig. 17c. Here, 41  $x$  output ports and 41  $y$  output ports were vertically interweaved forming a  $41 \times 41 = 1681$  pixels on the grid structure ( $1 \text{ cm} \times 1 \text{ cm}$ ). The distance between the electrodes is  $250 \mu\text{m}$ , so the resolution of movement is  $250 \mu\text{m}$  (inset of Fig. 17d). An object (PTFE, with nanoscale on surface) with a diameter of  $\sim 1.3 \text{ mm}$  (effective value) was introduced to slide along "G" curve on the surface of VTTS. The optical graph of the device is shown in Fig. 17d, and the inset is the high magnification graph. When the object slides across one of the electrode with the speed of  $0.4 \text{ m/s}$ , the peak of short current reaches  $\sim 5 \text{ nA}$ , and the SNR (signal-to-noise ratio) is about 50, which is sufficient enough for a sensor. A 3D structure and experimental curve recording the motion trajectory of target object is shown in Fig. 17e, f, in which all of the data were tested only by 82 output ports. The  $x-t$  and  $y-t$  curves reveal the 2D displacement component changing with time, while the  $x-y$  curves afford the real-time trajectory of moving object. In the same way, the average velocity, instantaneous velocity, and acceleration at any stage could be calculated. Besides, the approximate dimensions of the object is about  $1.3 \text{ mm}$ , which can also be clearly found according to the width of trajectory line displaying in  $x-y$  coordinates from Fig. 17f, and it is consistent with the experimental data.

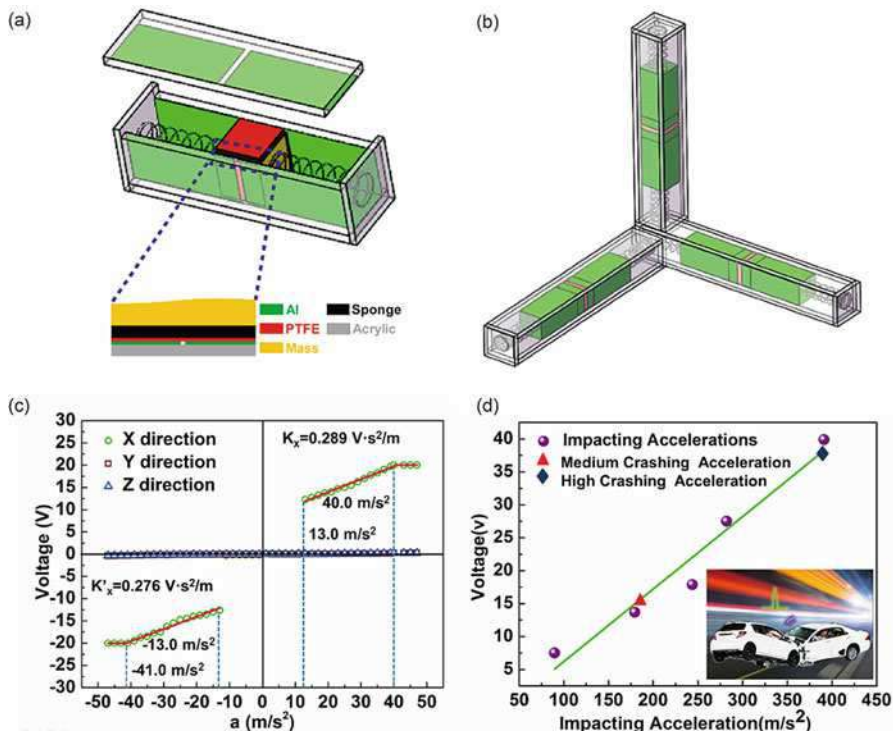
## Self-Powered Vehicle Collision Acceleration Sensor

A novel self-powered acceleration sensor based on triboelectric nanogenerator is proposed, which consists of an outer transparent shell and an inner mass-spring-damper mechanical system (Fig. 18a) (Pang et al. 2015). When the acceleration sensor is subject to a certain external acceleration, the PTFE film will slide against with the Al electrodes owing to the inertia force, and there will be a potential difference between the two electrodes. Owing to the linearly proportional relationship between the acceleration and the output voltage, the sensor can be used to





**Fig. 17** Self-powered velocity and trajectory tracking sensor array. (a) Schematic diagrams of electricity-generation process in the 1D direction for different types of electrode arrangements: individually grounded, and commonly grounded. (b) Relationship among the movement, time and measured current signal of the sliders with different sizes in comparison to the width of the electrode, and the expected output signal shape. (c) Schematic diagrams of 2D VTTS with crossed node and the trajectory of object. (d) Photo graph of VTTS with  $41 \times 41$  pixels in size of  $1 \times 1$  cm. (e) 3D structure of sensor with low node; the inset is the 3D structure of a node. (f) The measured 3D trajectory curves for the object (Reprinted with permission from Han et al. 2014b. Copyright 2014 Elsevier Ltd)



**Fig. 18** Self-powered vehicle collision acceleration sensor. (a) Schematic illustrations of 1D acceleration sensor based on single TENG. (b) Schematic diagram showing the structural design of the device composed of three TENGs. (c) Characteristic calibrations of the 3D acceleration sensor in one direction. (d) Relationship between the output voltage and the impact acceleration. With the speed of 80 km/h, the car collision acceleration is about 180 m/s<sup>2</sup>. When the car collision happened with a high speed of 120 km/h, the car collision acceleration can reach to about 400 m/s<sup>2</sup>. The general and high car collision accelerations are in the measurement range of the acceleration sensor (Reprinted with permission from Pang et al. 2015. Copyright 2015 ACS publications)

measure acceleration with great detection range and sensitivity. By coupling three independent and perpendicular TENGs in the three directions, a self-powered 3D acceleration sensor was developed (Fig. 18b). Due to the special structure design, when the 3D acceleration sensor is subjected to a certain external acceleration in one axis, the 3D acceleration sensor can generate voltage to characterize acceleration in this axis, and there is no output voltage signal for the other two axes (Fig. 18c). The three output voltages of the acceleration sensor can characterize the components of the vector acceleration in the three axes, respectively and independently. The 3D acceleration sensor has good performance with the detection range from about 13.0 to 40.0 m/s<sup>2</sup> and sensitivity of 0.289 V·s<sup>2</sup>/m in the axis direction. Furthermore,

high-speed car collision acceleration was simulated and successfully detected by the self-powered 3D acceleration sensor (Fig. 18d). If two 1200 kg car frontal collision happened, the maximum impact acceleration can reach about 180 and 400  $\text{m/s}^2$  corresponding to the running speed of 80 and 120 km/h, which are in the measurement range of the 3D acceleration sensor.

Unlike the traditional MEMS acceleration sensor (Hindrichsen et al. 2010), this 3D acceleration sensor can work without external powered and be easy fabricated. This study not only demonstrates a self-powered and high sensitivity 3D acceleration sensor but also expands TENG's application in automobile security system.

### Self-Powered Crack Detector and Rotating Speed Sensor

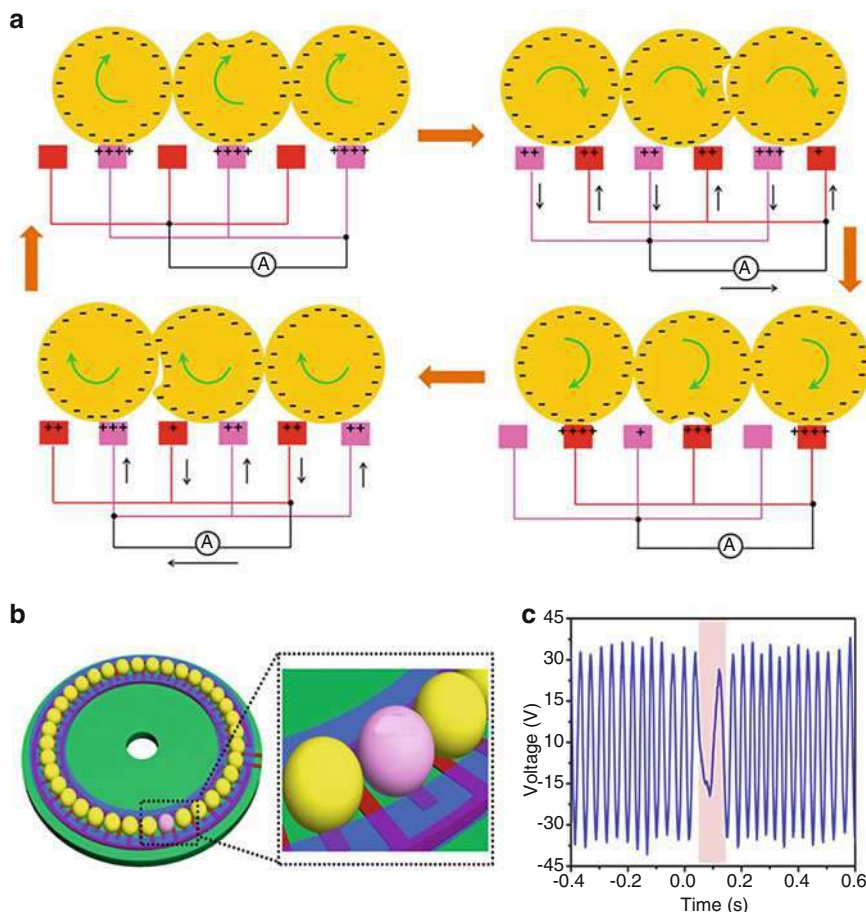
Nondestructive detection, which is able to inspect certain flaws by the means of sound, light, magnetic, and electrical properties without damaging the tested object, is an essential technology in the industrials applications. For bearing defects, nondestructive detection on the basis of electric effects is the common technique for the inspection of bearing failures. However, this method could have possible application limits for its sophisticated testing process, as well as dependence on external source. Here, a bearing-structure-based triboelectric nanogenerator (B-TENG) with interdigital electrodes was developed that cannot only collect energy from rotational kinetic energy but also serve as a self-powered and multifunctional sensor (Li et al. 2016).

As a crucial component in the real machine, the damage of those PTFE-bearing balls may bring a negative effect on the integral operation of the instruments. Here, we propose a self-powered nondestructive detection sensor for monitoring the balls' damage based on the B-TENG. The working principle is explained in Fig. 19a. When rolling on the disk, the PTFE ball with certain damage induces positive charges unevenly distributing on the Cu electrodes, which finally causes the reduction of electric potential difference between the interdigital electrodes, and the reduction tends to increase as the damage gets worse due to a more disordered charge distribution. As depicted in Fig. 19b, sketches demonstrate the conditions that all the balls are complete, one of the balls has a slight damage, and one of the balls has a serious damage, with their corresponding open-circuit voltage signals shown in Fig. 19c. Therefore, the damage of bearing balls can be well displayed by B-TENG.

---

### TENG for Tribo-Controlled Devices

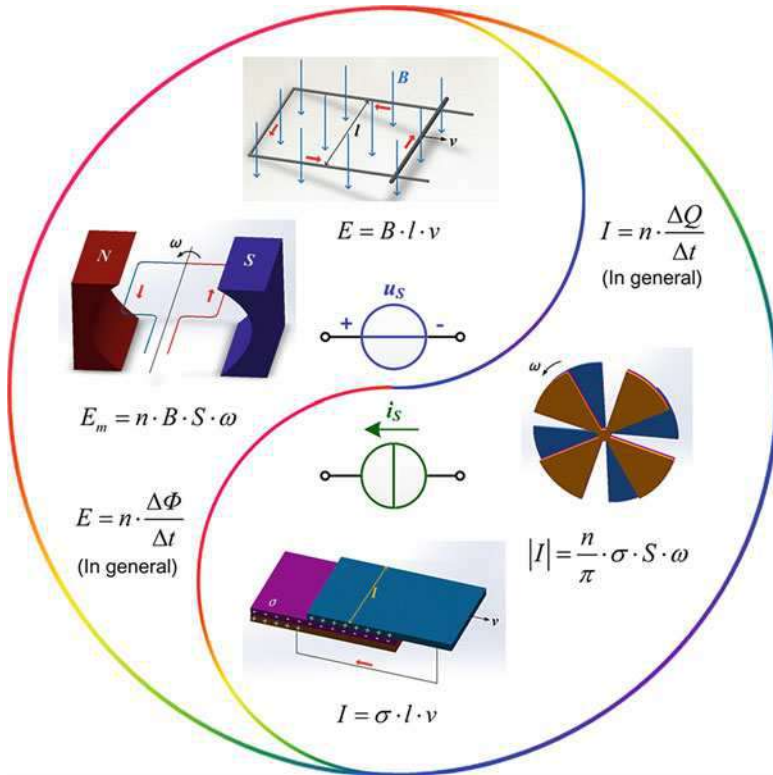
Besides the three applications above, the TENG can also be a role in place of the conventional voltage supply for capacitive devices, which has taken the advantage of high voltage and could be very useful and interesting in conjunction with micro-mechanics and microelectronics.



**Fig. 19** The damage detection of the bearing-structure based triboelectric nanogenerator (B-TENG). (a) Schematic diagrams of the working principles of the B-TENG when the PTFE balls have certain damage. (b) PTFE balls with a serious damage. (c) The output voltage of the B-TENG without any damage (Reprinted with permission from Li et al. 2016. Copyright 2016 IOP Publishing Ltd)

## Comparison of EMIG and TENG

Mechanical energy extensively exists in our living environment and industrial production (Tian et al. 2007). Since the electromagnetic induction effect was discovered by Michael Faraday in 1831, the electromagnetic induction generator (EMIG) has been the most important way for power generation from mechanical energy in conjunction to turbine engine (Williams et al. 2001). There has not been the second possibly equivalently technology that is as important as the EMIG for harvesting mechanical energy until the TENG was invented.



**Fig. 20** Comparison of electromagnetic induction generator (EMIG) and TENG in schematic diagram, fundamental principles and governing equations. The EMIG is close to an ideal voltage source while the TENG can be approximately considered as an ideal current source (Reprinted with permission from Zhang et al. 2014c. Copyright 2014 John Wiley & Sons, Inc)

By systematically analyzing the comparison of EMIG and TENG from their working mechanisms, governing equations, and output characteristics, the comparability and symmetry of the EMIG and TENG have been demonstrated (Zhang et al. 2014c; Fan et al. 2014) (Fig. 20). In the circuits, the TENG can be equivalent to a current source with a large internal resistance, while the EMIG is equivalent to a voltage source with a small internal resistance. For the EMIG connected in series with the load resistance and the TENG in parallel connection with the resistor, the equivalent transformation between them has been established to further validate the characteristics of the TENG as a current source and provide the experimental basis for the applications in conjunction with other devices.

The comparison of EMIG and TENG in internal impedance, impedance characteristic open-circuit voltage, and short-circuit current is summarized in Table 1. The internal construction of the EMIG is a group of coils, which is resistive and inductive with small impedance. The internal construction of the TENG is a dielectric layer, which is capacitive with large impedance. In comparison, the open-circuit voltage of



**Table 1** Comparison of EMIG and TENG in internal impedance, impedance characteristics open-circuit voltage, and short-circuit current

	EMIG	TENG
Approximate equivalence	Voltage source	Current source
Internal impedance	Small	Large
Impedance characteristics	Resistance, inductance	Capacitance
Open-circuit voltage	Low	High
Short-circuit current	High	Low

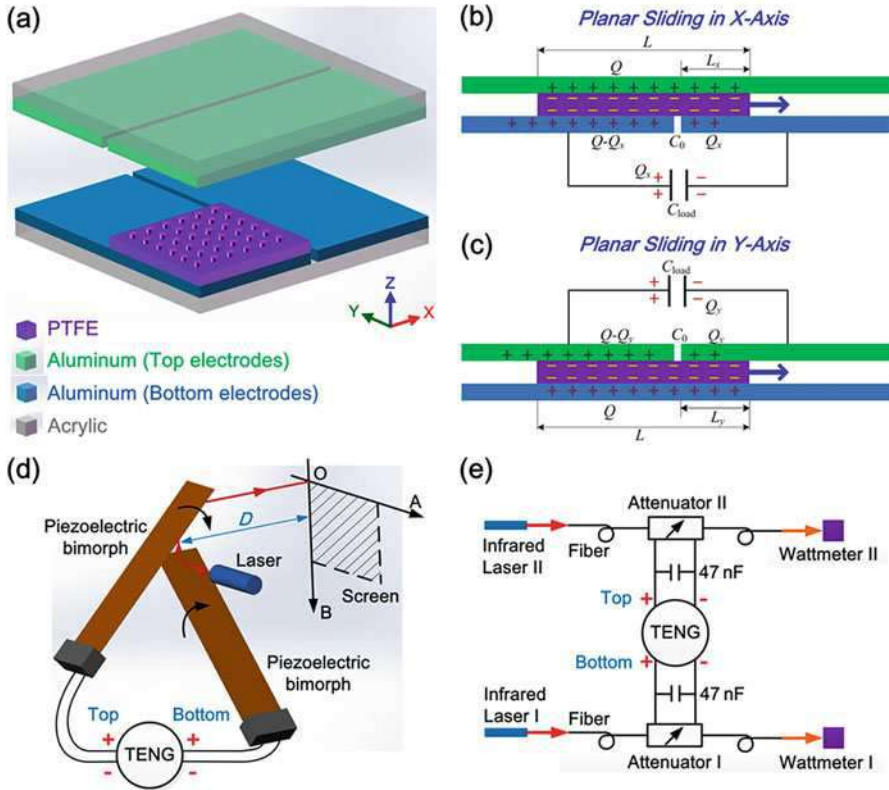
the EMIG is low and the short-circuit current is high, while the open-circuit voltage of the TENG is high and the short-circuit current is low.

In the past years since the TENG was invented, many efforts have been made to enhance the output current. Although great progress has been made, each step is not easy (Zhu et al. 2013b; Wang et al. 2014b). However, the open-circuit voltage is very easy to be high, which means the output voltage can be high when the TENG is applied on a load with large impedance such as a small capacitance. This feature makes the TENG can be a role of high voltage supply in specific conditions (Han et al. 2015b). Therefore, it could open up a new application of TENG as a control source using the output voltage to drive/modulate a capacitive device in place of the conventional voltage supply. Different from the TENG as a power source that relies on the current rather than the voltage, the control source has adopted the advantage point and avoided the shortcoming.

## TENG as a Control Source for MEMS Actuators

With the advantage of high output voltage, the TENG has been used as a control source for capacitive micromechanical and microelectromechanical system (MEMS) devices (Zhang et al. 2015). A planar sliding TENG has been first designed, which consists of a freestanding polytetrafluoroethylene (PTFE) triboelectric layer and two pairs of orthogonal electrodes (Fig. 21a). When the PTFE layer slides in plane, the dual-output voltages on a load capacitance are proportional to the displacements in X and Y directions, respectively, and independently (Fig. 21b, c).

The experimental results show that when the load capacitance is smaller, the larger voltage is applied. Based on the planar sliding TENG and two orthogonal positioned piezoelectric bimorphs, an active piezoelectric micro-actuator has been developed for two-dimensional (2D) optical direction modulation (Fig. 21d). The 2D moving distances of the light spot on the screen are linear to the sliding displacements  $L_x$  and  $L_y$  of the PTFE layer, respectively. While based on the planar sliding TENG and two MEMS infrared optical attenuators, an active electrostatic micro-actuator has been developed for dual-channel optical power modulation (Fig. 21e). The attenuation in each channel could be adjusted independently by the sliding displacement in the X and Y direction, respectively. As the piezoelectric and electrostatic micro-actuators are both small capacitive loads, the planar sliding of the



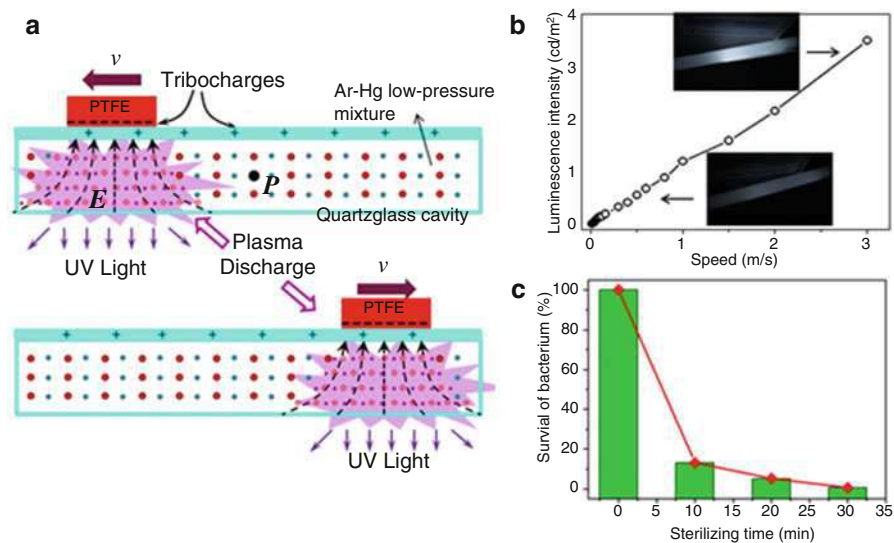
**Fig. 21** Active MEMS for optical modulation based on a planar sliding triboelectric nano-generator. (a) Schematic diagram of the planar sliding TENG with two pairs of orthogonal electrodes. (b, c) Charge distribution in the top and bottom electrodes and load capacitance when the PTFE layer slides in X and Y directions, respectively. (d) Active piezoelectric bimorph for 2D beam direction modulation. (e) Active electrostatic attenuator for double-channel optical power modulation (Reprinted with permission from Zhang et al. 2015. Copyright 2014 John Wiley & Sons, Inc)

PTFE layer has the same effects as supplying conventional high voltages to them. It is worth noting that the applied voltages are dependent on the sliding displacements of the PTFE layer, rather than the sliding speeds, which means that the active micro-actuators are the real-time state-controlled devices by the TENG.

This work has first presented the active micro-actuators driven by mechanical energy without external power or mechanical joints. The micro-/nano-motion in mechanics can be controlled by macroscopic motion, which could open up the important potential applications of independent and sustainable self-powered MEMS/NEMS. More importantly, the TENG has demonstrated great capabilities for modulating the capacitive micromechanics and found a new application for triboelectric-voltage-controlled devices.

## Triboelectrification-Induced Ultraviolet Emission from Plasmon Discharge

UV is a high-energy electromagnetic radiation that has been widely used in industrial production and the scientific research domain. In this work, a deep UV light emission was obtained using triboelectrification-induced plasma discharge without any extra power supply (Han et al. 2015c). By a mechanical friction between polymer and quartz glass, the triboelectric charges cause a changing high voltage electric field, which may bring plasma discharge of low-pressure gas (Ar-Hg) and give out 253.7 nm irradiation. The UV light caused by continuous friction can excite a trichromatic phosphor and afford a bright white light emission. A UV sterilization experiment shows that ~98% of *Escherichia coli* can be killed in 30 min by UV irradiation, which reveals that a self-powered sterilization apparatus with good sterilization effect was fabricated. This work provides a novel design to fabricate a self-powered UV light-emitting device using low-frequency mechanical friction and realizes the coupling of triboelectrification and plasma luminescence, which may further expand the application of UV light in special circumstances (Fig. 22).



**Fig. 22** Triboelectrification induced ultraviolet emission from plasmon discharge. **(a)** The principle of the UV light emitting device based on triboelectrification and plasma discharge luminescence. **(b)** The luminescence intensity changes at different sliding speeds. The *insets* are photographs of the luminescence cavity at different speeds. **(c)** Sterilization effects at different treatment times (Reprinted with permission from Han et al. 2015c. Copyright 2015 Springer)



## Triboelectrification-Induced Trace Memorization Device

By coupling TENG and a thin film of poly(vinylidene fluoride-trifluoroethylene) (P(VDF-TrFE)) polymer, we have demonstrated a novel self-powered memory system to memorize the motion in both 1D and 2D space (Chen et al. 2015b). The polymer (P(VDF-TrFE)) is a kind of ferroelectric materials that has been widely studied for organic memory devices, where two different orientations of dipole moments (up and down) can be considered as “0” and “1” signal for computation and information storage. The spontaneous polarization of the P(VDF-TrFE) is very sensitive to the applied electric field and less dependent on the conduction current, which is very suitable for the TENG’s output characteristic of high voltage and low current. Meanwhile, the dipole switching time of such materials is less than a millisecond, which is also compatible for the motion of TENG. All these facts inspired us to couple P(VDF-TrFE) thin film and TENG for realizing an effective and self-powered memory system to record mechanical displacement or electrical signals. The output voltage produced by the TENG during motion can polarize the dipole moments in the P(VDF-TrFE) film. Later, by applying a displacement current measurement to detect the polarization density in the ferroelectric film, we can directly read the motion information of the TENG. Currently, the ferroelectric thin film with a size of 3.1 mm<sup>2</sup> can record a minimum area changing of 30 mm<sup>2</sup>, and such resolution can still be possibly improved. The good stability and repeatability of the P(VDF-TrFE) thin film guaranteed the promising application of such memory technique for monitoring various mechanical motion. Furthermore, the flexibility of the P(VDF-TrFE) polymer can also open some possible application in the field of flexible electronics. Together with the self-powered capability of TENG, this memory system could potentially be used as a self-powered memory disk for electronic skin or some other soft electronics devices (Fig. 23).

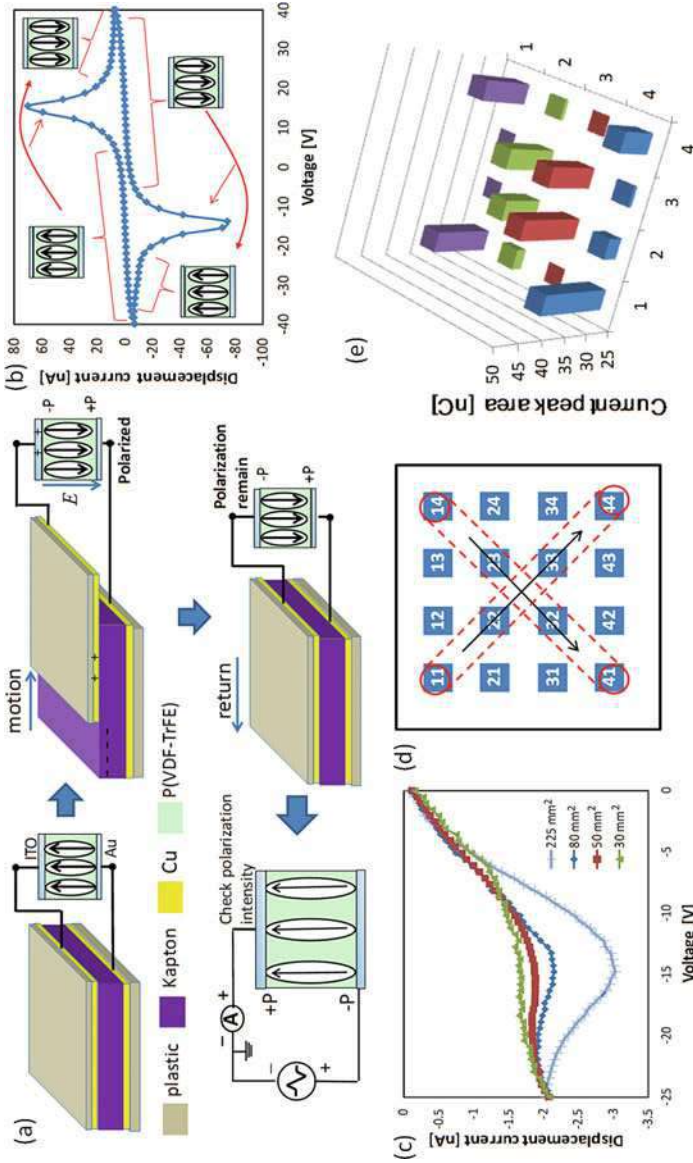
---

## Summary and Perspectives

Since the TENG was invented in 2012, four fundamental principle modes have been illustrated, including vertical contact-separation mode, lateral sliding mode, single-electrode mode, and freestanding triboelectric-layer mode. Based on these modes, various TENGs with different structures are developed depending on specific applications.

The first and direct goal of TENGs is to power micro- and small electronics. By developing various prototypes such as rotary disk, radial grating, fabrics, and multilayered films, the TENG can be used as a microscale power source by harvesting multiform ambient mechanical energy, including hand pressing, body motion, human walking, machine vibration, and wheel moving.

The TENG has also been used for harvesting energy from ocean waves, which has demonstrated enormous potential for general power application at large scale. By constructing a network of the TENGs, an average output power of more than 1 MW from 1 km<sup>2</sup> surface area can be expected for harvesting wave energy. This innovative



**Fig. 23** Triboelectricity induced trace memorization device. (a) Sketch of the device structure and four steps' working principle of the memory system based on sliding TENG and the P(VDF-TrFE) sample. (b) The DCM results of the P(VDF-TrFE)/MIM sample with the ramp voltage from  $-25$  to  $25$  V (double circulation). (c) The DCM results of P(VDF-TrFE) sample connected to electrodes with different size. (d) Illustration of the Kapton moving along a path of letter "X." (e) The calculated current peak area for each P(VDF-TrFE) sample in response to the motion trajectory of letter "X" (Reprinted with permission from Chen et al. 2015b. Copyright 2014 John Wiley & Sons, Inc)

and effective approach possibly prompts the TENG to the mega-scale blue energy harvesting from the ocean in the near future.

As the TENG converts a mechanical trigger into an electric signal, it can be used as a self-powered active sensor for sensing a dynamic mechanical action without power supply. The TENGs have been applied as various active sensors, such as self-powered wireless sensing node, transparent paper-based page mark and anti-theft sensor, self-powered velocity and trajectory tracking sensor array, and self-powered vehicle collision acceleration sensor.

Moreover, the TENG has demonstrated the symmetry and complementary with the traditional EMIG in mechanisms and characteristics, which has great advantages of high output voltage for capacitive devices, such as micro-actuators, ultraviolet emitter, and memorization device. The TENG can have a role in place of the conventional voltage supply for capacitive devices, which could be very useful and interesting for human-machine interfacing.

Up to now, the TENG's instantaneous energy efficiency has reached 70% (Tang et al. 2015), and its total energy conversion efficiency has reached 85% (Xie et al. 2014). The TENG is being studied actively by over 50 research groups around the world. Looking into the future, the TENGs have great applications in various kinds of mechanical energy harvesters, self-powered active sensors, and human-machine interfacing. It is likely to be a feasible and most effective choice for building self-powered systems to drive mobile electronics, sensor networks, flexible electronics, and wearable electronics. With the introduction of new materials and surface engineering technologies, the efficiency of the TENG and overall system performance are expected to be significantly improved. We anticipate its outstanding impact to health monitoring, medical care, environmental protection, infrastructure monitoring, and security.

---

## Acknowledgments

The authors thank the support of National Natural Science Foundation of China (Nos. 51,475,099 and 51,432,005); Beijing Natural Science Foundation (No. 4163077); the “thousands talents” program for the pioneer researcher and his innovation team, China; and the Youth Innovation Promotion Association, CAS, and National Key Research and Development Program of China (2016YFA0202704). We sincerely thank the following collaborators who made significant contributions to the work presented here: (not in particular order): Wei Tang, Changbao Han, Tao Jiang, Liang Xu, Xiangyu Chen, Limin Zhang, Yaokun Pang, Tao Zhou, Xiaohui Li, and many others.

---

## References

- Beeby SP, Torah RN, Tudor MJ, Glynn-Jones P, O'Donnell T, Saha CR, Roy S (2007) A micro electromagnetic generator for vibration energy harvesting. *J Micromech Microeng* 17:1257
- Brown KS (1999) Bright future—or brief flare—for renewable energy? *Science* 285:678–680

- Chen J, Yang J, Li ZL, Fan X, Zi YL, Jing QS, Guo HY, Wen Z, Pradel KC, Niu SM, Wang ZL (2015a) Networks of triboelectric nanogenerators for harvesting water wave energy: a potential approach toward blue energy. *ACS Nano* 9:3324–3331
- Chen X, Iwamoto M, Shi Z, Zhang L, Wang ZL (2015b) Self-powered trace memorization by conjunction of contact-electrification and ferroelectricity. *Adv Funct Mater* 25:739–747
- Cheng G, Lin Z, Lin L, Du Z, Wang ZL (2013) Pulsed nanogenerator with huge instantaneous output power density. *ACS Nano* 7:7383–7391
- Du W, Han X, Lin L, Chen M, Li X, Pan C, Wang ZL (2014) A three dimensional multi-layered sliding triboelectric nanogenerator. *Adv Energy Mater* 4:1301592
- Fan FR, Tian ZQ, Wang ZL (2012) Flexible triboelectric generator. *Nano Energy* 1:328–334
- Fan FR, Tang W, Yao Y, Luo J, Zhang C, Wang ZL (2014) Complementary power output characteristics of electromagnetic generators and triboelectric generators. *Nanotechnology* 25:135402
- Han CB, Du W, Zhang C, Tang W, Zhang L, Wang ZL (2014a) harvesting energy from automobile brake in contact and non-contact mode by conjunction of triboelectrification and electrostatic-induction processes. *Nano Energy* 6:59–65
- Han CB, Zhang C, Li XH, Zhang LM, Zhou T, Hu WG, Wang ZL (2014b) Self-powered velocity and trajectory tracking sensor array made of planar triboelectric nanogenerator pixels. *Nano Energy* 9:325–333
- Han CB, Zhang C, Tang W, Li XH, Wang ZL (2015a) High power triboelectric nanogenerator based on printed circuit board (PCB) technology. *Nano Res* 8:722–730
- Han CB, Jiang T, Zhang C, Li X, Zhang C, Cao X, Wang ZL (2015b) Removal of particulate matter emissions from vehicle using a self-powered triboelectric filter. *ACS Nano* 9:12552–12561
- Han CB, Zhang C, Tian J, Li X, Zhang L, Li Z, Wang ZL (2015c) Triboelectrification induced UV emission from plasmon discharge. *Nano Res* 8:219–226
- Henniker J (1962) Triboelectricity in polymers. *Nature* 196:474
- Hindrichsen CC, Almind NS, Brodersen SH, Lou-Moller R, Hansen K, Thomsen EV (2010) Triaxial MEMS accelerometer with screen printed PZT thick film. *J Electroceram* 25:108–115
- Jiang T, Zhang LM, Chen X, Han C, Tang W, Zhang C, Xu L, Wang ZL (2015) Structural optimization of triboelectric nanogenerator for harvesting water wave energy. *ACS Nano* 9:12562–12572
- Jing QS, Zhu G, Bai P, Xie YN, Chen J, Han RPS, Wang ZL (2014) Case-encapsulated triboelectric nanogenerator for harvesting energy from reciprocating sliding motion. *ACS Nano* 8:3836–3842
- Li X, Han C, Jiang T, Zhang C, Wang ZL (2016) A ball-bearing structured triboelectric nanogenerator for nondestructive damage and rotating speed measurement. *Nanotechnology* 27:085401
- Lin L, Wang SH, Xie YN, Jing QS, Niu SM, Hu YF, Wang ZL (2013) Segmentally structured disk triboelectric nanogenerator for harvesting rotational mechanical energy. *Nano Lett* 13:2916–2923
- Niu SM, Wang SH, Lin L, Liu Y, Zhou YS, Hu YF, Wang ZL (2013) Theoretical study of contact-mode triboelectric nanogenerators as an effective power source. *Energy Environ Sci* 6:3576–3583
- Niu SM, Liu Y, Wang SH, Lin L, Zhou YS, Hu YF, Wang ZL (2014a) Theoretical investigation and structural optimization of single-electrode triboelectric nanogenerators. *Adv Funct Mater* 24:3332–3340
- Niu S, Wang S, Liu Y, Zhou Y, Lin L, Hu Y, Pradel KC, Wang ZL (2014b) Theoretical study of grating structured triboelectric nanogenerators. *Energy Environ Sci* 7:2339–2349
- Pang YK, Li XH, Chen MX, Han CB, Zhang C, Wang ZL (2015) Triboelectric nanogenerators as a self-powered 3D acceleration sensor. *ACS Appl Mater Interfaces* 7:19076
- Pu X, Liu M, Li L, Zhang C, Pang Y, Jiang C, Shao L, Hu W, Wang ZL (2016) Efficient charging of li-ion batteries with pulsed output current of triboelectric nanogenerators. *Adv Sci* 3:1500255

- Tang W, Han C, Zhang C, Wang ZL (2014a) Cover-sheet-based nanogenerator for charging mobile electronics using low-frequency body motion/vibration. *Nano Energy* 9:121–127
- Tang W, Zhang C, Han C, Wang ZL (2014b) Enhancing output power of cylindrical triboelectric nanogenerators by segmentation design and multilayer integration. *Adv Funct Mater* 24:6684–6690
- Tang W, Zhou T, Zhang C, Han CB, Wang ZL (2014c) A power-transformed-and-managed triboelectric nanogenerator and its application in the self-powered wireless sensing node. *Nanotechnology* 25:225402
- Tang W, Jiang T, Fan FR, Yu A, Zhang C, Cao X, Wang ZL (2015) Liquid-metal electrode for high-performance triboelectric nanogenerator at an instantaneous energy conversion efficiency of 70.6%. *Adv Funct Mater* 25:3718–3725
- Tian BZ, Zheng XL, Kempa TJ, Fang Y, Yu NF, Yu GH, Huang JL, Lieber CM (2007) Coaxial silicon nanowires as solar cells and nanoelectronic power sources. *Nature* 449:885–889
- Wang ZL (2008) Self-powered nanotech. *Sci Am* 298:82–87
- Wang ZL (2014) Triboelectric nanogenerators as new energy technology and self-powered sensors – principles, problems and perspectives. *Faraday Discuss* 176:447–458
- Wang ZL, Wu W (2012) Nanotechnology – enabled energy harvesting for self-powered micro-/nanosystems. *Angew Chem Int Ed* 51:11700–11721
- Wang S, Lin L, Wang ZL (2012) Nanoscale triboelectric-effect-enabled energy conversion for sustainably powering portable electronics. *Nano Lett* 12:6339–6346
- Wang SH, Lin L, Xie YN, Jing QS, Niu SM, Wang ZL (2013) Sliding-triboelectric nanogenerators based on in-plane charge-separation mechanism. *Nano Lett* 13:2226–2233
- Wang S, Xie Y, Niu S, Lin L, Wang ZL (2014a) Freestanding triboelectric-layer-based nanogenerators for harvesting energy from a moving object or human motion in contact and non-contact modes. *Adv Mater* 26:2818–2824
- Wang S, Xie Y, Niu S, Lin L, Liu C, Zhou YS, Wang ZL (2014b) Maximum surface charge density for triboelectric nanogenerators achieved by ionized-air injection: methodology and theoretical understanding. *Adv Mater* 26:6720–6728
- Williams CB, Shearwood C, Harradine MA, Mellor PH, Birch TS, Yates RB (2001) Development of an electromagnetic micro-generator. *IEEE Proc-Circuits, Devices Syst* 148:337–342
- Xie Y, Wang S, Niu S, Lin L, Jing Q, Yang J, Wu Z, Wang ZL (2014) Grating-structured freestanding triboelectric-layer nanogenerator for harvesting mechanical energy at 85% total conversion efficiency. *Adv Mater* 26:6599–6607
- Xu L, Pang Y, Zhang C, Jiang T, Chen X, Luo J, Tang W, Cao X, Wang ZL (2017) Integrated triboelectric nanogenerator array based on air-driven membrane structures for water wave energy harvesting. *Nano Energy* 31:351–358
- Yang Y, Zhou YS, Zhang HL, Liu Y, Lee S, Wang ZL (2013) A single-electrode based triboelectric nanogenerator as self-powered tracking system. *Adv Mater* 25:6594–6601
- Zhang H, Yang Y, Zhong X, Su Y, Zhou Y, Hu C, Wang ZL (2013) Single-electrode-based rotating triboelectric nanogenerator for harvesting energy from tires. *ACS Nano* 8:680–689
- Zhang C, Zhou T, Tang W, Han CB, Zhang LM, Wang ZL (2014a) Rotating disk based direct-current triboelectric nanogenerator. *Adv Energy Mater* 4:1301798
- Zhang LM, Xue F, Du WM, Han CB, Zhang C, Wang ZL (2014b) Transparent paper-based triboelectric nanogenerator as page mark and anti-theft sensor. *Nano Res* 7:1215–1223
- Zhang C, Tang W, Han CB, Fan FR, Wang ZL (2014c) Theoretical comparison, equivalent transformation and conjunction operations of electromagnetic induction generator and triboelectric nanogenerator for harvesting mechanical energy. *Adv Mater* 26:3580–3591
- Zhang C, Tang W, Pang Y, Han CB, Wang ZL (2015) Active micro-actuators for optical modulation based on a planar sliding triboelectric nanogenerator. *Adv Mater* 27:719–726
- Zhang LM, Han CB, Jiang T, Zhou T, Li XH, Zhang C, Wang ZL (2016) Multilayer wavy-structured robust triboelectric nanogenerator for harvesting water wave energy. *Nano Energy* 22:87–94

- Zhou T, Zhang C, Han CB, Fan FR, Tang W, Wang ZL (2014) Woven structured triboelectric nanogenerator for wearable devices. *ACS Appl Mater Interfaces* 6:14695
- Zhou T, Zhang LM, Xue F, Tang W, Zhang C, Wang ZL (2016) Multilayered electret films based triboelectric nanogenerator. *Nano Res* 9:1442–1451
- Zhu G, Pan C, Guo W, Chen CY, Zhou Y, Yu R, Wang ZL (2012) Triboelectric-generator-driven pulse electrodeposition for micropatterning. *Nano Lett* 12:4960–4965
- Zhu G, Chen J, Liu Y, Bai P, Zhou YS, Jing QS, Pan CF, Wang ZL (2013a) Linear-grating triboelectric generator based on sliding electrification. *Nano Lett* 13:2282–2289
- Zhu G, Bai P, Chen J, Wang ZL (2013b) Power-generating shoe insole based on triboelectric nanogenerators for self-powered consumer electronics. *Nano Energy* 2:688–692
- Zi YL, Guo HY, Wen Z, Yeh MH, Hu CG, Wang ZL (2016) Harvesting low-frequency (< 5 Hz) irregular mechanical energy: a possible killer application of triboelectric nanogenerator. *ACS Nano* 10:4797–4805

---

**Part IX**

**Implantable Medical Devices**



# Microelectrode Array

Renxin Wang, Huaiqiang Yu, and Zhihong Li

## Contents

Overview .....	1380
Considerations of Implantable Microelectrode Arrays (MEAs) .....	1380
Biocompatibility .....	1380
Impedance .....	1381
Charge Transfer Capability .....	1382
MEAs for Central Nervous System .....	1383
Flexible Planar MEAs on the Surface of the Cortex .....	1384
Flexible Protruding MEAs on the Surface of the Cortex .....	1384
Penetrating MEAs Implanted into the Brain .....	1387
MEAs for Peripheral Nervous System .....	1393
MEAs for Retinal Prosthesis .....	1394
MEAs Placed on the Nerve .....	1401
Perspective .....	1406
References .....	1408

## Abstract

Microelectrode arrays (MEAs) have been applied as chronic interface with the neural system and play important roles in neural prosthesis for various diseases, including sensory and motor injuries such as blind, deaf, and paralyzed, and also mental diseases like depression, Parkinson's disease, and epilepsy. Thanks the

---

R. Wang

Science and Technology on Electronic Test and Measurement Laboratory, North University of China, Taiyuan, China

H. Yu

Research Institute, China Electronics Technology Group Corporation, Chongqing, China

Z. Li (✉)

National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Beijing, China

e-mail: [zhli@pku.edu.cn](mailto:zhli@pku.edu.cn)



inherent merits of microfabrication, MEAs show the advantages of low cost, mass production, high density, flexibility (optional), small footprint, integratability with ICs (integrated circuits), etc. In this chapter, we discuss about the general requirements and consideration of materials, design, and fabrication of MEAs in details. Then the various devices and their applications in the central nervous system and the peripheral nervous system are overviewed, respectively.

---

**Keywords**

Microelectrode array · Neural prosthesis · Neural stimulation · Neural recording

---

## Overview

Neural prosthesis is a term referred to the artificial device involving neural stimulation and recording, which is aimed to partly replace or repair the function of a damaged neural system. Numerous efforts have been put into neural prosthesis for various diseases, including sensory and motion injuries such as blind, deaf, and paralyzed, and also mental diseases like depression, Parkinson's disease, and epilepsy. With the development of neural prosthesis, microelectrode arrays (MEAs) have been paid more and more attention, which are considered as chronic interface with the neural system. Hence, as the electrical frontend of neural prosthesis directly bridges to neural system, the properties of MEAs would make enormous contributions to the prosthetic effect. Nervous systems consist of central nervous system and peripheral nervous system. Accordingly, in this chapter, MEAs are divided into two categories for the target nervous systems.

---

## Considerations of Implantable Microelectrode Arrays (MEAs)

There are various implantable microelectrode arrays for diverse applications, and some general considerations should be taken into account in term of their implanting and electric characters, including biocompatibility, impedance, charge transfer capability, flexibility, protruding, and microfabrication.

### Biocompatibility

In view of implantation application, biocompatibility should be seriously considered as the critical property. The biocompatibility includes material biocompatibility (Zhou and Greenbaum 2009), electrical biocompatibility (Cogan 2008), and mechanical biocompatibility (Hassler et al. 2011; Humayun 2001). Biocompatible materials should be adopted as structural materials of the microelectrode, where nontoxicity and long-term bio-stability are of first consideration (Zhou and Greenbaum 2009).

The electrode interface impedance and charge transfer capability of an electrode would play important roles in electrical biocompatibility, especially for stimulating

implants (Cogan 2008). Firstly, to electrically stimulate the neurons, electrical pulses with certain current density are required. Lower electrode interface impedance means lower voltage and less power consumption, less heat producing, and also less damage. On the other hand, for recording implants, low electrode interface impedance would be in favor of improving the signal-noise ratio (Cogan 2008). Secondly, irreversible Faradic reactions involving electron transferring might occur when electrical pulse is applied, which would bring damage to the tissues. Charge transfer capability, as a factor to judge the threshold of irreversible Faradic reaction, should be concerned.

As for mechanical biocompatibility, weight, shape, and flexibility should be emphasized. During surgical intervention, physical tissue damages are inevitable. Therefore, it is obvious that light, the small and flexible implant could reduce the damage degree (Humayun 2001). On the other hand, micromovement may occur after implantation due to mechanical mismatch, which may cause chronic damage or inflammation. It should be noted that any surgical intervention is accompanied by an inflammatory response as a normal physiological response (Hassler et al. 2011). Slight inflammatory response might be favorable for the fixation of implantation. However, continual inflammation may eventually encapsulate the implant. Hence, small and flexible properties are more preferable.

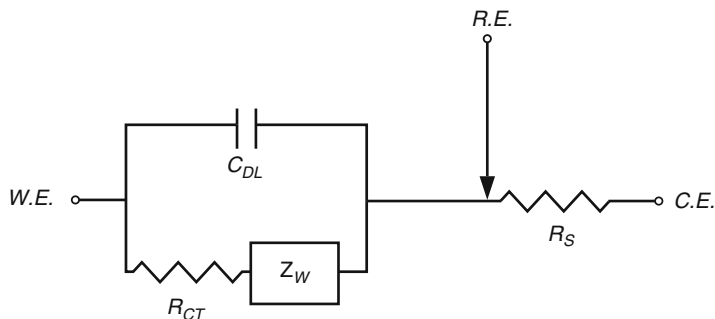
## Impedance

The interface impedance of an electrode is an important factor to determine whether the electrode could stimulate or record efficiently.

Herein, a widely used circuit model proposed by Randles (1947) is applied to describe the behavior of electrode-electrolyte impedance. As shown in Fig. 1, electrode impedance includes the series resistance ( $R_S$ ), the double layer capacitance ( $C_{DL}$ ), the charge transfer resistance ( $R_{CT}$ ), and the Warburg impedance ( $Z_W$ ).  $R_S$  comes from the passage of current through the resistive material, including the wires, the metal in the electrode, and the electrolyte.  $C_{DL}$  represents the buildup of charge at the electrode-electrolyte interface, which is proportional to the surface area of the electrode ( $C_{DL} = \epsilon \cdot A/d$ ).  $R_{CT}$  describes the charge transfer of chemical reaction and is only significant under very small signal test.  $R_{CT}$  is inversely dependent on the surface area ( $R_{CT} = \rho \cdot l/A$ ).  $Z_W$  corresponds to the effect of reactant transport on the rate of charge transfer, which increases with  $\frac{1}{\omega^{1/2}}$  ( $Z_W = \frac{\sigma}{(j\omega)^{1/2}}$ ), where  $\sigma$  is the Warburg coefficient. The Randles' model can be simplified under different conditions. At low frequencies, the impedance related to  $C_{DL}$ , which is proportional to  $1/\omega$ , is large enough to be omitted, compared to  $Z_W$  and  $R_{CT}$ , and can be replaced by an open circuit (Wang et al. 2010). Then,

$$\tilde{Z}_{LF} = R_S + \frac{\rho l}{A} + \frac{\sigma}{(j\omega)^{1/2}}. \quad (1)$$

On the other hand, at high frequencies ( $Z_W \ll R_{CT}$ ),  $Z_W$  can be replaced by a short circuit. Then,



**Fig. 1** The primary equivalent circuit describing the interface impedance of MEA and electrolyte

$$\tilde{Z}_{HF} = R_S + \frac{R_{CT}}{j\omega C_{DL}R_{CT} + 1} = R_S + \frac{1}{A} \frac{\rho l d}{j\omega \epsilon \rho l + d}. \quad (2)$$

According to Eqs. 1 and 2, the impedances in both low and high frequencies are decreasing with  $A$  and  $\omega$ . In the present work, 1 kHz is the typical frequency used to detect the electrical properties of the microelectrodes. At this condition, the high frequency model (2) is more suitable for the equivalent circuit. With this model, the phase can also be calculated.

## Charge Transfer Capability

Faradaic reaction means an electrochemical reaction in which chemical species are oxidized or reduced. If the transferred charges exceed a certain reversible limit, the irreversible Faradaic reaction starts up, which may produce harmful ions into tissues (Cogan 2008). Charge transfer capability of material is the factor to judge the reversible limit. Hence, the materials with considerable charge transfer capabilities should be utilized for electrical stimulation (Zhou and Greenbaum 2009).

The noble metals and their compounds such as gold, platinum, iridium oxide are widely taken as stimulating electrode materials. The charge transfer capability is  $0.04 \text{ mC/cm}^2$  for gold,  $0.1 \text{ mC/cm}^2$  for platinum, and  $1 \text{ mC/cm}^2$  for iridium oxide (Margalit et al. 2002), respectively.

## Flexibility

In order to facilitate surgery operation and fulfill effective stimulation/recording, many MEAs are made of rigid materials. However, as aforementioned, the rigid substrate may not conform to the curving surface of target tissue and lead to loss of proper contact (Hassler et al. 2011). Furthermore, such rigid MEAs may cause irreversible tissue damage during applications such as chronic implantation in tissue. Therefore, the flexible substrate seems to be more favorable in conforming to the curving surface and resisting the movement (Zhou et al. 2009; Humayun 2001).

### **Protruding**

Many reports have indicated that protruding electrode especially microneedle structure could be favorable in reducing the impedance and improving stimulation/recording (Wang et al. 2012).

To be planar or protruding, it is a question. However, there is no universal answer for diverse implant applications. It depends on the cost and feasibility (Tombran-Tink et al. 2007).

### **Microfabrication**

With regard to aforementioned considerations, size and material choice should be well designed (Humayun 2001). On the other hand, heterogeneous integration of different materials especially flexible materials is a challenge to microfabrication (Wang et al. 2012). Besides, how to manufacture high-density electrode channels on the limited space also requires elaborated scheme (Rodger et al. 2008). Obviously, the difficulty in producing 3D structure would be dramatically increased. Hence, it should be noted that microfabrication plays an important role in MEA design, to which should be paid attention.

---

## **MEAs for Central Nervous System**

Central nervous system consists of the brain and spinal cord. The brain is still a big scientific mystery. In order to understand the function of the brain, researchers have been struggling to study the brain and seek new ways to treat and finally prevent its disorders. As an important and effective tool, the implantable microelectrode is fabricated and used to record action potentials from neurons and stimulate specific nervous tissues in brain.

In the 1950s, discharges of single neurons in brain of mammals were recorded for a week or more by the first use of implantable microelectrodes, which were made from the 80- $\mu\text{m}$  stainless-steel wires with insulation exposed at the tip (Strumwasser 1958). Such manually fabricated microwire arrays were widely used. However, because of nonautomated fabrication, it is difficult to guarantee the shape, size, pitch, and impedance level of the electrode arrays. There was no revolutionary change in microelectrode techniques in the last several decades.

In the 1970s, Wise et al. utilized silicon micromachining and integrated-circuit fabrication techniques with inherent high accuracy and repeatability to fabricate the first silicon-based multi-electrode microprobes (Wise and Angell 1975). Since then, silicon surface and bulk micromachining techniques have been introduced to develop implantable microelectrode arrays.

According to implantation location, these microelectrode arrays could be classified into three major catalogues: the first one is flexible planar MEAs placed inside the skull and on the surface of the cerebral cortex, which are usually used in electrocorticography (ECoG); the second is flexible protruding MEAs that are also placed on the cortex, but they have rigid 3D electrodes that can penetrate the cortex

shallowly and form a better contact between the electrode sites and the neural tissues; the third is microprobe arrays penetrated into the brain, which can be placed deep inside the brain and be close to the target neurons to record the action potentials or to apply the stimulus. Here we focus on the design, fabrication, and in vivo test of these implantable MEAs, as shown in Table 1.

## Flexible Planar MEAs on the Surface of the Cortex

Flexible MEAs have been widely studied in implantation applications for their biocompatibility and low invasiveness, where flexible materials such as parylene and polyimide are chosen.

Rodger et al. demonstrated a kind of flexible bio-interfaced MEAs with ultrathin electronics, whose substrate was bioresorbable silk fibroin (Kim et al. 2010). Yamagiwa et al. reported a flexible self-curling and sticking ECoG MEA to conform to the cortex in one direction (Yamagiwa et al. 2013). Lee developed a kind of MEA integrated with aligned silicon-nanowire FET switches (Lee 2017), where silicon nanowires were transferred to a flexible substrate by wafer bonding and thinning processes (Figs. 2, 3, and 4).

It should be noted that the arithmetic product of impedance with foot area of microelectrode is usually big, which would dramatically influence the stimulation and recording. To solve this problem of planar MEA, Jang et al. presented a kind of planar MEA based on graphene, ZnO nanowires, and conducting poly[3,4-ethylenedioxythiophene](PEDOT) that could be capable of flexibility and low impedance (Ryu et al. 2017). ZnO nanowires were introduced to increase the effective surface area and consequently realize the low impedance. However, this MEA was not suitable for long-term application because the nanowires were too brittle to resist tissue displacement (Fig. 5).

## Flexible Protruding MEAs on the Surface of the Cortex

In order to get close to the target neurons in the cortex and enlarge the effective electrode area, flexible protruding MEAs are developed to perform more effective and accurate electrical conduction.

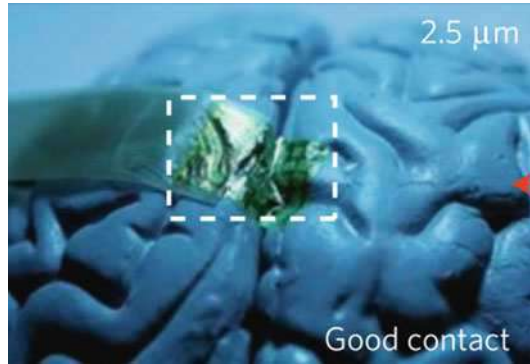
Zhou et al. demonstrated a pyramid-shaped polymer MEA coated by patterned metal (Zhou et al. 2009). Böhringer et al. proposed a polyimide-based interconnect approach to realize individual needle-shaped microelectrodes (Holman et al. 2002). However, this approach was aimed to perform accurate single-point stimulation and recording. Li et al. proposed a novel process to fabricate flexible microneedle electrode array. The process included creation of a silicon microneedle electrode array using an SOI substrate followed by its release using parylene films as the support material. The proposed structure provided flexibility of a parylene thin film and rigidity of a silicon structure (Wang et al. 2012). Furthermore, to raise the height of microneedle, an improved process involved with dicing and KOH was developed (Wang et al. 2017) (Figs. 6 and 7).

**Table 1** Contrasts of various reported intracortical MEAs for superficial brain stimulation

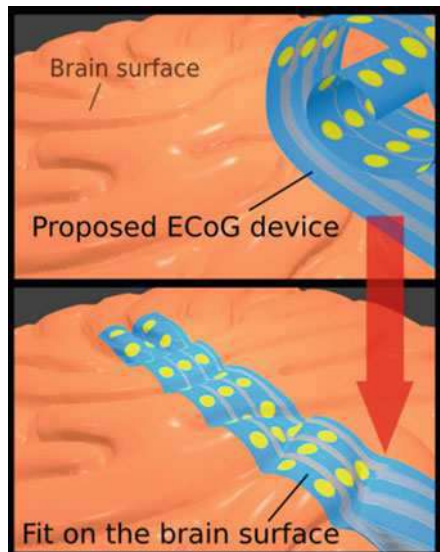
Citation	Substrate material	Rigid or flexible	Planar or protruding	Protruding shape and height	Electrode material	Foot area ( $\mu\text{m}^2$ )	Impedance ( $\text{k}\Omega@1\text{kHz}$ )
(Kim et al. 2010)	Polyimide	Flexible	Planar	None	Au	250,000	—
(Yamagiwa et al. 2013)	Parylene	Flexible	Planar	None	Pt	785,000	3.2
(Lee 2017)	Polyimide	Flexible	Planar	None	Au	706	161
(Ryu et al. 2017)	PEDOT	Flexible	Planar	None	Au on nanowire	640,000	0.52
(Zhou et al. 2009)	Polyimide	Flexible	Protruding	Pyramid, 56 $\mu\text{m}$	Polyimide + Pt	6400	—
(Holman et al. 2002)	Polyimide	Flexible	Protruding	Cone, 400 $\mu\text{m}$	Silicon	—	—
(Wang et al. 2012)	Parylene	Flexible	Protruding	Nail, 80 $\mu\text{m}$	Silicon+Pt	7850	11.4
(Wang et al. 2017)	Parylene	Flexible	Protruding	Pyramid, 190 $\mu\text{m}$	Silicon+Au	14,400	90

“—” means the content is not mentioned in the report

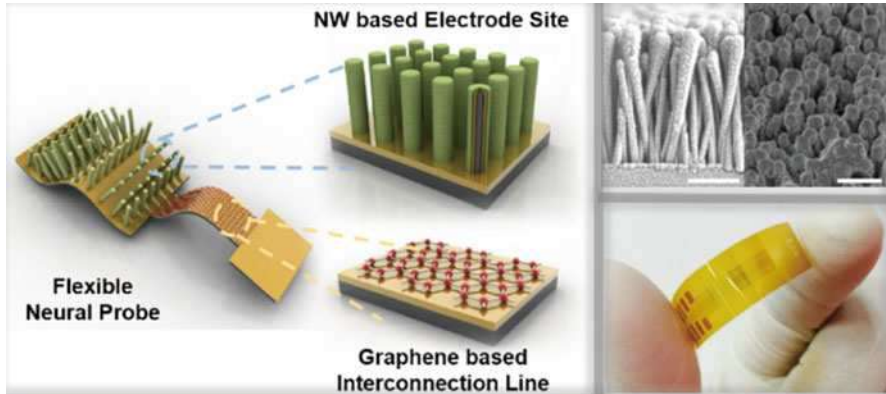
**Fig. 2** Schematic image of ultrathin and bioresorbable MEA



**Fig. 3** Schematic image of self-curling and sticking MEA



**Fig. 4** Photo of MEA with silicon-nanowire FET switches



**Fig. 5** Schematic image and photo of the MEA based on the graphene and nanowire

## Penetrating MEAs Implanted into the Brain

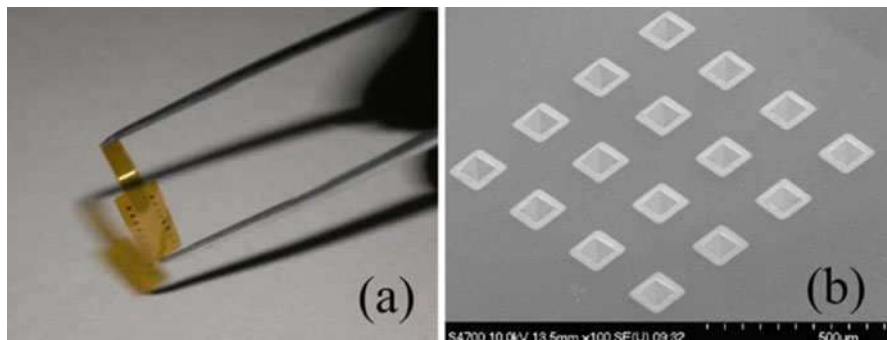
In order to measure extracellular action potentials or send electrical impulses to specific targets in the cerebral cortex or beneath the cortex, MEAs must be penetrated into the brain and be close to the cell membrane. This type of electrode has one or more long shafts with enough robustness and the electrode sites are distributed on the shank or at the tip. According to the material of shaft, MEAs penetrated into the brain are classified into three types: silicon-based microprobes, metal-based microprobes, and polymer-based microprobes.

### Silicon-Based Microprobes

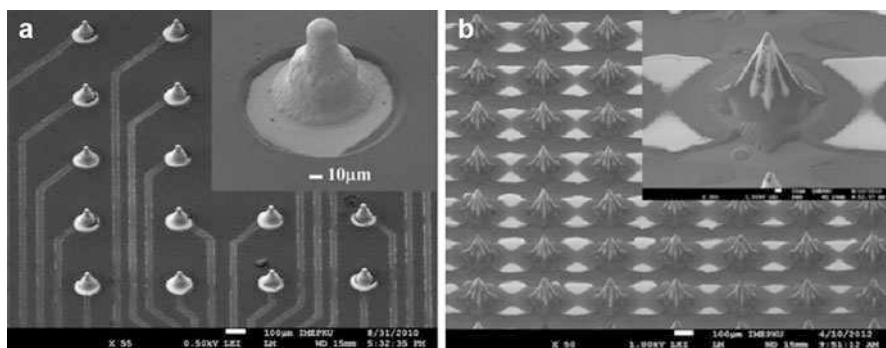
Since the 1990s, a few of researchers took advantages of silicon MEMS fabrication process and developed silicon-based microprobes incorporating high-density microelectrode arrays. There are two major types of silicon needle structure: vertically oriented and horizontally oriented. A well-known example is the Utah electrode array reported by Normann et al., which is a  $10 \times 10$  silicon needle array (Normann et al. 1999), as shown in Fig. 8. Each needle has only one independent electrode site exposed at the tip. A few of microfabrication technologies including thermal diffusion, diamond dicing, and wet etching are used to create sharpened needle array. The length of the needle is limited by the thickness of the silicon wafer (typically 1.5 mm), and therefore this type of electrode cannot reach the deep region in the brain while the substrate must rest on the surface of the brain. These arrays have been used in cat visual cortex and sensory cortex (Rousche and Normann 1999). Recently, in order to restore arm and hand function for paralyzed persons, Utah arrays are widely used to record neuron activities from the motor cortex (Velliste et al. 2008; Ethier et al. 2012).

The other type of silicon-based microprobe is called Michigan probe, whose needles are horizontally oriented. In the 1985, Najafi, Wise, and Mochizuki utilized dopant-selective etch (boron etch-stop) process to fabricate the silicon probe





**Fig. 6** Photo and SEM of pyramid-shaped MEA



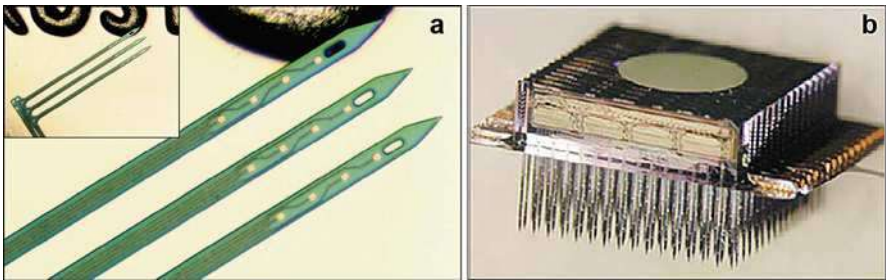
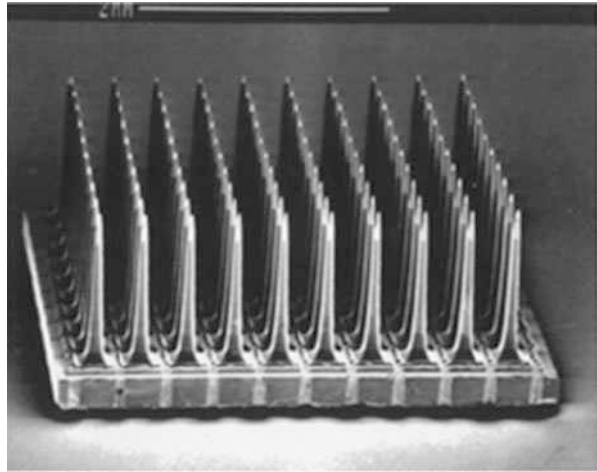
**Fig. 7** SEMs of tip-shaped (a) and pyramid-shaped (b) MEAs

incorporating MEAs (Najafi et al. 1985), and then they developed a variety of probes including single-shaft, multi-shaft, and 3-D-stacked layouts (Wise et al. 2004), as shown in Fig. 9. The thickness of the shank is limited by the maximum doping depth (on the order of 15  $\mu\text{m}$ ).

Since the 2000s, the work of silicon-based Michigan-type probe has focused on the use of DRIE technology combining with polymer thin-film process to improve electrode site's density, biocompatibility and flexibility of the electrode. In 2008, Huang et al. reported an integrated parylene-cabled silicon probe for neural prosthetics (Huang et al. 2008). The silicon-based probes were shaped by deep reactive ion etching (DRIE) technology and fully encapsulated by parylene C (poly-paraxylene C, a USP Class VI polymer). Meanwhile, a flexible cable was integrated for electrical interconnection between the electrode and external circuitry, as shown in Fig. 10.

In 2011, Merriam et al. described a 3-D 160-site MEAs integrated parylene cable, which employed the well-established 2-D probe process flow developed at the University of Michigan. 2D probes were assembled in a slotted platform that were

**Fig. 8** A scanning electron micrograph (SEM) of the Utah electrode array



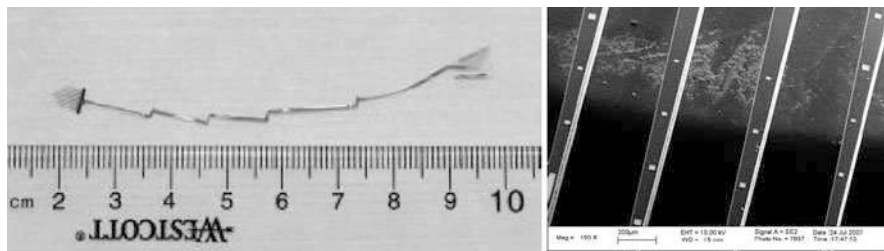
**Fig. 9** Pictures of Michigan probes: (a) a three-shank 12-site recording probe; (b) a 3-D 1024-site 128-channel neuroelectronic interface

fabricated by DRIE etches with a 1.18-mm<sup>2</sup> footprint. The parylene cable integrated in the platform was batch fabricated to route the electrode sites to a site-selection board, as shown in Fig. 11.

### Metal-Based Microprobes

Because of silicon's brittleness, the silicon-based microprobes are subjected to breakage during handling and insertion (Huang et al. 2008; Motta and Judy 2005). In addition, the fabrication process is normally time-consuming and expensive, because the most in-plane silicon substrate has to be removed using wet etching or DRIE.

To address these issues of silicon-based probes, metal-based microprobes were developed as an alternative. In general, microprobes made of ductile metals have higher mechanical performance than silicon-based ones in chronic implantation. A long metal shank will be elastically or plastically deformed when it is significantly



**Fig. 10** Released silicon probe with parylene flexible cable and SEM picture of the electrodes on the silicon shanks

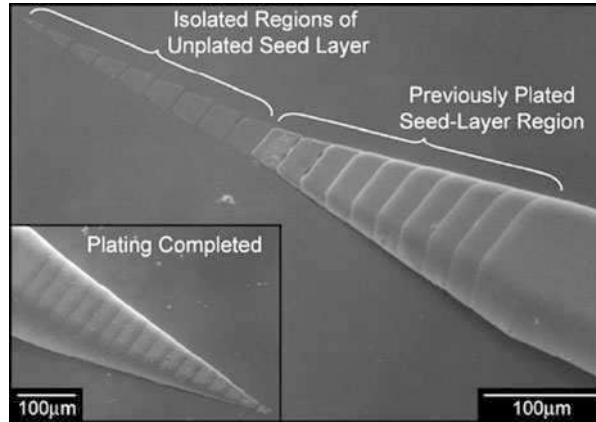
**Fig. 11** Photograph of a 160-site 3-D MEAs



bended, while one made of fragile material such as silicon or glass tends to break. Electroplating techniques are widely used to form the metal shank. Gold, copper, and nickel can be deposited with standard electroplating processes. Gold has high ductility and malleability and good compatibility. However, the gold-based probe lacks of robustness for inserting into the cortex; furthermore, gold electroplating has a low deposition rate and needs relatively expensive baths. Copper can easily and quickly be electroplated, but it also lacks of robustness and it is an unsuitable implant material due to high toxicity. Compared with copper, nickel electroplating has better mechanical performance, which can provide enough stiffness of probe. However, there are also some concerns about allergic reaction and toxicity of nickel.

In 2005, Motta et al. published nickel microprobes incorporating MEAs for deep brain stimulation (Motta and Judy 2005). The seed-layer was designed to create 3-D

**Fig. 12** SEM image of electrodeposition of the mechanical layer in the probe shank; the inset image on the lower left shows the finished tip of the electroplated microprobe

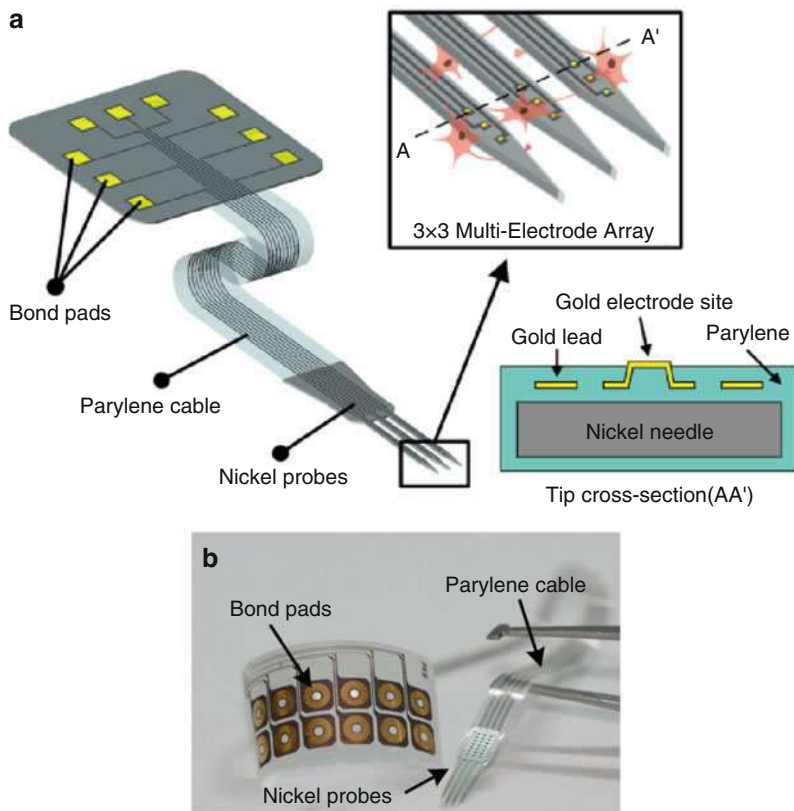


shank structure (Fig. 12). The electroplated nickel probe shank with the thickness of  $55\ \mu\text{m}$  was coated with  $10\text{-}\mu\text{m}$ -thick gold layer and silicon nitride for biocompatibility. In this work, 30-day implantation in the subthalamic nucleus of four rats demonstrated that gold-coated nickel microprobes show small astrocytic reaction and less glial fibrillary acidic protein (GFAP) reactivity than nickel probes.

The brain's Young's module is about  $1 \sim 10\ \text{kPa}$ . Due to the high Young's module, both silicon oxide ( $57 \sim 92\ \text{GPa}$ ) and silicon nitride ( $130 \sim 146\ \text{GPa}$ ) are not the suitable insulation material for implant. In addition, they are not chemically stable, so there has been risk of chemical failure after being implanted in the body. In order to address these issues, polymer materials such as polyimide, parylene, and silicone were widely used for insulation.

Fomani and Mansour described a penetrating MEA with electroplated gold microprobes and a flexible polyimide-based cable (Fomani and Mansour 2010). The thickness of the gold layer is  $15\ \mu\text{m}$ . However, the work did not show the results of microprobe mechanical test or implantation experiment. Takeuchi et al. developed a specific process to fold the planar nickel probe structures to create a 3D probe array with polyimide interconnections (Takeuchi et al. 2004). The probe shank consists of a  $5\text{-}\mu\text{m}$ -thick electroplated nickel layer and two  $5\text{-}\mu\text{m}$ -thick polyimide layers. This flexible probe array was inserted into a rat's brain without fracturing. And the MEA had successfully recorded neural signals in the rat's visual cortex.

In 2013, Yu et al. reported the work on electroplated nickel multi-electrode microprobes with a flexible parylene cable (Yu et al. 2013), as shown in Fig. 13. In order to improve biocompatibility, the nickel probes are fully encapsulated with parylene C that has been widely used as a coating material for biomedical implants. Due to the specific fabrication process, the electrodes are bumped out of the parylene surface, which allows a better contact with the neural tissue. Using these microprobes, the spontaneous neural activity in rat's hippocampus has been recorded and the wing beat of tethered worker honeybees has been successfully triggered by delivering the stimulating pulse into the optic lobes.



**Fig. 13** Schematic view of the proposed microprobes with flexible parylene cable as well as close-ups of stimulation and recording electrode sites: (a) 3-D drawing (not to scale) and (b) a typical device layout

### Flexible Polymer-Based Microprobes

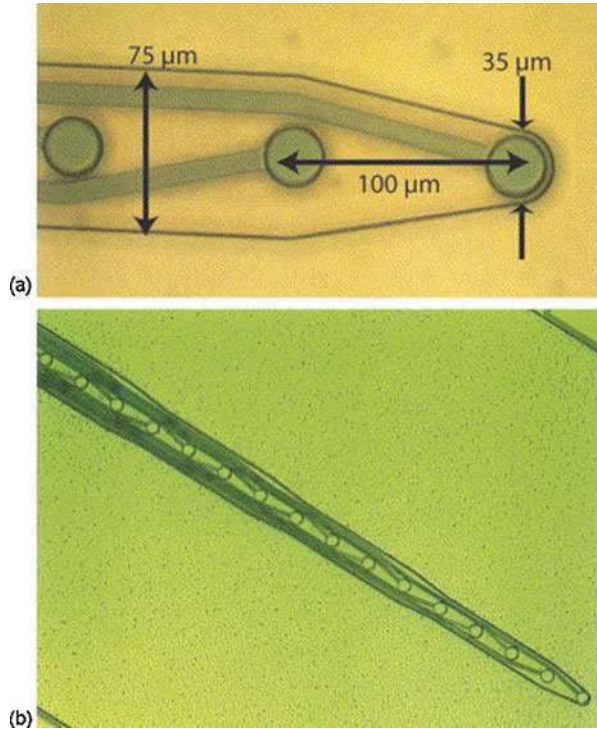
Due to the flexibility, polymer-based microprobes with good biocompatibility are suitable for long-term implantation in the brain. Polyimide and parylene are usually used as the mechanical support layer. The advantage of these soft materials is that their shapes can deform with the micro-motion of neural tissues and avoid fracturing, which will cause less damage of neural tissues than rigid materials.

In 2006, Cheung et al. reported a flexible polymer-based probe with 16-channel electrode for neural recordings *in vivo*, as shown in Fig. 14 (Cheung et al. 2007). The probe thickness is 15  $\mu\text{m}$  and the electrode site's diameter is 25  $\mu\text{m}$ . These probes obtained similar recording results, compared to silicon-based microelectrodes fabricated at University of Michigan in a chronic implantation experiment.

It is evitable that polymer-based probes with a long shank length are compliant, and therefore the probe tends to buckle during insertion. To address this issue, Lee et al. reported a polyimide-based probe using silicon as stiffness improving material (Lee et al. 2004). The results in the work showed that the probe with a silicon backbone layer of 5 ~ 10  $\mu\text{m}$  was robust enough to penetrate the pia of the rodent as



**Fig. 14** Pictures of the flexible polyimide-based neural probe with 16 electrode sites



shown in Fig. 15. However, the drawback of the work is that the use of silicon-on insulator (SOI) substrate increases the fabrication cost. Figure 16 shows the fabrication process.

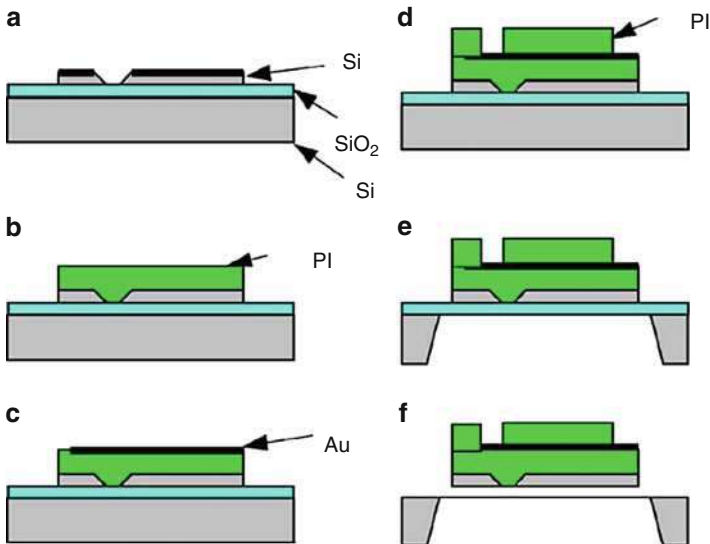
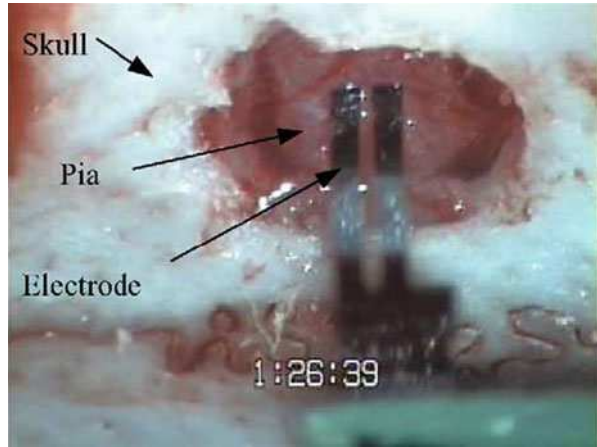
Without adding process complexity or increasing probe thickness and volume, Egert et al. described a parylene-based penetrating microprobe with integrating vertical parylene stiffeners as shown in Fig. 17 (Egert et al. 2011). The impedance of the electrode was characterized to be 350 k $\Omega$  for Pt, and 1.5 M $\Omega$  for Au. The probe can be easily inserted into nicked pia of the lamb.

In general, penetrating microprobes with MEAs should consider the factors including the flexibility of electrode, the biocompatibility of material, the design of probe structure, the probe stiffness, electrical properties to meet the chronic implantation in the brain.

## MEAs for Peripheral Nervous System

Peripheral nervous system (PNS) consists of the nerves and ganglia outside of the brain and spinal cord. The PNS mainly serves as connection of the CNS to the limbs and organs, including some sensory neural system such as retina and cochlea. By the way, the most widespread and matured one is cochlear prosthesis, which has been put into clinical application for several decades. Hence, academic attentions and

**Fig. 15** Optical microscope picture of the penetration test into rat's pia

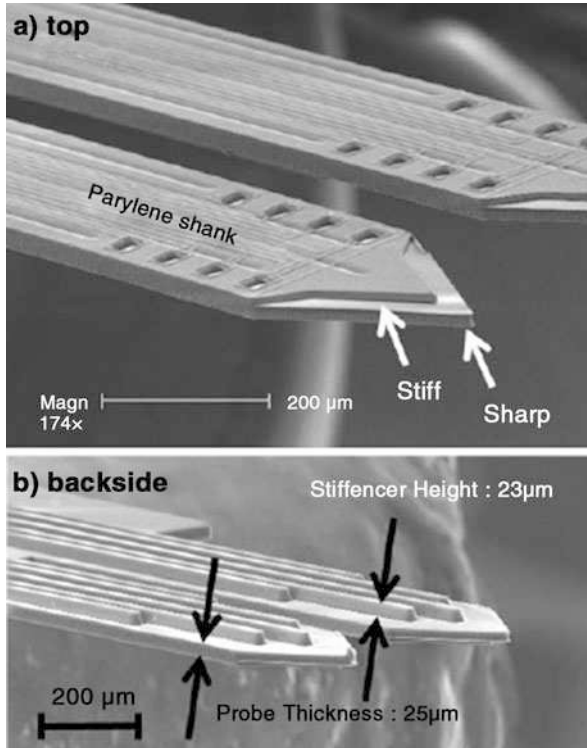


**Fig. 16** Fabrication procedure of the PI electrode with flexible region to accommodate micro-movement

ideas are focused on other undeveloped prosthesis and there is no individual section for cochlear prosthesis in this chapter.

## MEAs for Retinal Prosthesis

Retinal prosthesis is aimed to restore vision by stimulating the retina with controlled pattern. They could be divided into subretinal and epiretinal prosthesis according to



**Fig. 17** Parylene microprobes with engineered stiffness

the location of MEAs. Accordingly, various MEAs were developed, which could be sorted according to the flexibility and penetration (Fig. 18), as shown in Table 2.

### Rigid Planar MEAs

Li et al. presented an optical switchable MEA for retinal prosthesis, which is realized by optically modulating the resistance of photosensitive hydrogenated amorphous silicon ( $\alpha$ -Si:H) (Zhang et al. 2011) (Fig. 19).

Optobionics Inc. presented a rigid silicon prosthesis, which consisted of 3500 photodiode electrodes (Chow et al. 2004). However, the rigid and flat substrate was obviously not conformed to the shape of the retina, though subretinal implantation may make the retina deform to be flat.

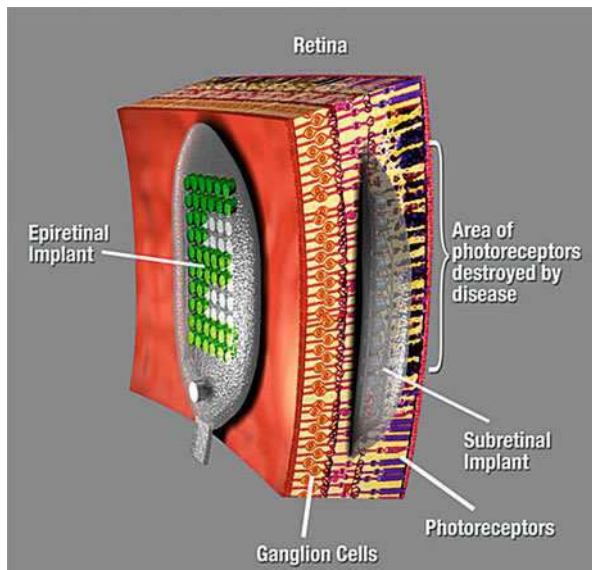
### Flexible Planar MEAs

To make the MEAs conform to the retina, flexible planar MEAs were developed.

Tai and Humayun et al. (Humayun 2001; Rodger et al. 2008) reported parylene-based planar flexible MEAs for the epiretinal prosthesis. It was successfully applied in Argus<sup>®</sup> II Retinal Prosthesis System, which was the first vision prosthesis to acquire regulatory authorization in both Europe and the USA (Luo and Cruz 2016).



**Fig. 18** Retinal prosthesis concept. The image could be captured by a video camera, which could be transmitted to the implanted MEA with a pattern



Humayun et al. presented a PDMS-based MEA, which was distorted toward a spherical curvature to fit the curved retina (Weiland et al. 2006). Intelligent Medical Implants AG (Zurich, Switzerland) performed clinical trial of an epiretinal implant, which featured a polyimide-based MEA comprised of 49 electrodes coated with IrOx (Keseru et al. 2012). Furthermore, larger MEA with 12 mm diameter was fabricated, which was flexible enough to be folded during implanting operation (Waschkowski et al. 2014) (Figs. 20, 21, and 22).

Zrenner developed an active subretinal polyimide-based MEA with 1500 micro-photodiodes, where the current pulses were determined through the amount of light received by the photodiodes (Gekeler et al. 2004; Zrenner et al. 2011). Retina Implant GmbH (Reutlingen, Germany) took clinical trials using this MEA in 2006, illustrating sufficient vision sensations for daily recognition of the blinds were regained (Gekeler et al. 2007). A novel retinal prosthetic MEA based on LCP was presented, which formed a thin and conformable device together with integrated circuits (Jeong et al. 2013).

Generally speaking, flexible planar MEAs are the most possible ones to be commercialized in retinal prosthesis owing to the flexibility and feasibility, in spite of their limited stimulating area.

### Rigid 3D MEAs

To reduce the distance of electrode and retina and so to improve the stimulating effects, rigid 3D MEAs were proposed.

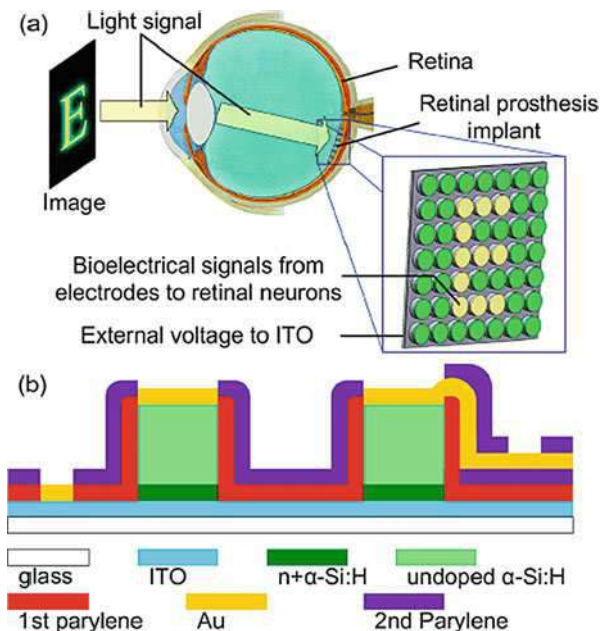
Palanker et al. presented protruding electrode arrays made of SU-8 for penetrating deep into the retina (Palanker et al. 2005). Johnson et al. proposed curved microwire glass MEA to bridge the gap between the flat silicon chip and the curved retina

**Table 2** Contrasts of various reported MEAs for retinal prosthesis

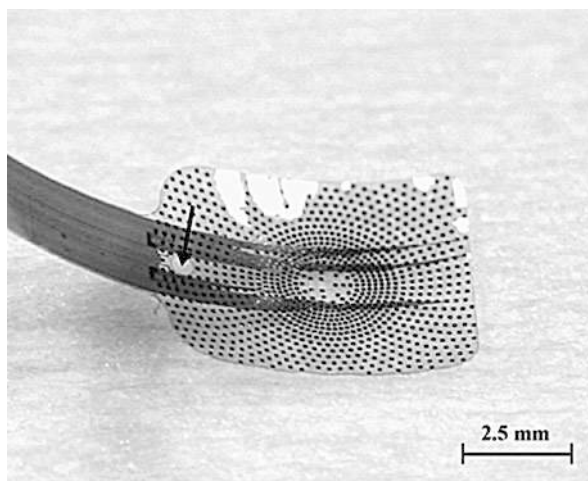
Citation	Implantation strategy	Substrate material	Rigid or flexible	Planar or 3D	Electrode material	Opening foot area ( $\mu\text{m}^2$ )	Impedance ( $\text{k}\Omega@1\text{kHz}$ )
(Zhang et al. 2011)	Epiretinal	Glass	Rigid	Planar	Au	7850	190
(Chow et al. 2004)	Subretinal	Silicon	Rigid	Planar	Iridium oxide	81	—
(Rodger et al. 2008)	Epiretinal	Parylene	Flexible	Planar	Iridium	4416	1.2
(Weiland et al. 2006)	Epiretinal	PDMS	Flexible	Planar	Au	70,650	—
(Keseru et al. 2012)	Epiretinal	Polyimide	Flexible	Planar	IrOx	7850	—
(Waschkowski et al. 2014)	Epiretinal	Polyimide	Flexible	Planar	IrOx	7850	2
(Zrenner et al. 2011)	Subretinal	Polyimide	Flexible	Planar	TiN	10,000	300
(Gekeler et al. 2007)	Subretinal	Polyimide	Flexible	Planar	Au	2500	—
(Jeong et al. 2013)	Epiretinal	Liquid crystal polymer	Flexible	Planar	Au	31,400	7.6
(Palanker et al. 2005)	Subretinal	SU-8	Rigid	3D	IrOx	78.5	50
(Johnson et al. 2004)	Epiretinal	Glass	Rigid	3D	Au	20,000	—
(Bhandari et al. 2008)	Epiretinal	Silicon	Rigid	3D	Silicon+Ir	15,386	—
(Ko and Lee 2017)	—	Polyimide	Flexible	3D	Au	1963	29
(Roessler et al. 2009)	Epiretinal	Polyimide	Flexible	3D	IrOx	7850	—
(Rui et al. 2011)	—	Parylene	Flexible	3D	Pt black	2,826	12.58
(Wang et al. 2010)	Subretinal	Parylene	Flexible	3D	Pt	7850	6.4

“—” means the content is not mentioned in the report

**Fig. 19** Optical switchable MEA. (a) Operating principle. (b) Schematic profile



**Fig. 20** Photo of parylene-based MEA applied in Argus<sup>®</sup> II



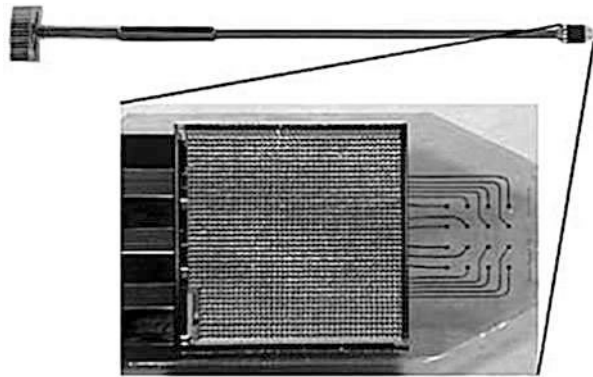
(Johnson et al. 2004). Bhandari et al. proposed a novel high aspect ratio silicon-based convoluted MEA for conforming to the curved surfaces of the retina (Bhandari et al. 2008) (Figs. 23 and 24).

However, the rigid substrate may inevitably bring physical injury to retina, which is rarely adopted in clinical surgery for retinal prosthesis.

**Fig. 21** Photo of polyimide-based MEA proposed by Intelligent Medical Implants AG



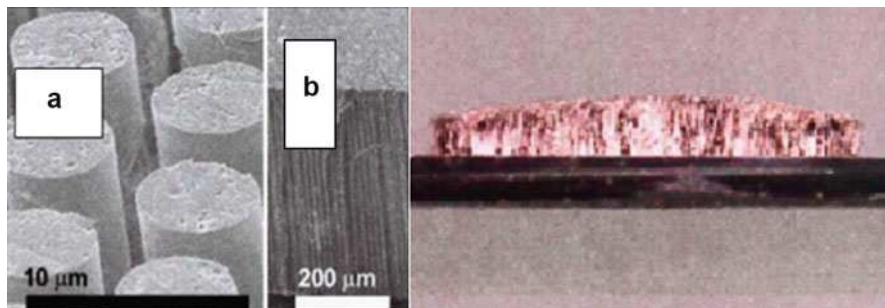
**Fig. 22** Photo of active polyimide-based MEA by Retina Implant GmbH



### Flexible 3D MEAs

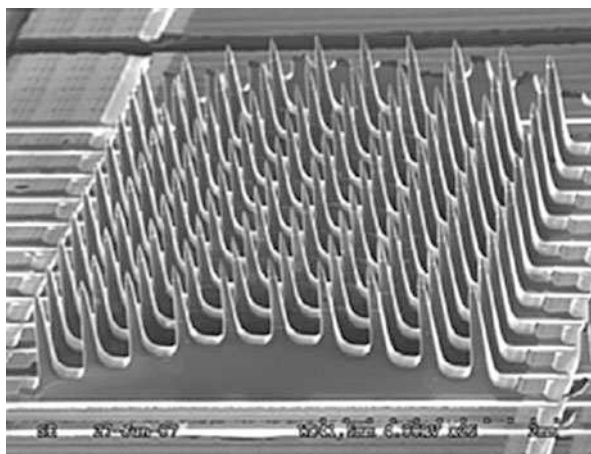
In order to combine the superiorities of flexibility and 3D, the 3D MEAs' designs on flexible substrates came out.

Ko et al. presented a comprehensive electrical characterization of 2D and 3D microelectrodes (Ko and Lee 2017). A device named EPIRET3 was developed, consisting of IrOx MEA for stimulation, the microcircuits packaged on the polyimide substrate, and the microcoil located underneath the circuits (Fig. 25).



**Fig. 23** (a) SEM and (b) photo of the microwire glass electrode

**Fig. 24** SEM of silicon-based convoluted MEA



Liu et al. fabricated hemispherical-shaped MEA based on parylene in order to increase effective stimulating area and decrease the impedance (Rui et al. 2011). Pt microelectrode has an impedance magnitude of 79.43 k $\Omega$ @1KHz, whereas Pt-black coated one is 12.58 k $\Omega$ @1KHz. However, as their fabrication was based on melting photoresist technique, the 3D microelectrode was limited in aspect ratio and not able to focus on electrical field (Fig. 26).

Li et al. proposed a process to fabricate 3D hollow tip-shaped microelectrodes on the parylene substrate (Wang et al. 2010). Employing a shape-transferring technique, the parylene 3D MEA with a sharp tip profile was readily replicated from the silicon tip mold. In the process, an aluminum-assisted lift-off technique was developed to realize high-quality lift-off process on a highly rugged substrate without using ultrasonic vibration. The electrical and preliminary surgical characterization showed the superiority of this 3D MEA compared to the planar one (Fig. 27).

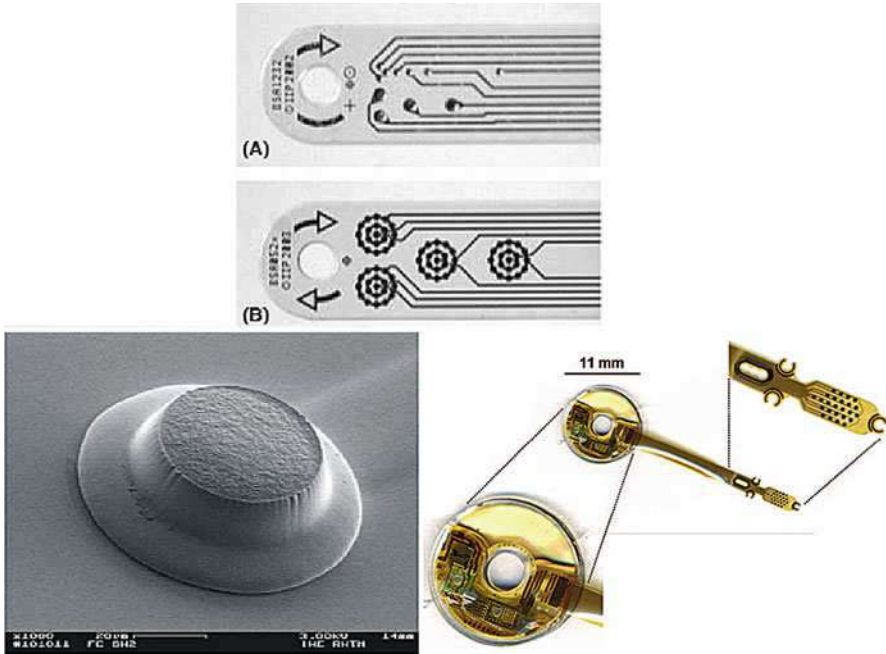


Fig. 25 Illustrations of EPIRET3

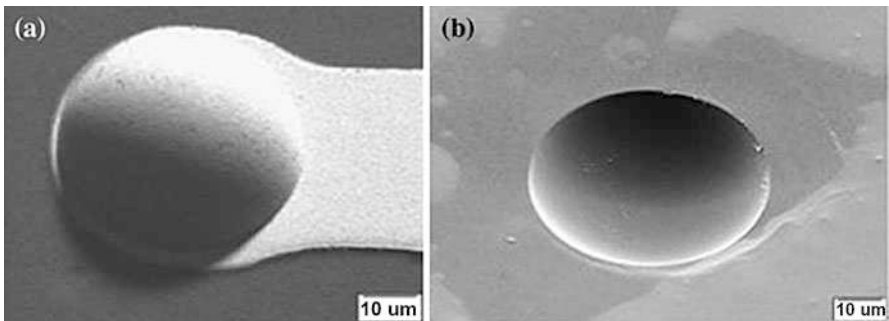
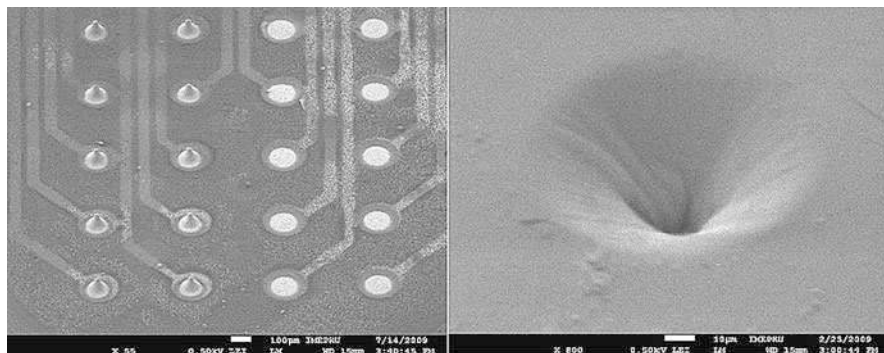


Fig. 26 Front side and back side of hollow hemispherical-shaped electrode

### MEAs Placed on the Nerve

Among the most successful PNS interface electrodes are the cuff electrodes, which hold the target nerve close to the electrode sites that are distributed on the inside walls of a polymeric insulating sheath. So far, cuff electrodes have been widely and successfully utilized in both basic research and clinical practice, such as vagus nerve stimulation (Engineer et al. 2011; De Ferrari 2011; Castoro 2011), recording





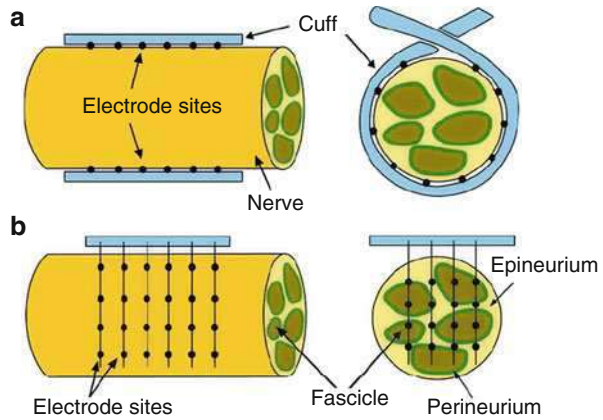
**Fig. 27** Front side and back side of hollow tip-shaped electrode

electroneurograms (ENGs) (Struijk et al. 1999; Yoo and Durand 2005; Chu 2012), treatment of fecal incontinence (Matzel et al. 2001), bladder control (Rijkhoff 1994), activation of paralyzed extremities (Polasek et al. 2009), and treatment of obstructive sleep apnea (OSA) (Eastwood 2011).

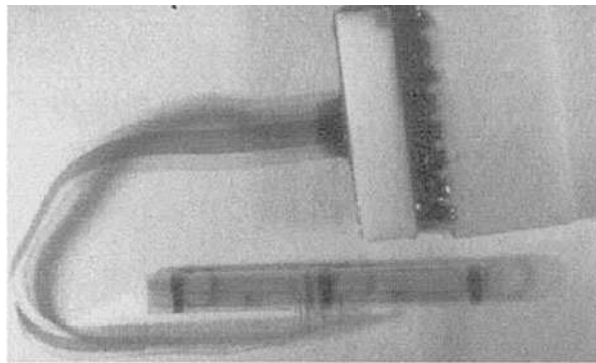
Cuff electrodes offer several advantages over other commonly used electrodes such as intramuscular, epimysial, and surface electrodes (Crago et al. 1980; Peckham and Knutson 2005). First, the stimulus intensity required for nerve activation is decreased. Therefore, tissue damage induced by system power and stimulation can be minimized. Second, multiple contacts inside one cuff electrode can achieve selective stimulation of different axonal fascicles within the nerve, producing potentially different functional outputs. Third, the electrode and its lead wires are less prone to mechanical failure and movement relative to the target tissues, which are caused by repeated muscle contraction after implantation. Further, in contrast with intrafascicular electrodes that are placed within fascicles in the nerve and directly contacted with the target fibers, such as fine wires (Lawrence et al. 2003), and silicon or polyimide probes (Branner and Normann 2000; Boretius 2010), cuff electrodes are safer and more reliable for chronic implantation, though they have a lower selectivity for stimulation and recording. As shown in Fig. 28, the electrode sites of cuff electrodes are placed on the surface of the nerve, while intrafascicular electrodes must penetrate the epineurium around the entire nerve and the perineurium around the fascicles (Branner et al. 2001). There are some reports showing that the cuff electrodes have been in use for as long as 15 years (Glenn and Phelps 1985; Waters et al. 1985).

Traditional cuff electrodes are fabricated by manually assembling discrete components, with little possibility of minimizing the device dimensions and increasing the number of electrode sites for both stimulation and recording to large populations of peripheral nerve fibers (Pratt et al. 1991; Loeb 1993; Carp et al. 2005; Naples et al. 1988). Microfabrication technologies with advantages of small device size, high reproducibility, batch fabrication, and system integration allow developing flexible, high-density microelectrode arrays (MEAs) for neural interface systems. There are

**Fig. 28** Comparison of (a) the cuff electrode and (b) the intrafascicular electrode. The positions of electrode sites are shown in both pictures (from the side and the front)



**Fig. 29** Micrographs of the cuff electrode with an attached plastic connector



two types of microfabricated cuff electrodes: self-curling spiral cuff electrode and self-locking cuff electrode.

In 2000, Rodríguez et al. reported a microfabricated polyimide-based self-curling spiral cuff electrode with light weight and high flexibility (Rodríguez 2000), which were composed of a 10- $\mu\text{m}$ -thick polyimide layer and three platinum strip electrodes with a 5 mm inter-electrode distance (Fig. 29). In order to obtain a self-curling cuff, first, the planar electrode with a polyimide-metal-polyimide sandwich structure was fabricated by standard fabrication process flow. Then, the planar electrode was inserted, rolled, and fixed in a specially designed tool to keep the final cuff-shape in a temper step at 340 °C for 2 h. After the temper step, the cuffs were released from the tool and remained stable in a cylinder spiral shape.

In their study, the polyimide PI2611 was used to overcome the problems including substrate stress fracture and high water absorption, previously reported from in vivo test of other types of polyimide such as PI2525 (González and Rodríguez 1997). The cuff electrodes induced only a very mild foreign body reaction and did not change the nerve shape over a 2–6 month implantation period, as shown in Fig. 30.



**Fig. 30** A cuffed nerve  
2 months post implantation



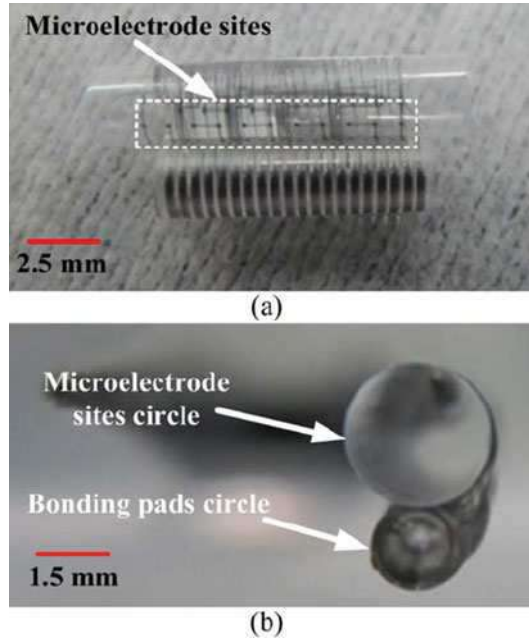
As an improved version, Navarro et al. developed a polyimide spiral cuff electrode with 12 platinum electrode sites (Navarro et al. 2001) and successfully demonstrated selective fascicular stimulation of the rat sciatic nerve using this type of cuff electrodes. In 2014, Kang et al. reported a parylene-based spiral cuff electrode (Kang et al. 2015), and cylinder spiral cuff shape is realized by the difference in self-stress of different parylene layers due to the heat treatments, as shown in Fig. 31. Moreover, the multichannel neural recording and stimulation of the sciatic nerve were successfully demonstrated in this work.

A relatively small cuff may cause the nerve compression-induced injury during chronic implantation (Mackinnon et al. 1984; Larsen et al. 1998), while an oversized cuff will lead to a poor contact between the electrode sites and the target nerve (McNeal and Bowman 1985). Therefore, the Association for the Advancement of Medical Instrumentation (AAMI) recommends that the internal cuff diameter should be 50% larger than the external diameter of the target nerve for safe implantation.

In order to address this issue, Yu et al. developed a self-locking parylene cuff electrode (Yu et al. 2014). This type of cuff electrode is easy to implant and the cuff diameter can be adjusted to fit the nerve properly during implantation through a self-locking structure, as shown in Fig. 32.

As shown in Fig. 33, the microelectrode array is composed of several central working electrode sites and two outer reference electrode sites. The cuff diameter

**Fig. 31** Photographs of the self-closed parylene cuff electrode. (a) Top view; (b) Cross-sectional view

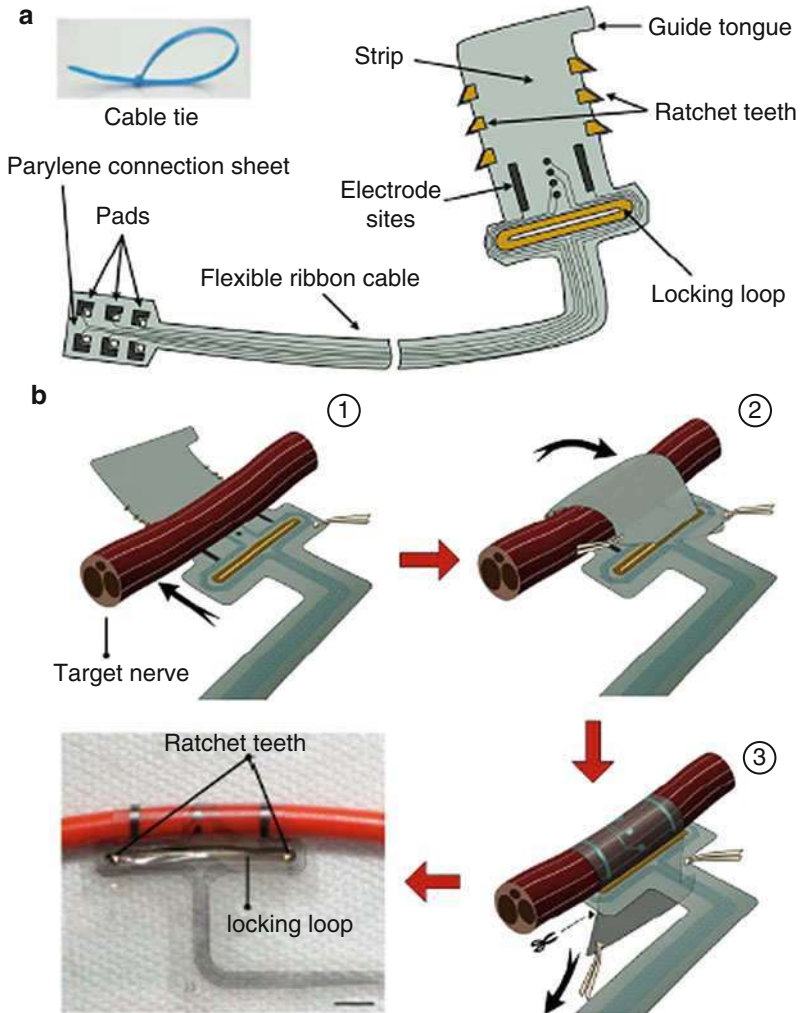


can be adjusted from 1.5 to 2.0 mm with an increment of 0.25 mm. Both the length of the opening in locking loop and the parylene strip are 8 mm, while the length of the projection portion of the ratchet teeth is typically 100  $\mu\text{m}$ . Acute stimulation tests have proved its feasibility of selective stimulation of tibial and peroneal fascicles within the rat sciatic nerve, while chronic implantations (11 weeks) in rats (sciatic nerve) have also demonstrated the safety of these cuff electrodes.

In 2017, Cobo et al. reported a parylene cuff electrode with a similar adjustable locking mechanism. However, these electrodes integrated microfluidics for localized drug delivery (Cobo et al. 2017), as shown in Fig. 34, which could induce axonal sprouting from the fascicles towards the embedded electrodes. The close contact between the axon and the electrode is expected to yield high-quality neural recordings.

Besides the cylinder cuff electrodes described above, there are other types of microelectrode that are also placed on the nerve. For example, Lee et al. reported a microfabricated flexible and adjustable sling electrode (Lee et al. 2016), as shown in Fig. 35. The sling design enables helical electrode distribution with less pressure like the Huntington Helix electrode developed by Agnew et al. at the Huntington Medical Research Institutes (Agnew et al. 1989). The several designed holes have to be sutured to secure the cuff electrodes.

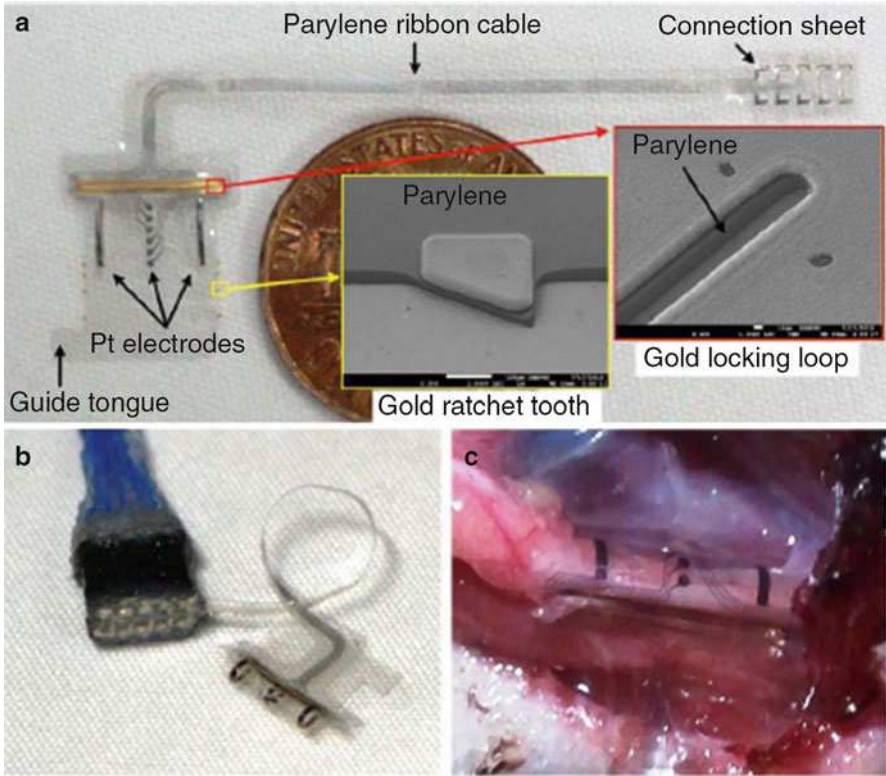
In conclusion, the electrodes implanted on the nerve tend to be lighter, thinner, more flexible, and more biocompatible. The electrodes should cause less damage to the nerve, while maintaining good electrical properties during long-term implantation.



**Fig. 32** Schematic view (not to scale) of (a) the parylene self-locking cuff electrode with flexible parylene cable and (b) the implantation procedure. The photo shows a fabricated cuff electrode wrapped around a hook-up wire (1.3 mm in diameter) and the scale bar is 2 mm

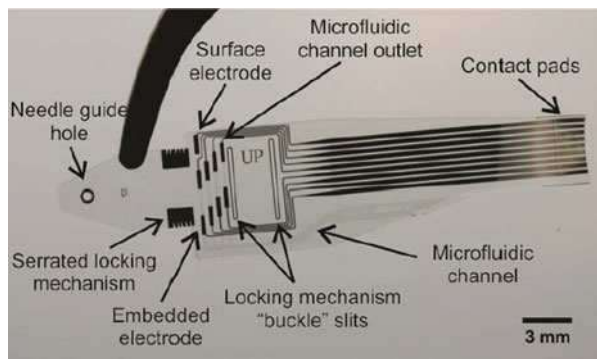
## Perspective

With the development of design and fabrication technology, the diversity and utilization of MEAs are flourishing. Nevertheless, the ultimate purpose is making them usable in practical implantations. There is no permanent criterion to judge the superiority of MEA. Numerous MEAs are being researched, involving improvement and innovative creation. Delightfully, many MEAs are put into clinical trials and

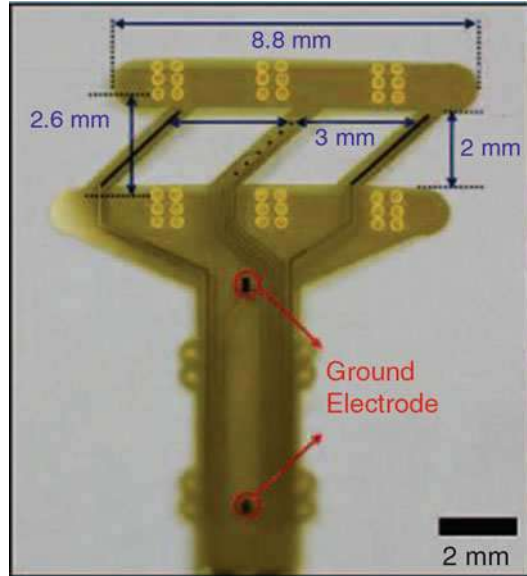


**Fig. 33** Fabrication results: (a) the parylene-based self-locking cuff electrode released from silicon substrate and details of the gold ratchet and locking loop, (b) the packaged cuff electrode with a cuff diameter of 1.7 mm in a closed state, (c) the cuff electrode snugly implanted around the rat sciatic nerve at the mid-thigh. The scale bars in the SEM pictures of the ratchet tooth and locking loop are 100  $\mu\text{m}$  and 10  $\mu\text{m}$ , respectively

**Fig. 34** Photograph of fabricated parylene cuff electrode with microfluidics



**Fig. 35** Pictures of polyimide sling electrode



positive reactions are reported. The development of MEAs would be continued until the electrical neural prosthesis could be substituted for optogenetics or biological repairing such as generic repairing, which seems to take a long time.

## References

- Agnew WF, McCreery DB, Yuen TGH (1989) Histologic and physiologic evaluation of electrically stimulated peripheral nerve: considerations for the selection of parameters. *Ann Biomed Eng* 17(1):39–60
- Bhandari R et al (2008) A novel method of fabricating convoluted shaped electrode arrays for neural and retinal prostheses. *Sens Actuators A Phys* 145–146(1–2):123–130
- Boretius T (2010) A transverse intrafascicular multichannel electrode (TIME) to interface with the peripheral nerve. *Biosens Bioelectron* 26(1):62–69
- Branner A, Normann RA (2000) A multielectrode array for intrafascicular recording and stimulation in sciatic nerve of cats. *Brain Res Bull* 51(4):293–306
- Branner A, Stein RB, Normann RA (2001) Selective stimulation of cat sciatic nerve using an array of varying-length microelectrodes. *J Neurophysiol* 85(4):1585–1594
- Carp JS et al (2005) Long-term spinal reflex studies in awake behaving mice. *J Neurosci Methods* 149(2):134–143
- Castoro MA (2011) Excitation properties of the right cervical vagus nerve in adult dogs. *Exp Neurol* 227(1):62–68
- Cheung KC, Renaud P, Tanila H (2007) Flexible polyimide microelectrode array for in vivo recordings and current source density analysis. *Biosens Bioelectron* 22(8):1783–1790
- Chow AY et al (2004) The artificial silicon retina microchip for the treatment of vision loss from retinitis pigmentosa. *Arch Ophthalmol* 122:460–469
- Chu JU (2012) Improvement of signal-to-interference ratio and signal to noise ratio in nerve cuff electrode systems. *Physiol Meas* 33(6):943–967

- Cobo AM et al (2017) A parylene cuff electrode for peripheral nerve recording and drug delivery. In: IEEE MEMS, pp 506–509
- Cogan SF (2008) Neural stimulation and recording electrodes. *Annu Rev Biomed Eng* 10:275–309
- Crago PE, Peckham PH, Thrope GB (1980) Modulation of muscle force by recruitment during intramuscular stimulation. *IEEE Trans Biomed Eng* 12:679–684
- De Ferrari GM (2011) Chronic vagus nerve stimulation: a new and promising therapeutic approach for chronic heart failure. *Eur Heart J* 32(7):847–855
- Eastwood PR (2011) Treating obstructive sleep apnea with hypoglossal nerve stimulation. *Sleep* 34(11):1479–1486
- Egert D, Peterson RL, Najafi K (2011) Parylene microprobes with engineered stiffness and shape for improved insertion. In: Solid-state sensors, actuators and microsystems conference (Transducers), pp 198–201
- Engineer ND et al (2011) Reversing pathological neural activity using targeted plasticity. *Nature* 470(7332):101–104
- Ethier C et al (2012) Restoration of grasp following paralysis through brain-controlled stimulation of muscles. *Nature* 485:368–371
- Fomani AA, Mansour RR (2010) Flexible neural microelectrode arrays reinforced with embedded metallic micro-needles. In: 2010 I.E. sensors, pp 1601–1604
- Gekeler F et al (2004) Subretinal electrical stimulation of the rabbit retina with acutely implanted electrode arrays. *Graefes Arch Clin Exp Ophthalmol* 242:587–596
- Gekeler F et al (2007) Compound subretinal prostheses with extra-ocular parts designed for human trials: successful long-term implantation in pigs. *Graefes Arch Clin Exp Ophthalmol* 245:230–241
- Glenn WW, Phelps ML (1985) Diaphragm pacing by electrical stimulation of the phrenic nerve. *Neurosurgery* 17(6):974–984
- González C, Rodríguez M (1997) A flexible perforated microelectrode array probe for action potential recording in nerve and muscle tissues. *J Neurosci Methods* 72(2):189–195
- Hassler C, Boretius T, Stieglitz T (2011) Polymers for neural implants. *J Polym Sci B Polym Phys* 49(1):18–33
- Holman G et al (2002) Silicon micro-needles with flexible interconnections. In: IEEE-EMBS, pp 255–260
- Huang R et al (2008) Integrated parylene-cabled silicon probes for neural prosthetics. In: MEMS 2008, pp 240–243
- Humayun MS (2001) Intraocular retinal prosthesis. *Trans Am Ophthalmol Soc* 99:271–300
- Jeong J et al (2013) Advancements in fabrication process of microelectrode array for a retinal prosthesis using liquid crystal polymer (LCP). In: 35th annual international conference of the IEEE EMBS, pp 5295–5298
- Johnson L et al (2004) Electrical stimulation of isolated retina with microwire glass electrodes. *J Neurosci Methods* 137(2):265–273
- Kang X, Liu JQ, Tian H (2015) Self-closed parylene cuff electrode for peripheral nerve recording. *J Microelectromech Syst* 24(2):319–332
- Keseru M et al (2012) Acute electrical stimulation of the human retina with an epiretinal electrode array. *Acta Ophthalmol* 90:1–8
- Kim D et al (2010) Dissolvable films of silk fibroin for ultrathin conformal bio-integrated electronics. *Nat Mater* 9:511–517
- Ko H, Lee S (2017) Electrical characterization of 2D and 3D microelectrodes for achieving high resolution sensing in retinal prostheses with in vitro animal experimental results. *Microsyst Technol* 23(2):473–481
- Larsen JO et al (1998) Degeneration and regeneration in rabbit peripheral nerve with long-term nerve cuff electrode implant: a stereological study of myelinated and unmyelinated axons. *Acta Neuropathol* 96(4):365–378
- Lawrence SM, Dhillon GS, Horch KW (2003) Fabrication and characteristics of an implantable, polymer-based, intrafascicular electrode. *J Neurosci Methods* 131:9–26

- Lee S (2017) Top-down fabrication of silicon nanowire arrays for large scale integration on a flexible substrate for achieving high resolution neural microelectrodes. *Microsyst Technol* 23(2):491–498
- Lee K, Singh A, He J (2004) Polyimide based neural implants with stiffness improvement. *Sens Actuators B Chem* 102(1):67–72
- Lee S, Yen SC, Liao LD (2016) Flexible sling electrode for bidirectional neural signal recording and selective stimulation. In: *IEEE MEMS*, pp 375–378
- Loeb GE (1993) The distal hindlimb musculature of the cat: interanimal variability of locomotor activity and cutaneous reflexes. *Exp Brain Res* 96(1):125–140
- Luo YH-L, da Cruz L (2016) The Argus<sup>®</sup> II retinal prosthesis system. *Prog Retin Eye Res* 50:89–107
- Mackinnon SE et al (1984) Chronic nerve compression – an experimental model in the rat. *Ann Plast Surg* 13(2):112–120
- Margalit E et al (2002) Retinal prosthesis for the blind. *Surv Ophthalmol* 47(4):335–356
- Matzel KE et al (2001) Chronic sacral spinal nerve stimulation for fecal incontinence: longterm results with foramen and cuff electrodes. *Dis Colon Rectum* 44(1):59–66
- McNeal DR, Bowman BR (1985) Selective activation of muscles using peripheral nerve electrodes. *Med Biol Eng Comput* 23(3):249–253
- Motta PS, Judy JW (2005) Multielectrode microprobes for deep-brain stimulation fabricated with a customizable 3-D electroplating process. *IEEE Trans Biomed Eng* 52(5):923–933
- Najafi K, Wise KD, Mochizuki T (1985) A high-yield IC-compatible multichannel recording array. *IEEE Trans Electron Devices* 32(7):1206–1211
- Naples GG et al (1988) A spiral nerve cuff electrode for peripheral nerve stimulation. *IEEE Trans Biomed Eng* 35(11):905–916
- Navarro X, Valderrama E, Stieglitz T (2001) Selective fascicular stimulation of the rat sciatic nerve with multipolar polyimide cuff electrodes. *Restor Neurol Neurosci* 18(1):9–21
- Normann RA et al (1999) A neural interface for a cortical vision prosthesis. *Vis Res* 39(15):2577–2587
- Palanker D et al (2005) Design of a high-resolution optoelectronic retinal prosthesis. *J Neural Eng* 2:S105–S120
- Peckham PH, Knutson JS (2005) Functional electrical stimulation for neuromuscular applications. *Annu Rev Biomed Eng* 7:327–360
- Polasek KH et al (2009) Stimulation stability and selectivity of chronically implanted multicontact nerve cuff electrodes in the human upper extremity. *IEEE Trans Neural Syst Rehabil Eng* 17(5):428–437
- Pratt CA, Chanaud CM, Loeb GE (1991) Functionally complex muscles of the cat hindlimb. *Exp Brain Res* 85(2):281–299
- Randles JEB (1947) Kinetics of rapid electrode reactions. In: *Discuss Faraday SOC*, pp 11–19
- Rijkhoff NJ (1994) Selective stimulation of sacral nerve roots for bladder control: a study by computer modeling. *IEEE Trans Biomed Eng* 41(5):413–424
- Rodger D et al (2008) Flexible parylene-based multielectrode array technology for high-density neural stimulation and recording. *Sens Actuators B Chem* 132(2):449–460
- Rodríguez FJ (2000) Polyimide cuff electrodes for peripheral nerve stimulation. *J Neurosci Methods* 98(2):105–118
- Roessler G et al (2009) Implantation and explantation of a wireless epiretinal retina implant device: observations during the EPIRET3 prospective clinical trial. *Invest Ophthalmol Vis Sci* 50:3003–3008
- Rousche J, Normann A (1999) Chronic intracortical microstimulation (ICMS) of cat sensory cortex using the Utah intracortical electrode array. *IEEE Trans Rehabil Eng* 7(1):56–68
- Rui Y et al (2011) Parylene-based implantable Pt-black coated flexible 3-D hemispherical microelectrode arrays for improved neural interfaces. *Microsyst Technol* 17(3):437–442
- Ryu M et al (2017) Enhancement of interface characteristics of neural probe based on graphene, ZnO nanowires, and conducting polymer PEDOT. *ACS Appl Mater Interfaces* 9:10577–10586

- Struijk JJ et al (1999) Cuff electrodes for long-term recording of natural sensory information. *IEEE Eng Med Biol Mag* 18(3):91–98
- Strumwasser F (1958) Long-term recording from single neurons in brain of unrestrained mammals. *Science* 127(3296):469–470
- Takeuchi S et al (2004) 3D flexible multichannel neural probe array. *J Micromech Microeng* 14(1):104–107
- Tombran-Tink J, Barnstable CJ, Rizzo III JF (2007) VISUAL PROSTHESIS AND OPHTHALMIC DEVICES
- Velliste M et al (2008) Cortical control of a prosthetic arm for self-feeding. *Nature* 453(7198):1098–1101
- Wang R et al (2010) Fabrication and characterization of a parylene-based 3D microelectrode array for use in retinal prosthesis. *J Microelectromech Syst* 19:367–374
- Wang R et al (2012) A flexible microneedle electrode array with solid silicon needles. *J Microelectromech Syst* 21(5):1084–1089
- Wang R et al (2017) A microneedle electrode array on flexible substrate for long-term EEG monitoring. *Sens Actuators B Chem* 244:750–758
- Waschkowski F et al (2014) Development of very large electrode arrays for epiretinal stimulation (VLARS). *Biomed Eng Online* 13:1–15
- Waters RL et al (1985) Functional electrical stimulation of the peroneal nerve for hemiplegia. Long-term clinical follow-up. *J Bone Joint Surg Am* 67(5):792–793
- Weiland JD et al (2006) Implantation of an inactive epiretinal poly (dimethyl siloxane) electrode array in dogs. *Exp Eye Res* 82:81–90
- Wise KD, Angell JB (1975) A low-capacitance multielectrode probe for use in extracellular neurophysiology. *IEEE Trans Biomed Eng* 3:212–219
- Wise KD et al (2004) Wireless implantable microsystems: high-density electronic interfaces to the nervous system. *Proc IEEE* 92(1):76–97
- Yamagiwa S, Ishida M, Kawano T (2013) Self-curling and -sticking flexible substrate for ECOG electrode array. In: *IEEE MEMS*, pp 480–483
- Yoo PB, Durand DM (2005) Selective recording of the canine hypoglossal nerve using a multi-contact flat interface nerve electrode. *IEEE Trans Biomed Eng* 52(8):1461–1469
- Yu H et al (2013) Electroplated nickel multielectrode microprobes with flexible parylene cable for neural recording and stimulation. *J Microelectromech Syst* 22(5):1199–1206
- Yu H et al (2014) A parylene self-locking cuff electrode for peripheral nerve stimulation and recording. *J Microelectromech Syst* 23(5):1025–1035
- Zhang X et al (2011) Characterization of a light switchable microelectrode array for retinal prosthesis. *Appl Phys Lett* 99:253702
- Zhou DD, Greenbaum E (2009) Implantable neural prostheses 1: devices and applications
- Zhou H et al (2009) A new process for fabricating tip-shaped polymer microstructure array with patterned metallic coatings. *Sens Actuators A Phys* 150:296–301
- Zrenner E et al (2011) Subretinal electronic chips allow blind patients to read letters and combine them to words. *Proc Roy Soc B-Biol Sci* 278:1489–1497





# Electrodes for Nerve Recording and Stimulation

Jing-Quan Liu, Hong-Chang Tian, Xiao-Yang Kang, and Ming-Hao Wang

## Contents

Background .....	1414
Flexible Microelectrodes for Recording and Stimulation .....	1416
Materials for Flexible Microelectrodes .....	1417
Research Progress of Flexible Microelectrodes .....	1418
Flexible Microelectrodes for Neural Interface .....	1418
Flexible Microelectrodes with Fluidic Channels .....	1422
Flexible MEMS Microelectrodes with $\mu$ LEDs .....	1424
Research Progress of Materials for Flexible Microelectrodes .....	1426
Future Development Prospect .....	1432
References .....	1432

## Abstract

With the rapid development of MEMS fabrication technologies, versatile microelectrodes with different structures and functions have been designed and fabricated. The flexible MEMS microelectrodes exhibit multiaspect excellent characteristics compared to stiff microelectrodes based on silicon or SU-8, which comprising: *lighter weight, smaller volume, better conforming to neural tissue, and lower fabrication cost.*

This chapter mainly reviewed key technologies on flexible MEMS microelectrodes for neural interface in recent years, including: design and fabrication technology, *fluidic channels,  $\mu$ LEDs*, and electrode-tissue interface modification technology for performance improvement. Furthermore, the future directions of flexible MEMS microelectrodes were described including *transparent and stretchable* microelectrodes with characteristics of *multifunction, high-density,*

J.-Q. Liu (✉) · H.-C. Tian · X.-Y. Kang · M.-H. Wang

Shanghai Jiao Tong University, Shanghai, China

e-mail: [jqliu@sjtu.edu.cn](mailto:jqliu@sjtu.edu.cn); [hchtian@sjtu.edu.cn](mailto:hchtian@sjtu.edu.cn); [xykang@sjtu.edu.cn](mailto:xykang@sjtu.edu.cn); [M.H.Wang89@sjtu.edu.cn](mailto:M.H.Wang89@sjtu.edu.cn)

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,

[https://doi.org/10.1007/978-981-10-5945-2\\_43](https://doi.org/10.1007/978-981-10-5945-2_43)

1413

*biodegradation*, and next-generation electrode-tissue interface modifications facilitated electrode *efficacy* and implantation *safety*.

The goal of this chapter is to provide the reader a broader overview of flexible MEMS technologies that can be applied together to solve problems in neural interface.

---

**Keywords**

MEMS · Microelectrodes · Neural Interface · Conducting Polymer · Nanotechnology

---

## Background

The typical configuration of implantable artificial nerve system is illustrated in Fig. 1. As the tissue-machine interface, the MEMS microelectrodes are interconnected with the biological tissue through the implantable devices, which can transmit electrophysiological signals and coordinate brain, nerve, and muscle orderly. Therefore, the MEMS microelectrodes play an important role in the overall efficiency of whole nerve system.

With the rapid development of microfabrication technologies, tiny biomedical devices with diverse structures can be manufactured, which can minimize the tissue damage during and after implantation for both short-term and long-term requirement (Receveur et al. 2007; Hong et al. 2014). Nowadays, some dense electrode arrays and tenuous electrodes are developed to perform the complex and precise electrophysiological study, which can provide excellent spatial selectivity and low power consumption (Kim et al. 2009; Kozai et al. 2012). Even though these electrodes with smaller size could reduce the damage to the tissue, it would inevitably weaken the performance and safety of the electrodes (Grill et al. 2009). That is because as the size of electrode decrease, its impedance will increase accordingly that induce the drop in charge storage capacity (CSC). As a result, it will generate the poor recording signal and require high current density for stimulation that may damage the tissue. Under this circumstance, the interface materials play a significant role in improving the electrode performance. The ideal materials for electrode-tissue interface should have several properties. First, it should be with electrical property containing low impedance, high CSC, and high charge injection limit. The electrode impedance contributes noise, and higher impedance electrodes are expected to have a lower signal-to-noise ratio. In addition, high electrode impedance in combination with the distributed capacitance between the electrode and the recording amplifier will reduce the electrodes' high-frequency response. Second, it can stabilize for long-term working after implantation without significant property variation. The stability of microelectrodes is important because possible irreversible or harmful electrochemical reactions that might occur during stimulation and glial coatings are always accompanied by the chronic implantations. Third, it should be biocompatible with tissue without inducing severe tissue response, toxicity, or even necrosis. Figure 2 shows the characteristics for ideal implantable electrode-tissue interface.

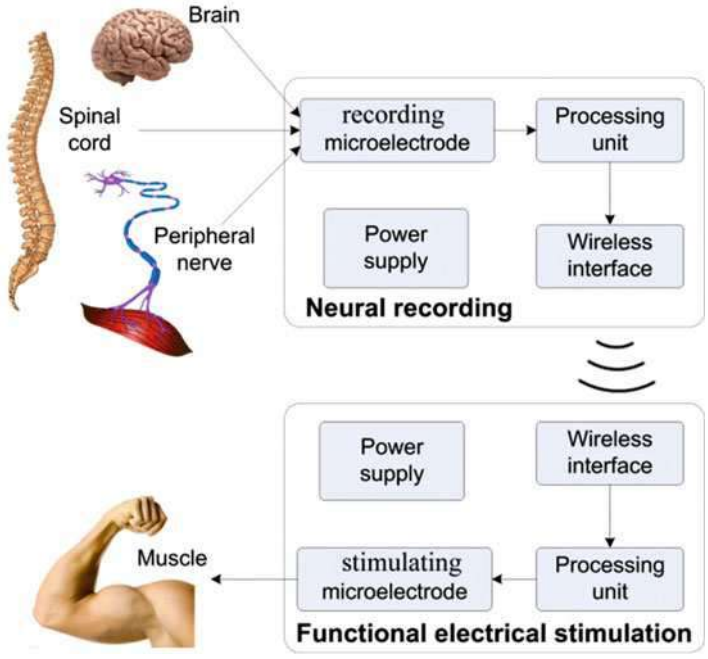
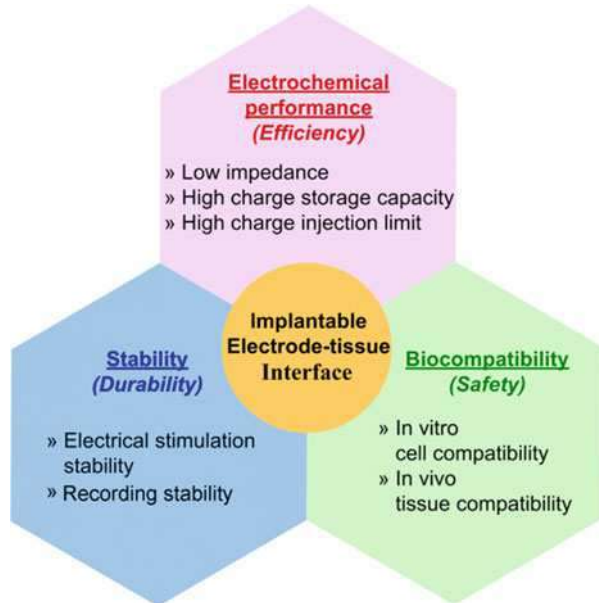


Fig. 1 Structural illustration of implantable artificial nerve system

Fig. 2 Characteristics of ideal implantable electrode-tissue interface



## Flexible Microelectrodes for Recording and Stimulation

One of the most important components for artificial prostheses is the microelectrodes acting as tissue-machine interface (Cogan 2008; Yoshida et al. 2010; Ortiz-Catalan et al. 2012). To better functionalize in living muscle and nerve tissue, the biomedical microelectrodes should meet the following requirements: (1) small dimensions that reduce the tissue damage, (2) excellent performances that ensure the effective operation of the prostheses device, and (3) good biocompatibility that enables relatively long-term implantation without inducing severe immune response. Based on these requirements, various microelectrodes have been developed to perform electrical stimulation and electrophysiological signal recording for paralysis recovery after spinal cord injury. Among these electrodes, neural probes made by University of Michigan and electrodes array made by University of Utah were widely used in central nerve prostheses applications (Hochberg et al. 2006; Wise et al. 2008), while longitudinal intrafascicular (LIFEs) electrodes were commonly applied in peripheral nerve and intramuscular studies (Yoshida et al. 2010; Farina et al. 2008). LIFEs offer a means of interfacing to restricted subsets of axons within fasciculated peripheral nerves and they are constructed from thin insulated conducting wires, such as Pt-Ir or metallized Kevlar fibers (Navarro et al. 2005). Cuff electrodes are composed of an insulating tubular sheath that completely encircles the nerve and contains two or more electrode contacts exposed at their inner surface that are connected to insulated lead wires. Cuff electrodes are extraneural electrodes which can provide simultaneous interface with many axons in the nerve, whereas intrafascicular and sieve electrodes inserted in the nerve may interface small groups of axons within a nerve fascicle.

There is a significant symptom if only the electrical interaction between electrodes and muscle/nerve tissue were performed without nutrition factor delivery that would eventually lead to denervation-induced skeletal muscle atrophy (Mitch and Goldberg 1996; Thomas et al. 1997; Midrio 2006). In recent years, the microelectrodes integrated with microchannels for fluidic drug delivery were developed to solve this problem (Jessin et al. 2011; Altuna et al. 2013; Pongrácz et al. 2013). Moreover, microfluidic neural interfaces hold immense potential for basic neuroscience research and clinical medicine (Jeong et al. 2015). *In vivo* neuropharmacology enables delivery of pharmacological agents deep into the brain to help dissect complex neural circuits and neurotransmitter/receptor systems (Anthony et al. 2014) and treat neurodegenerative diseases and brain tumors. However, most studies focused on stiff electrodes made of silicon or SU-8 for applications on central nerve system, only a few flexible electrodes made of parylene, polyimide (PI), and polydimethylsiloxane (PDMS) were proposed (Metz et al. 2004; Takeuchi et al. 2005; Gao et al. 2013). And, the problem which limits the precise stimulation with the polymer-based flexible electrodes mentioned above is the electrode sites distributed on one side of the electrode only. Moreover, most of the microelectrodes mentioned above were developed for neural applications, and few electrodes were designed for intramuscular research. Recently, the microelectrodes for intramuscular electrical stimulation and electromyogram (EMG) recording were generally

restricted to crude wire electrodes with simple construction and single function (Jarc et al. 2013; Memberg et al. 2014). For the current situations, it is necessary to design and fabricate the multifunctional microelectrodes with circumferential distributed electrode sites for intramuscular prostheses. However, its limitations such as unpredictable current pathways, electrical artifacts, and nonselectivity of the target neurons increase the demands for a new technology. Direct optical stimulation of neural cells in the brain tissue, genetically modified by expressing channelrhodopsin-2 (ChR2), has drawn much attention over the past few years. This technique, called optogenetics, can target specific types of neurons with submillisecond temporal precision. For *in vitro* and head-fixed *in vivo* applications, numerous optical stimulation systems have been reported such as a laser-coupled optical fiber (Zhang et al. 2007), light-emitting diodes (LEDs), band-filtered white light, and a focused laser beam through a microscope (Hira et al. 2009). For experiments with freely behaving subjects, however, only limited light delivery methods such as a laser-coupled optical fiber and a head-mountable single LED system (Iwai et al. 2011) are available because the tethered optical fiber restricts a natural behavior of the subjects.

## Materials for Flexible Microelectrodes

In order to enhance the electrode performance of implantable MEMS microelectrodes, the ideal materials for implantable electrode-tissue interface should satisfy requirements of excellent electrochemical performance, high stability, and well biocompatibility. For neural stimulation electrodes, possible irreversible or harmful electrochemical reactions might occur during stimulation. *In vivo* studies have identified Pt dissolution (Robblee et al. 1983) and iridium oxide delamination (Cogan et al. 2004) as electrode degradation processes. Thus, the materials for stimulation electrodes should be stable without corrosion under *in vivo* implantation. In general, electrodes used in neural stimulation can be divided into two categories. Macroelectrodes exhibit high CSC and low charge density; they are typically placed on the surface of the target tissue and have a GSA larger than approximately  $100,000 \mu\text{m}^2$ . Microelectrodes have the opposite behavior, exhibiting low CSC and high charge density. Microelectrodes typically penetrate the target and have surface areas that are smaller than approximately  $10,000 \mu\text{m}^2$ . An obvious advantage of microelectrodes is the ability to stimulate a comparatively small volume of tissue, which should, with a sufficient number of electrodes, improve the selectivity and spatial resolution of functional responses. For neural recording electrodes, action potentials are recorded in close proximity to the target neurons, and, for single-unit recording, the microelectrode GSA should be no larger than approximately  $2000\text{--}4000 \mu\text{m}^2$ . The objective with single-unit neural recording is to measure action potentials with a useful signal-to-noise ratio,  $\sim 3:1$  or greater, and to do this chronically. The amplitude of action potentials in the CNS can be quite large, more than  $500 \mu\text{V}$ , but is more typically on the order of  $100 \mu\text{V}$ , and often is smaller. In general, the majority of the noise signal encountered in single-unit recording arises from the

multitude of undifferentiated background action potentials (neural noise). However, electrode impedance does contribute noise, and higher impedance electrodes are expected to have a lower signal-to-noise ratio. Thus, the materials for recording electrodes should have low impedance.

Even though various materials for electrode-tissue interface have been developed, the most widely used materials remain noble metals, such as platinum, gold, iridium, tungsten, and their alloys. These metals are chosen to be the materials for electrode-tissue interface due to their excellent chemical stability without serious erosion after implantation. However, the electrochemical performance of MEMS microelectrodes are restricted mostly by their high electrochemical impedance and low CSC of bare metallic electrode-tissue interface. Some electrode-tissue interfaces were processed with porous structure in a special profile to form a rough surface, which would consequently improve the effective surface area of electrode sites. In recent years, emerging carbon nanomaterials, such as carbon nanotubes and grapheme, etc., act as electrode-tissue interface to improve MEMS microelectrode performance for their multispect excellent properties. Although the carbon nanomaterials possess the advantages of extremely large specific area and excellent electrochemical performance, they have the poor bonding with electrode substrate and may induce nanotoxicity by litters in tissue that will highly limit their applications.

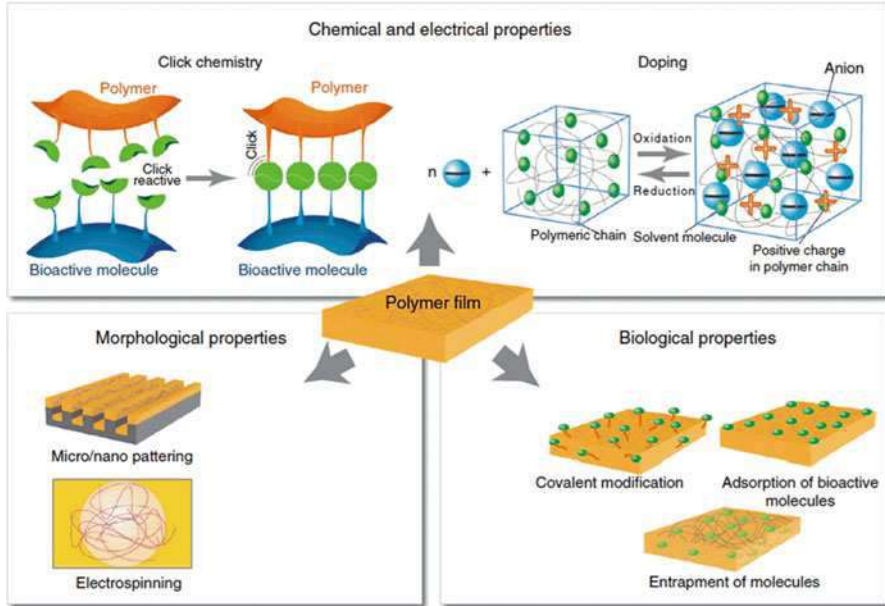
Conducting polymers have attracted much attention and broadly applied in different biomedical research due to their unique characteristics of low electrochemical impedance, high CSC, favorable plasticity, electrostriction, and biocompatibility. Meanwhile, conducting polymers are capable to meet the requirements of materials for electro-tissue interface. In addition, conducting polymers also possess the other characteristics, including modification through doping, electrically controlled drug releasing, molding through micro-nano processing and electro-spinning, and surface modification by biochemical molecules, as shown in Fig. 3 (Vallejo-Giraldo and Biggs 2014). As two kinds of conducting polymers that are widely used as electrode-tissue interface, poly (3,4-ethylenedioxythiophene) (PEDOT) exhibits better performance than polypyrrole (PPy) in terms of electrical stimulation (better electrochemical performance) and cell culture (longer neurites growth) (Vallejo-Giraldo and Biggs 2014).

---

## Research Progress of Flexible Microelectrodes

### Flexible Microelectrodes for Neural Interface

With the rapid development of MEMS fabrication technologies, researchers have developed various kinds of biomedical microelectrodes applied on electrical stimulation and electrophysiological signal recording for paralysis recovery. Among these microelectrodes, neural probes University of Michigan and electrodes array made by University of Utah are widely utilized in central nerve system studies as stiff MEMS microelectrodes for neural prosthesis (Hochberg et al. 2006; Wise et al. 2008). Many



**Fig. 3** Schematic illustration of characteristics of conducting polymers (Vallejo-Giraldo and Biggs 2014)

research efforts have been devoted to developing novel flexible MEMS microelectrodes due to their excellent characteristics compared with stiff microelectrodes, such as lighter weight, smaller volume, better conforming to neural tissue, and lower fabrication cost (Kozai et al. 2012; Ferguson et al. 2009; Kim et al. 2010; Rui et al. 2011). Owing to these advantages, the flexible MEMS microelectrodes for neural interface have attracted extensive attentions and considered to have broad prospects in future. As described below, the flexible MEMS microelectrodes can be classified into three types: wire electrode, thin film electrode, and mesh electrode. As shown in Table 1, the comparison results of different kinds of the flexible MEMS microelectrodes for neural interface are discussed in this chapter.

John E. Ferguson and A. David Redish et al. from University of Minnesota developed a tetrode made from four microwire electrodes for neural signal recording on living animals in 2009 (Ferguson et al. 2009). As shown in Fig. 4a, the tetrode was composed of four Ni-Cr alloy microwires with diameter of 12.7  $\mu\text{m}$ , which are coated with polyimide as insulation on the surface. The microwire electrodes were cut to expose the cross section as electrode sites, and the electrode sites were electrodeposited with gold to improve their electrochemical performance. The microwire electrode was easy to be fabricated, and the electrode with micro-dimension was suitable for cortical implantation with little tissue damage. However, the microwire electrode was inconvenient to be implanted into target position at deep brain area precisely.



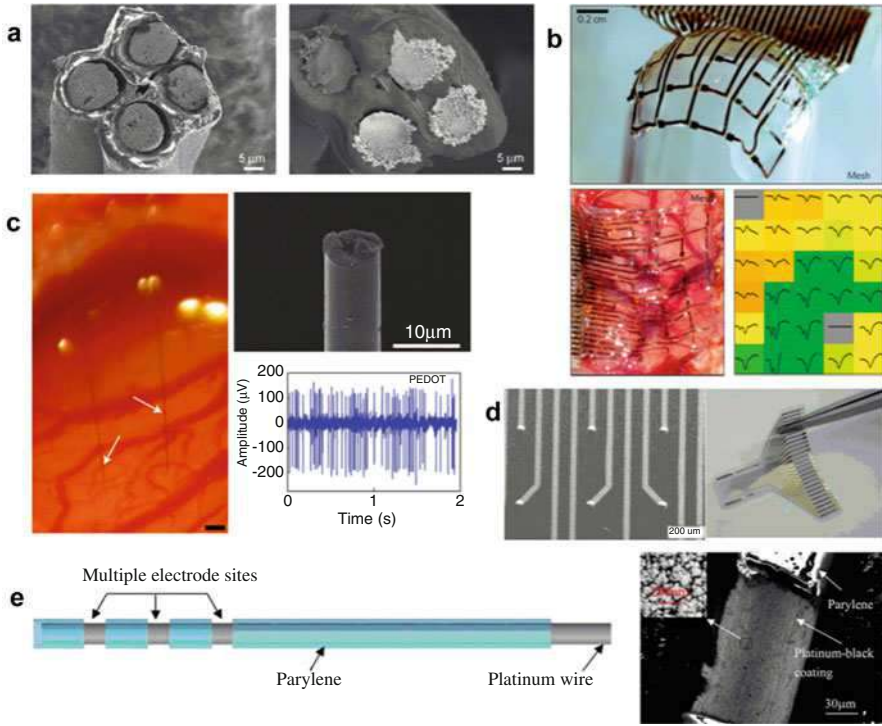
**Table 1** Flexible MEMS microelectrodes for neural interface

Type	Materials		Dimension of electrode site	Fluidic channels	References
	Conduction	Insulation			
Wire electrode	Ni/Cr	Polyimide	12.7 $\mu\text{m}$ of diameter	Without	Ferguson et al. (2009)
Mesh electrode	Au	Polyimide/fibroin	500 $\mu\text{m}$ $\times$ 500 $\mu\text{m}$	Without	Kim et al. (2010)
Wire electrode	Carbon	Poly-p-xylene	7 $\mu\text{m}$ of diameter	Without	Kozai et al. (2012)
Thin film electrode	Pt	Parylene	50 $\mu\text{m}$ of diameter	Without	Rui et al. (2011)
Wire electrode	Pt	Parylene	100 $\mu\text{m}$ of diameter	Without	Rui et al. (2012)
Thin film electrode	Pt	Polyimide	50 $\mu\text{m}$ $\times$ 50 $\mu\text{m}$	With	Metz et al. (2004)
Thin film electrode	Au	Parylene	40 $\mu\text{m}$ $\times$ 40 $\mu\text{m}$	With	Ziegler and Takeuchi (2006)
Thin film electrode	Pt	PDMS/polyimide	–	With	Gao et al. (2013)
Wire electrode	Au	Parylene/Teflon	100 $\mu\text{m}$ of diameter	With	Tian et al. (2014a)

Dae-Hyeong Kim and John A. Rogers et al. from University of Illinois at Urbana-Champaign developed a thin film microelectrode array based on polyimide for electrocorticogram (ECoG) recording in 2010. The microelectrode array was reinforced by silk fibroin, which functioned as biodegradable substrate, to improve the conformal attachment on the brain tissue surface (Kim et al. 2010). As displayed in Fig. 4b, the thickness of 5  $\times$  6 grid-like thin film microelectrode array was approximately 2.5  $\mu\text{m}$ , and the area of the electrode site was 500  $\mu\text{m}$   $\times$  500  $\mu\text{m}$ . In addition, as shown in Fig. 4b, the thin film microelectrode array could be attached on the sphere surface tightly, when the silk fibroin dissolved. The biodegradable surface coating with silk fibroin facilitated the conformal cover of the mesh electrode on the rough surface of brain. Moreover, the mesh electrode could be fabricated thinner and the electrode sites could be designed smaller to further improve the conformal attachment on brain and accuracy of neural recording.

Takashi D. Yoshida Kozai, Nicholas A. Kotov, and Daryl R. Kipke et al. from University of Michigan developed composite fiber electrodes consisting of carbon fiber core and poly-p-xylene insulation coating in 2012 (Kozai et al. 2012). As shown in Fig. 4c, the diameter of carbon fiber core was 7  $\mu\text{m}$ . The poly-p-xylene insulation layer with thickness of 800 nm was coated on the carbon fiber surface by chemical vapor deposition. Moreover, conducting polymer was electrochemically deposited on its cross section to improve the electrochemical performance. The ultra-small dimension facilitated the penetration of microelectrode into brain tissue and induced little tissue damage. In addition, the mechanical property of the carbon fiber microelectrodes was well suited for the inherent mechanical property of brain tissue. The carbon fiber electrode could be further fabricated into multiple channels to realize the complex neural recording.





**Fig. 4** Research progress of flexible MEMS microelectrodes for neural interface. (a) Tetrode composed of four microwire electrodes (Ferguson et al. 2009); (b) Thin film microelectrode array with silk fibroin covered as substrate (Kim et al. 2010); (c) Carbon fiber microelectrode (Kozai et al. 2012); (d) Thin film microelectrode with 3D raised hemispherical electrode sites (Rui et al. 2011); (e) Microwire electrode with multiple cylindrical electrode sites (Rui et al. 2012)

Yuefeng Rui and Jingquan Liu et al. from Shanghai Jiao Tong University developed the flexible 3D microelectrode array with raised hemispherical electrode sites in 2011 (Rui et al. 2011). The electrode sites with diameter of  $50\ \mu\text{m}$  were arranged in  $5 \times 5$  array, and the gap between two adjacent electrodes was  $600\ \mu\text{m}$ . As demonstrated in Fig. 4d, compared with flat electrode sites, the microscale 3D hemispherical electrode sites facilitated the contact with nerve tissue by increasing the effective contact area and reducing the interfacial resistance, which could improve the electrical stimulation and neural signal recording performance. In addition, the research group developed flexible microwire electrodes in 2012 (Rui et al. 2012). As shown in Fig. 4e, compared with flat microelectrode array, the microwire electrodes could not only facilitate the implantation process and reduce the tissue damage but also be arbitrarily bended to adapt specific circumstances. The cylindrical electrode sites of microwire electrode contacted well with the bioactive tissue, thus the impedance per unit area was decreased. Furthermore, the electrode sites with diameter of  $100\ \mu\text{m}$  were electrodeposited with platinum black by ultrasonic current pulses to improve their electrochemical performance. Compared with

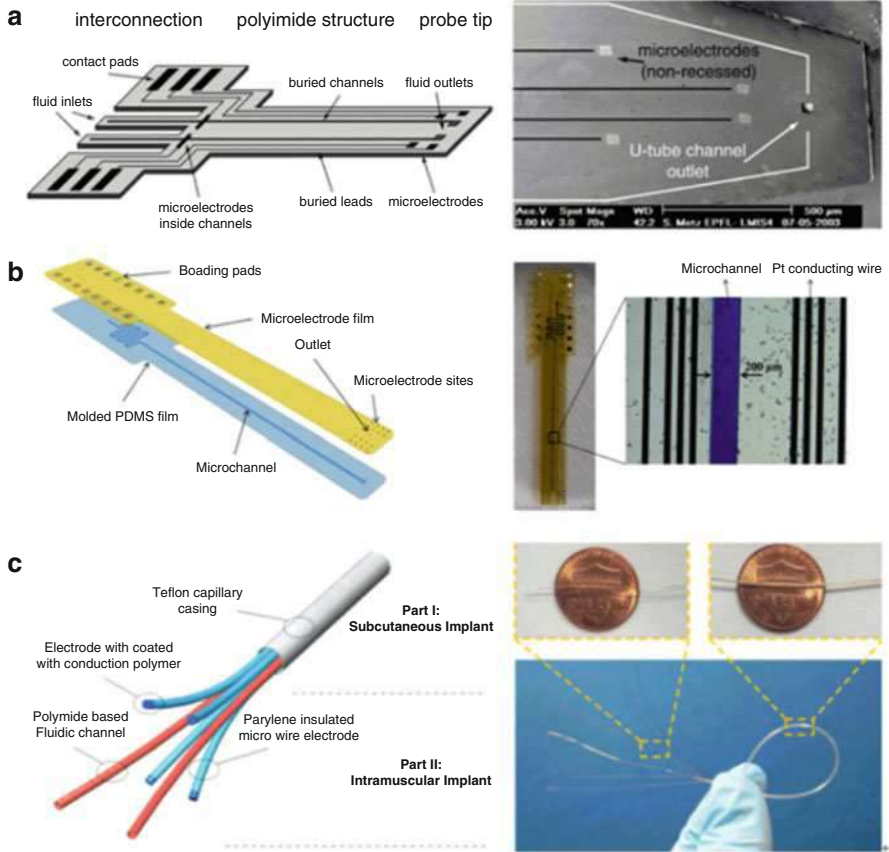
conventional microwire electrode, the cylindrical electrode sites facilitated the tight attachment on the nerve and muscle tissue. The area of the electrode sites could be further reduced to improve the spatial selectivity.

## Flexible Microelectrodes with Fluidic Channels

While the researchers focus on developing MEMS microelectrodes with smaller dimension, more complex structure and denser electrode sites distribution, they are also devoting many efforts to multifunctionalizing the MEMS microelectrodes. The nerve conduction and muscle contraction actions of denervated paralyzed nerve and muscle tissue could be restored by electrical stimulation based on MEMS microelectrodes in artificial neural system (Yoshida et al. 2010; Farina et al. 2008). However, long-term lack of neural nutritional factors would eventually lead to denervation atrophy of nerve and muscle tissue, which would result in irreversible loss of natural conduction of nerve system and contraction function of muscle (Mitch and Goldberg 1996; Thomas et al. 1997; Midrio 2006). Therefore, some MEMS microelectrodes integrated with fluidic channels have been developed for drug delivery (Metz et al. 2004; Gao et al. 2013; Ziegler and Takeuchi 2006; Tian et al. 2014a, 2015). Based on the inherent electrical stimulation and electrophysiological signal recording properties, the novel MEMS microelectrodes can also be utilized to deliver fluidic drugs, nutritional factors, and neural transmitters to target nerve and muscle tissue sites. The flexible MEMS microelectrodes with fluidic channels were mainly made from polyimide, PDMS, and parylene.

S. Metz and A. Bertsch et al. from École polytechnique fédérale de Lausanne developed flexible MEMS microelectrodes with fluidic channels based on polyimide (PI) in 2004 (Metz et al. 2004). As displayed in Fig. 5a, the thickness of microelectrode array was 10 ~ 60  $\mu\text{m}$ , and the dimension of electrode sites was 50  $\mu\text{m}$   $\times$  50  $\mu\text{m}$ . The cross-section dimension of inner fluidic channels was 5  $\mu\text{m}$   $\times$  50  $\mu\text{m}$  or 20  $\mu\text{m}$   $\times$  200  $\mu\text{m}$ , and the cross-section dimension of fluidic channel exits was 30  $\mu\text{m}$   $\times$  30  $\mu\text{m}$  or 50  $\mu\text{m}$   $\times$  50  $\mu\text{m}$ . The thickness of PI microelectrode array was relatively small, which was suitable for cortical implantation and neural recording. However, the one-sided distribution of electrode sites affected the functional scope of neural recording and stimulation.

Shoji Takeuchi and Dominik Ziegler et al. from Institute of Industrial Science and University of Tokyo developed flexible MEMS microelectrodes with fluidic channels based on parylene in 2006 (Ziegler and Takeuchi 2006). The thickness of microelectrode array was 18  $\mu\text{m}$ , and the dimension of electrode sites was 40  $\mu\text{m}$   $\times$  40  $\mu\text{m}$ . The cross-section dimension of inner fluidic channels was 15  $\mu\text{m}$   $\times$  80  $\mu\text{m}$ , and the cross-section dimension of fluidic channel exits was 100  $\mu\text{m}$   $\times$  100  $\mu\text{m}$ . The parylene microelectrode array was very thin, which could facilitate the conformal cover on brain cortex. However, the parylene thin film electrode was too thin to withstand the internal pressure produced by tissue motion, which might lead to the closure of the fluidic channel and disability of drug delivery.



**Fig. 5** Research progress of flexible MEMS microelectrodes with fluidic channels for neural interface. (a) Thin film microelectrode array based on polyimide (PI) (Metz et al. 2004); (b) Thin film microelectrode array based on poly-dimethylsiloxane (PDMS) (Gao et al. 2013); (c) microelectrodes-integrated polyimide microfluidic channels and parylene microwire electrodes (Tian et al. 2015)

Kunpeng Gao and Gang Li et al. from Shanghai Institute of Micro-System and Information Technology (Chinese Academy of Sciences) developed flexible MEMS microelectrodes with fluidic channels based on poly-dimethylsiloxane (PDMS) in 2013 (Gao et al. 2013). As shown in Fig. 5b, the thickness of the microelectrode array was 125  $\mu\text{m}$ . The cross-section dimension of inner fluidic channels was 50  $\mu\text{m} \times 200 \mu\text{m}$ . The thickness of the PDMS microelectrode array was relatively large, since it was unsuitable for conformal attachment on tissue with high flexibility.

Hongchang Tian and Jingquan Liu et al. from Shanghai Jiao Tong University developed flexible MEMS microelectrodes integrated with polyimide microfluidic channels and parylene microwire electrodes in 2014 (Tian et al. 2014a, 2015). As shown in Fig. 5c, the integrated flexible microelectrode was composed of three parts: (1) the microwire electrodes as electrical interfaces, (2) the PI capillaries (outer

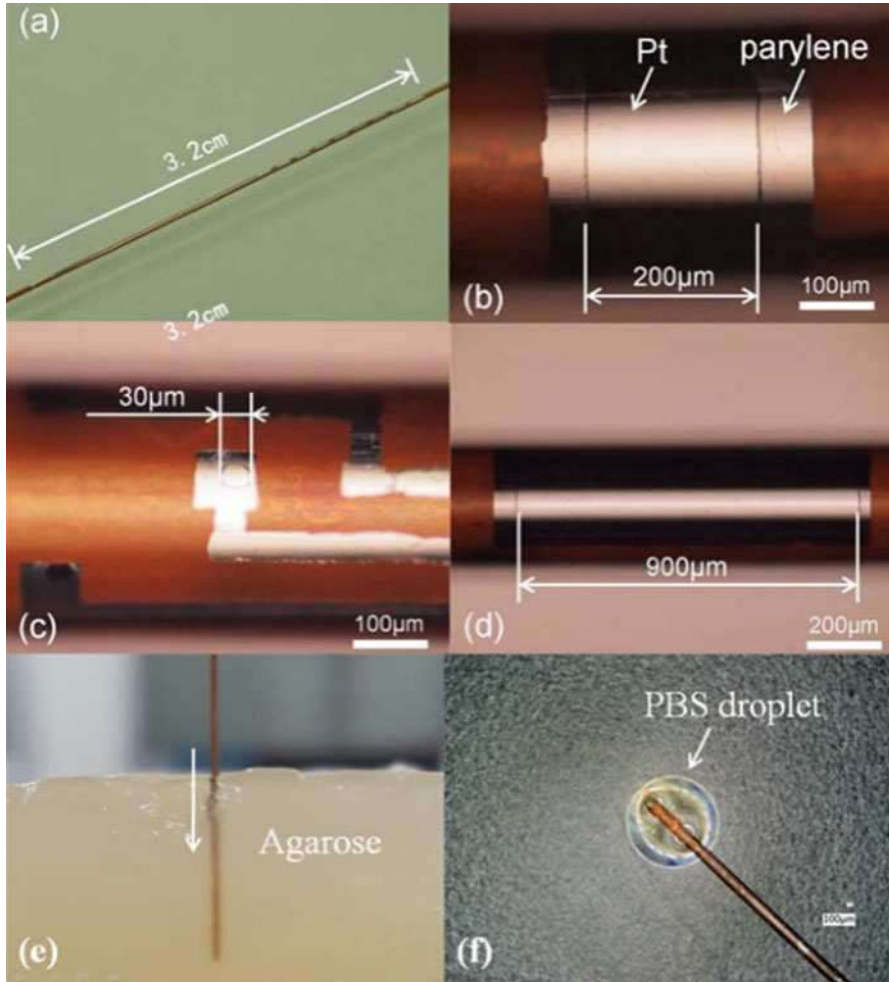
diameter of 110  $\mu\text{m}$  and wall thickness of 10  $\mu\text{m}$ ) for fluidic drug releasing as chemical interfaces, and (3) the Teflon capillary (outer diameter of 650  $\mu\text{m}$  and wall thickness of 140  $\mu\text{m}$ ) for packaging. The integrated microelectrode with drug delivery function was easy to be fabricated and change its parameters. More electrode sites and fluidic channels with smaller dimensions were required to realize more complex and precise neural recording and stimulation.

Minghao Wang and Jingquan Liu et al. of Shanghai Jiao Tong University reports a novel flexible neural probe fabricated by cylindrical substrates lithography for Brain-Computer-Interface (BCI) applications (Wang et al. 2017). The electrode sites were patterned on cylindrical surface to acquire high space selectivity and the microchannel was integrated to deliver drugs or optical stimulation. The unique cylindrical substrates lithography has been reported to have high pattern resolutions ( $\pm 1$   $\mu\text{m}$  alignment precision) and high reliability (Yang et al. 2014). Using this technology, postcrimping process of the substrate is not needed compared to the plane lithography. Therefore, this method was more time-saving and reliable. Fig. 6a–d shows the photographs of the whole design of the flexible probe. The probe has an outer diameter of 330  $\mu\text{m}$  and an inner diameter of 250  $\mu\text{m}$  and the total length is 3.2 cm. There are eight electrode sites with a diameter of 30  $\mu\text{m}$  and a counter electrode site with a width of 200  $\mu\text{m}$  distributed on the surface of the probe. Figure 6e, f shows the assembled probe inserted into agarose and delivering drug. The drug-delivering ability makes the probe suitable for chronic implantation by releasing anti-inflammatory agent or nerve growth factor (NGF).

## Flexible MEMS Microelectrodes with $\mu\text{LEDs}$

Attribute to the development of MEMS technology, microelectrocorticography ( $\mu\text{ECoG}$ ) electrodes can record from large-area cortical surface field potentials with high-density electrodes at mesoscopic scales. On the other hand, optogenetics has gained substantial interest over the last 10 years, which can excite or inhibit a specific neuron type with expression of light sensitive ion channels or pumps. It provides an ideal possibility to manipulate specific neural circuits with the combination of  $\mu\text{ECoG}$  electrodes. As the stimulation light source,  $\mu\text{LED}$  is a superior choice than laser, whereas only few researches on  $\mu\text{ECoG}$  are reported using  $\mu\text{LEDs}$ . Thus, it is urgent to develop a novel photoelectric neural interface combining  $\mu\text{LEDs}$  and  $\mu\text{ECoG}$  electrodes for precise simultaneous optical stimulation and recording.

Ki Yong Kwon and Wen Li et al. of Michigan State University presents a wireless-enabled, flexible optrode array with multichannel microlight-emitting diodes ( $\mu\text{LED}$ ) for bidirectional wireless neural interface (Kwon et al. 2014). The array integrates wirelessly addressable  $\mu\text{LED}$  chips with a slanted polymer optrode array for precise light delivery and neural recording at multiple cortical layers simultaneously. The wirelessly powered-slanted optrode array contains 32 embedded  $\mu\text{LED}$  light sources on a  $2.5 \times 2.5$   $\text{mm}^2$  flexible substrate, with  $4 \times 4$  channels per each hemisphere to cover both visual cortices in rats. Integrated LED light sources powered by the wireless switched-capacitor stimulator (SCS) enable a truly untethered system.

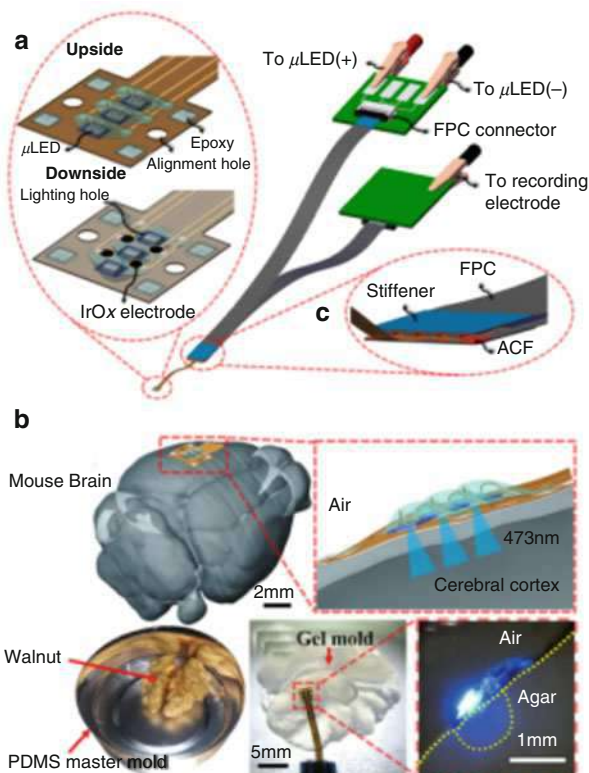


**Fig. 6** The photographs of (a) the fabricated neural probe and (b) the counter electrode site, (c) the electrode site and (d) the bonding pad of the probe. (e) The probe inserted to agarose and (f) delivering drug (Wang et al. 2017; Yang et al. 2014)

Bowen Ji and Jingquan Liu et al. of Shanghai Jiao Tong University developed a novel integrated  $\mu$ LED- $\mu$ ECoG neural interface using wire-bonding technology combining with iridium oxide (IrOx) microelectrodes (Ji et al. 2017). An overall schematic of the  $\mu$ LED microelectrodes was shown in Fig. 7a, with a  $2 \times 2 \text{ mm}^2$  footprint. The three bare  $\mu$ LED chips (C460TR2227-0328, Cree Inc., USA), with  $220 \times 270 \times 50 \text{ }\mu\text{m}^3$  in dimension and 460 nm in peak wavelength, were arranged in a line with luminous surface down and aligned above the holes on PI substrate to allow the light propagation without obstacle. As can be seen from the downside view, four electrode sites with a diameter of 200  $\mu\text{m}$  were modified with IrOx and



**Fig. 7** (a) Concept diagram of the  $\mu$ LEDs microelectrode. (b) Illustration of the  $\mu$ LEDs microelectrode attaching on the translucent agar gel mold (Ji et al. 2017)



distributed around three  $\mu$ LEDs. These two subarrays were individually connected to flexible printed circuit (FPC) with anisotropic conductive film (ACF, AC2056R, HITACHI, Japan) using a pulse hot-pressing machine. The illustration of  $\mu$ LED microelectrodes attaching on the surface of mouse cerebral cortex (model from Allen Brain Atlas) is shown in Fig. 7b. To assess the attachment effects, a walnut-shaped agar gel mold from PDMS was applied to mimic mouse cerebral cortex. The  $\mu$ LED microelectrodes can be attached on the cortical surface of a mouse expressing ChR2 to realize synchronized light modulating and neural signal recording.

## Research Progress of Materials for Flexible Microelectrodes

Functional interfaces should have the ability to combine different physical stimulations, such as electrical (Radisic et al. 2004), magnetic (Dobson 2008), mechanical (Svennersten et al. 2011), or optical (Wells et al. 2005), for further exploration and manipulation of stimulus-sensitive cell with modified substrates. Majority of existing studies have already incorporated conductive biomaterials functioning as electrode-tissue interface in neural engineering studies (Tandon et al. 2009; Cho and Borgens 2013; Martins et al. 2013). Park et al. and Huang et al. reported that the

differentiation and maturity of neural stem cells could be promoted by electrical stimulation on graphene plate and carbon nanotube rope, respectively (Park et al. 2011; Huang et al. 2012). Zhao reported the effects of skeletal myogenesis on comb pattern substrate adjusted by different frequency of electrical stimulating (Zhao 2009). Lee and his colleagues demonstrated that polypyrrole coated nanofibres were able to induce directional growth of neurons (Lee et al. 2009). These researches suggested that conducting substrates with variable structure had great potential on modulating excitable cells in neural engineering.

Conducting polymer possesses various excellent characteristics including high CSC, low impedance, excellent plasticity, and volume electrostrictive effect (Abidian and Martin 2009; Poole-Warren et al. 2010; Yoon and Jang 2009). Meanwhile, due to the good biocompatibility, conducting polymers is widely applied in various biomedical area such as biomedical imaging (Au et al. 2013), biosensor (Yang et al. 2011; Arter et al. 2010), artificial muscle (Plesse et al. 2010), drug release controller (Simon et al. 2009), cancer biomarker (Bangar et al. 2009), and neural interface (Abidian et al. 2009; Asplund et al. 2010). One of the most important application of conducting polymer is using as the electrode-tissue interface material in neural engineering, which can be facilely fabricated into multiple structures (Gomez et al. 2007), modified by different doping (Abidian et al. 2012; Bongo et al. 2013) and regulated to undertake electrical stimulation (Hsiao et al. 2013; Quigley et al. 2012).

The fundamental properties of conducting polymers including surface morphology, electrical stimulating performance, stability, and biocompatibility highly depend on the characteristics of negatively charged dopants which are also termed as counterions. For instance, macromolecules with high mechanical properties have the ability to enhance the stability of conducting polymer composites (Luo et al. 2011). Similarly, nanomaterials with excellent conductivity, such as carbon nanotubes, are capable of improving the electrical performance of composite film (Albahrani et al. 2015). Therefore, it is very important to assess the influence on doping different counterion components into the conducting polymer. As one of the frequently used conducting polymer for interfaces to different cells, polypyrrole (PPy) was doped with generally accessible molecules as electrode-tissue interface. The conducting polymer electrode-tissue interface mainly referred to PEDOT, which was combined with water soluble molecules, biomolecules, carbon nanotube (CNT), perchlorate ( $\text{ClO}_4^-$ ) oxide groups, and graphene oxide (GO). Different kinds of the conducting polymer (PEDOT) electrode-tissue interface for neural interface are discussed in this chapter and presented and reviewed in Table 2 for comparison.

Xinyan Cui and David C. Martin et al. from University of Michigan doped polystyrenesulfonate (PSS) as negatively charged counter ion into PEDOT to form PEDOT/PSS composite as electrode-tissue interface (Cui and Martin 2003). As exhibited in Fig. 8a, since the rough and porous structure of PEDOT/PSS composite greatly increased the effective area of electrode-tissue interface, it would improve the electrochemical performance. Therefore, the PEDOT/PSS composite became one of the most widely used conducting polymer materials.

Maria Asplund and Hans von Holst et al. from Royal Institute of Technology separately doped hyaluronic acid (HA), heparin, and fibrinogen as negatively charged

**Table 2** Conducting polymer (PEDOT) electrode-tissue interface for neural interface

Dopant	Structure	Dopant type	References
Polystyrenesulfonate (PSS)	Film	Soluble molecule	Hsiao et al. (2013)
Hyaluronic acid (HA)	Film	Biomolecule	Quigley et al. (2012)
Heparin	Film	Biomolecule	Quigley et al. (2012)
Fibrinogen	Film	Biomolecule	Quigley et al. (2012)
Carbon nanotube (CNT)	Porous	Carbon nanomaterial	Luo et al. (2011)
Perchlorate ( $\text{ClO}_4^-$ )	Nanotube	Negatively charged ion	Al-bahrani et al. (2015)
Graphene oxide (GO)	Film	Carbon nanomaterial	Cui and Martin (2003), Asplund et al. (2008), Tian et al. (2014b)

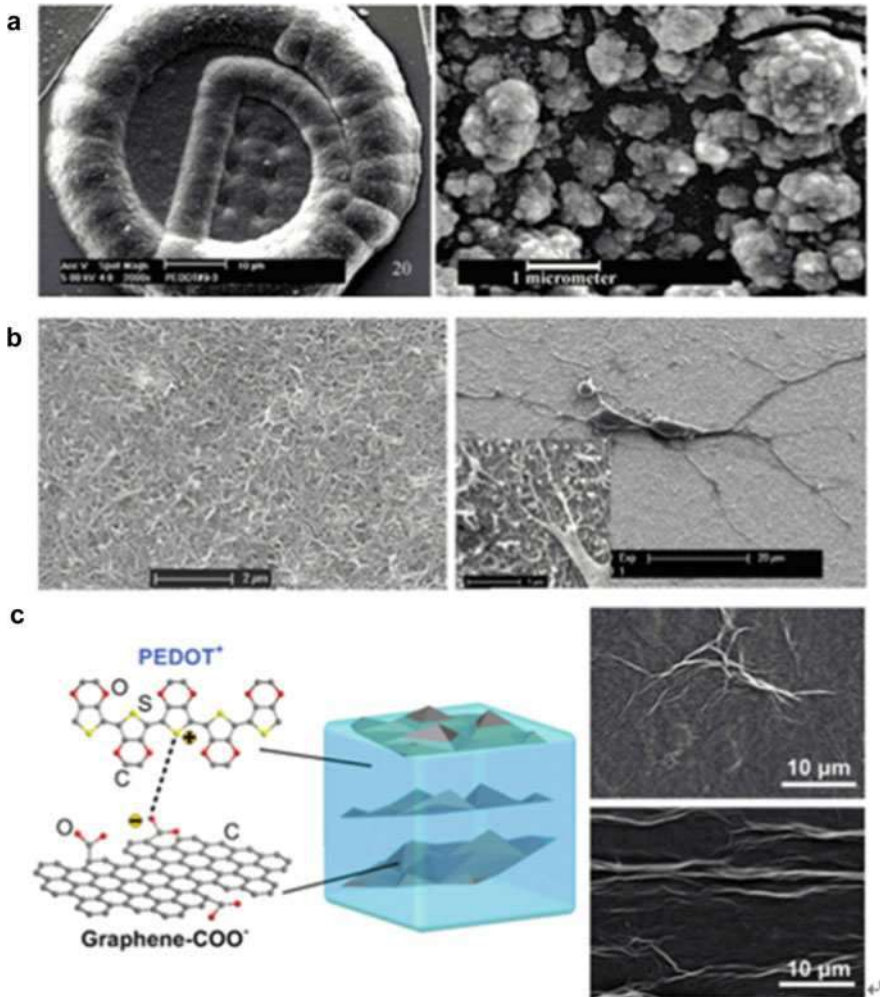
counter ion into PEDOT to form PEDOT/biomolecule composites as electrode-tissue interface (Asplund et al. 2008). These three composite electrode-tissue interface doped with different biomolecules varies from each other in their characteristics like surface morphology, effective area determined by surface roughness, and the electrochemical performance. The addition of biomolecules might improve the biocompatibility of the conducting polymer electrode-tissue interface. Furthermore, bioactive drugs and nutrition factors could also be added into the PEDOT/biomolecule composites to produce the drug-loaded functionalized electrode-tissue interface.

Xiliang Luo and Xinyan T. Cui et al. from University of Pittsburgh doped acidified carbon nanotube (CNT) as negatively charged counter ion separately into PEDOT to form PEDOT/CNT composite as electrode-tissue interface (Luo et al. 2011). As exhibited in Fig. 8c, the PEDOT/CNT composite exhibited rougher surface than other PEDOT electrode-tissue interfaces, which facilitated the improvement of electrochemical performance by increasing the effective area. The neuron cells grew well and tightly adhered on the PEDOT/CNT composite film, which indicated that the PEDOT/CNT neural interface possessed good biocompatibility. The porous structure of PEDOT/CNT composite was attributed to addition of carbon nanotube, which largely increased the effective area. Moreover, due to its electrical and mechanical performance, the PEDOT/CNT composite possessed excellent electrochemical property and stability.

Mohammad Reza Abidian and Daryl R. Kipke et al. from University of Michigan doped perchlorate ( $\text{ClO}_4^-$ ) as negatively charged counter ion separately into PEDOT to form hollow nanotube structure as electrode-tissue interface (Abidian et al. 2009). The PEDOT nanotube intersected and stacked together to form loose and porous extensional organization on the electrode surface, which improved the effective surface area of electrode-tissue interface. The hollow structure of PEDOT nanotube further increased the electrode-tissue interface area, which resulted in the improvement of electrochemical performance.

Hongchang Tian and Jingquan Liu et al. from Shanghai Jiao Tong University doped graphene oxide (GO) as negatively charged counter ion separately into PEDOT to form PEDOT/GO nanocomposite film as electrode-tissue interface



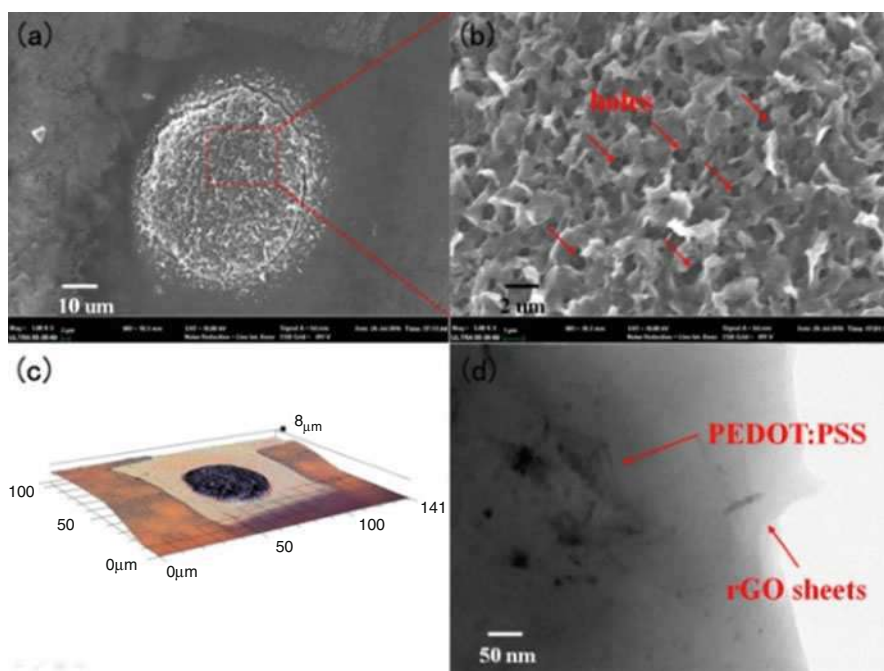


**Fig. 8** Research progress of the electrode-tissue interface modification technology of flexible MEMS microelectrodes for neural interface. (a) PEDOT/PSS composite film (Cui and Martin 2003); (b) PEDOT/carbon nanotube (CNT) composite film (Luo et al. 2011); (c) PEDOT/graphene oxide (GO) composite film (Tian et al. 2014b)

(Tian et al. 2014b, c, d). As shown in Fig. 8c, in PEDOT/GO composite film, GO was randomly distributed as the structural material to form three-dimensional cross-over networks, while PEDOT which served as stable charge transfer medium was interspersed among the interspaces of graphene nets. Like rebar in concrete, GO doping enhanced the mechanical property of conducting polymer film. Meanwhile, the encapsulated conducting polymer prevents GO from dispersing to tissue during recording or stimulation process, which greatly reduces the possibility of cytotoxicity induced by the diffusion of carbon nanomaterial. Like carbon nanotube, the

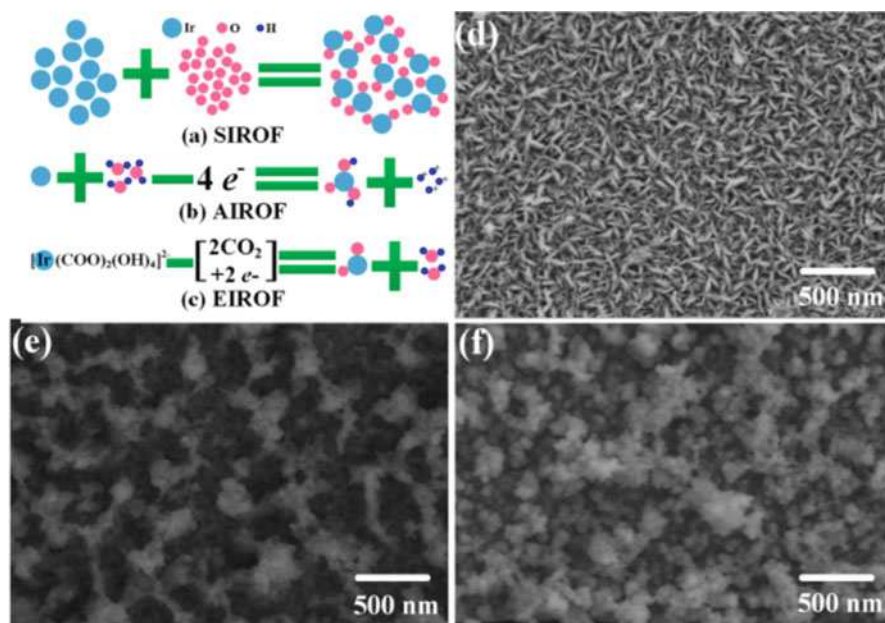
nanoscale GO also possessed multiple excellent properties, which could improve the performance of the conducting polymer electrode-tissue interface.

Minghao Wang and Jingquan Liu et al. of Shanghai Jiao Tong University reports for the first time the use of reduced graphene oxide enhanced conductive polymer (PEDOT:PSS-rGO) to improve the electrochemical properties, biocompatibility and mechanical stability of microelectrodes for high-quality neural recording (Wang et al. 2017). A flexible electrochemical method was adopted to realize the codeposition of PEDOT:PSS-rGO composites with L-Ascorbic acid on microelectrode sites without post-reduction. The SEM pictures shown in Fig. 9 illustrate the PEDOT:PSS-rGO has a porous wrinkle structure with large effective surface area which is beneficial to charge transfer and storage. The electrochemical tests demonstrate that the PEDOT:PSS-rGO coatings can decrease the impedance of microelectrodes by two orders of magnitude and increase the CSC by ten times. The microelectrodes modified with PEDOT:PSS-rGO recorded higher quality spikes ( $\text{SNR} > 10$ ) than the uncoated. Moreover, the microelectrodes modified with PEDOT:PSS-rGO had significantly higher amplitude and fewer low-frequency artifacts in LFP recordings. The calculated power spectra of the LFP signals illustrate the PEDOT:PSS-rGO-modified microelectrodes had higher LFPs sensitivity and common-mode noise suppression capability than the unmodified ones.



**Fig. 9** The SEM pictures of the deposited PEDOT:PSS-rGO on microelectrode site at (a) 1000 $\times$ , (b) 5000 $\times$  magnification. (c) The 3D microscopy pictures of the PEDOT:PSS-rGO modified neural probe at lateral view and (d) the TEM picture of the as-deposited PEDOT:PSS-rGO (Wang et al. 2017)

Xiaoyang Kang and Jingquan Liu et al. from Shanghai Jiao Tong University also developed iridium oxide ( $\text{IrO}_x$ ) as an important electrochemical modification material for neural interface, which was extremely valuable in neural stimulation and recording applications (Kang et al. 2014a, b, c, d, 2015). There were various kinds of preparation methods for  $\text{IrO}_x$ , which included sputtering iridium oxide film (SIROF), activated iridium oxide film (AIROF), and electrodeposited iridium oxide film (EIROF), as shown in Fig. 10a–c. For SIROF,  $\text{IrO}_x$  was formed by the combination of iridium atom and oxygen atom under vacuum condition. For AIROF, iridium atom reacted with water to form  $\text{IrO}_x$  hydrate. For EIROF,  $\text{IrO}_x$  hydrate was formed by drawing off carbon dioxide from iridium complex compound. The SEM of the iridium oxide as SIROF, AIROF, and EIROF are shown in Fig. 10d–f, respectively. The SIROF prepared under optimal condition exhibited dendrite surface morphology with porous structure. The AIROF displayed rough and porous structure which could facilitate fast ion exchange. The EIROF were suitable for short-time electrical stimulation due to its relatively high CSC and low impedance. The SIROF possessed better stability than the other two  $\text{IrO}_x$ , because there was no water existed in sputtered  $\text{IrO}_x$ . Therefore, it would be more suitable for long-term electrical stimulation. However, the AIROF was more suited to neural recording, because the phase angle shift of the electrochemically activated  $\text{IrO}_x$  was the smallest among all these three  $\text{IrO}_x$ .



**Fig. 10** Research progress of iridium oxide as an important electrochemical modification material for neural interface. (a–c) Three preparation methods for iridium oxide as SIROF, AIROF, and EIROF; (d–f) SEM of the iridium oxide as SIROF, AIROF, and EIROF (Asplund et al. 2008; Tian et al. 2014b, c, d; Kang et al. 2014a)

## Future Development Prospect

In the research field of flexible MEMS microelectrodes for neural interface, the development of microelectrodes with characteristics of tiny dimension, multifunction, high density, and biodegradation will be the goals. Moreover, transparent and stretchable MEMS microelectrodes which can facilitate fluorescence observation of neural tissue and promote conformal covering on brain tissue will become a new direction. Furthermore, in terms of electrode modifications for electrode-tissue interface, the ideal tissue engineered interface proposed by Ulises A. Aregueta-Robles et al. should incorporate the combined coating approaches of conductive polymers, hydrogels, and attachment factors with neural cells, which can give consideration to each requirement of electrode-tissue interface (Aregueta-Robles et al. 2014).

In recent years, the optogenetics applications in neuroscience have attracted much attention for neuroscientists. Although electrical stimulation methods exhibit remarkable advantages in controlling and exploring the function of discrete brain regions and providing therapeutic solutions, it cannot genetically target specified neuron types. This defect could be overcome with genetically encoded actuators (Warden et al. 2014). Therefore, flexible MEMS microelectrodes integrated with the optical stimulation capability will become a very important issue for neural interface in future.

In addition, the rising nanotechnology and biomedical engineering will offer a new opportunity for the development of flexible MEMS microelectrodes for neural interface. The interdisciplinary research of microfabrication technology with nanotechnology and biomedical engineering will lead the developing orientation in future (Yang et al. 2012; Aregueta-Robles et al. 2014; Warden et al. 2014; Zhang et al. 2014; Yu et al. 2014). The combination of these interdisciplinary subjects will undoubtedly collide to burst shining sparks and greatly improve human life and understanding the world.

---

## References

- Abidian MR, Martin DC (2009) Multifunctional Nanobiomaterials for neural interfaces. *Adv Funct Mater* 19(4):573–585
- Abidian MR, Ludwig KA, Marzullo TC, Martin DC, Kipke DR (2009) Interfacing conducting polymer nanotubes with the central nervous system: chronic neural recording using poly (3, 4-ethylenedioxythiophene) nanotubes. *Adv Mater* 21(37):3764–3770
- Abidian MR, Daneshvar ED, Egeland BM, Kipke DR, Cederna PS, Urbanchek MG (2012) Hybrid conducting polymer-hydrogel conduits for axonal growth and neural tissue engineering. *Adv Healthcare Mater* 1(6):762–767
- Al-bahrani MR, Ahmad W, Mehnane HF, Chen Y, Cheng Z, Gao Y (2015) Enhanced electrocatalytic activity by RGO/MWCNTs/NiO counter electrode for dye-sensitized solar cells. *Nano-Micro Lett* 7(3):298–306
- Altuna EB, Cid E, Aivar P, Gal B, Berganzo J, Gabriel G, Guimera A, Villa R, Fernandez LJ, Menendez de la Prida L (2013) SU-8 based microprobes for simultaneous neural depth recording and drug delivery in the brain. *Lab Chip* 13(7):1422–1430

- Anthony TE, Dee N, Bernard A, Lerchner W, Heintz N, Anderson DJ (2014) Control of stress-induced persistent anxiety by an extra-amygdala septohypothalamic circuit. *Cell* 156(3):522–536
- Aregueta-Robles UA, Woolley AJ, Poole-Warren LA, Lovell NH, Green RA (2014) Organic electrode coatings for next-generation neural interfaces. *Front Neuroeng* 7:15
- Arter JA, Taggart DK, McIntire TM, Penner RM, Weiss GA (2010) Virus-PEDOT nanowires for biosensing. *Nano Lett* 10(12):4858–4862
- Asplund M, von Holst H, Inganäs O (2008) Composite biomolecule/PEDOT materials for neural electrodes. *Biointerphases* 3(3):83–93
- Asplund M, Nyberg T, Inganäs O (2010) Electroactive polymers for neural interfaces. *Polym Chem* 1(9):1374–1391
- Au KM, Lu Z, Matcher SJ, Armes SP (2013) Anti-biofouling conducting polymer nanoparticles as a label-free optical contrast agent for high resolution subsurface biomedical imaging. *Biomaterials* 34(35):8925–8940
- Bangar MA, Shirale DJ, Chen W, Myung NV, Mulchandani A (2009) Single conducting polymer nanowire chemiresistive label-free immunosensor for cancer biomarker. *Anal Chem* 81(6):2168–2175
- Bongo M, Winther-Jensen O, Himmelberger S, Strakosas X, Ramuz M, Hama A, Stavrinidou E, Malliaras GG, Salleo A, Winther-Jensen B (2013) PEDOT:gelatin composites mediate brain endothelial cell adhesion. *J Mater Chem B* 1:3860–3867
- Cho Y, Borgens RB (2013) Electrically controlled release of the nerve growth factor from a collagen-carbon nanotube composite for supporting neuronal growth. *J Mater Chem B* 1(33):4166–4170
- Cogan SF (2008) Neural stimulation and recording electrodes. *Annu Rev Biomed Eng* 10:275–309
- Cogan SF, Guzelian AA, Agnew WF, Yuen TG, McCreery DB (2004) Over-pulsing degrades activated iridium oxide films used for intracortical neural stimulation. *J Neurosci Meth* 137(2):141
- Cui XY, Martin DC (2003) Electrochemical deposition and characterization of poly(3,4-ethylenedioxythiophene) on neural microelectrode arrays. *Sensor Actuat B-Chem* 89(1–2):92–102
- Dobson J (2008) Remote control of cellular behaviour with magnetic nanoparticles. *Nat Nanotechnol* 3(3):139–143
- Farina D, Yoshida K, Stieglitz T, Koch KP (2008) Multichannel thin-film electrode for intramuscular electromyographic recordings. *J Appl Physiol* 104(3):821–827
- Ferguson JE, Boldt C, Redish AD (2009) Creating low-impedance tetrodes by electroplating with additives. *Sensor Actuat A-Phys* 156(2):388–393
- Gao KP, Li G, Liao LY, Cheng J, Zhao JL, Xu YS (2013) Fabrication of flexible microelectrode arrays integrated with microfluidic channels for stable neural interfaces. *Sensor Actuat A-Phys* 197:9–14
- Gomez N, Lee JY, Nickels JD, Schmidt CE (2007) Micropatterned polypyrrole: a combination of electrical and topographical characteristics for the stimulation of cells. *Adv Funct Mater* 17(10):1645–1653
- Grill WM, Norman SE, Bellamkonda RV (2009) Implanted neural interfaces: biochallenges and engineered solutions. *Annu Rev Biomed Eng* 11:1–24
- Hira R, Honkura NJ, Maruyama Y, Augustine G, Kasai H, Matsuzaki M (2009) Transcranial optogenetic stimulation for functional mapping of the motor cortex. *J Neurosci Meth* 179(2):258–263
- Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD, Donoghue JP (2006) Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 442(7099):164–171
- Hong X, Wu Z, Chen L, Wu F, Wei L, Yuan W (2014) Hydrogel microneedle arrays for transdermal drug delivery. *Nano-Micro Lett* 6(3):191–199
- Hsiao YS, Kuo CW, Chen P (2013) Multifunctional Graphene-PEDOT microelectrodes for on-chip manipulation of human Mesenchymal stem cells. *Adv Funct Mater* 23(37):4649–4656

- Huang YJ, Wu HC, Tai NH, Wang TW (2012) Carbon Nanotube rope with electrical stimulation promotes the differentiation and maturity of neural stem cells. *Small* 8(18):2869–2877
- Iwai Y, Honda S, Ozeki H, Hashimoto M, Hirase H (2011) A simple head-mountable led device for chronic stimulation of optogenetic molecules in freely moving mice. *Neurosci Res* 70(1):124–127
- Jarc M, Berniker M, Tresch MC (2013) FES control of isometric forces in the rat Hindlimb using many muscles. *IEEE Trans Bio-Med Eng* 60(5):1422–1430
- Jeong JW, Mccall JG, Shin G, Zhang Y, Alhasani R, Kim M (2015) Wireless optofluidic systems for programmable in vivo pharmacology and optogenetics. *Cell* 162(3):662–674
- Jessin J, Yuefa L, Jinsheng Z, Jeffrey AL, Yong X (2011) Microfabrication of 3D neural probes with combined electrical and chemical interfaces. *J Micromech Microeng* 21(10):105011
- Ji BW, Kang XY, Wang MH, Bao BF, Tian HC, Yang B, Chen X, Wang XL, Liu JQ (2017) Photoelectric neural interface combining wire-bonding LEDs with iridium oxide microelectrodes for optogenetics, MEMS 2017, Las Vegas, 22–26 Jan
- Kang XY, Liu JQ, Tian HC, Zhang C, Yang B, NuLi Y, Zhu HY, Yang CS (2014a) Controlled activation of iridium film for AIROF microelectrodes. *Sensor Actuat B-Chem* 190:601–611
- Kang XY, Liu JQ, Tian HC, Yang B, Nuli YN, Yang CS (2014b) Fabrication and electrochemical comparison of SIROF-AIROF-EIROF microelectrodes for neural interfaces. *IEEE Eng Med Biol*:478–481
- Kang XY, Liu JQ, Tian HC, Yang B, NuLi YN, Yang CS (2014c) Optimization and electrochemical characterization of RF-sputtered iridium oxide microelectrodes for electrical stimulation. *J Microelectromech Syst* 24(2)
- Kang XY, Liu JQ, Tian HC, Du JC, Yang B, Zhu HY, NuLi YN Yang CS (2014d) Fabrication and degradation characteristic of sputtered iridium oxide neural microelectrodes for Fes application, MEMS 2014, San Francisco, 26–30 Jan, 616–619
- Kang XY, Liu JQ, Tian HC, Yang B, Nuli YN, Yang CS (2015) Self-closed Parylene cuff electrode for peripheral nerve recording. *J Microelectromech Syst* 24(2):319–332
- Kim S, Bhandari R, Klein M, Negi S, Rieth L, Tathireddy P, Toepper M, Oppermann H, Solzbacher F (2009) Integrated wireless neural interface based on the Utah electrode array. *Biomed Microdevices* 11(2):453–466
- Kim H, Viventi J, Amsden JJ, Xiao JL, Vigeland L, Kim YS, Blanco JA, Panilaitis B, Frechette ES, Contreras D, Kaplan DL, Omenetto FG, Huang YG, Hwang KC, Zakin MR, Litt B, Rogers JA (2010) Dissolvable films of silk fibroin for ultrathin conformal bio-integrated electronics. *Nat Mater* 9(6):511–517
- Kozai TDY, Langhals NB, Patel PR, Deng XP, Zhang HN, Smith KL, Lahann J, Kotov NA, Kipke DR (2012) Ultrasmall implantable composite microelectrodes with bioactive surfaces for chronic neural interfaces. *Nat Mater* 11(12):1065–1073
- Kwon KY, Lee HM, Ghovanloo M, Weber A, Li W (2014) A wireless slanted optrode array with integrated micro LEDs for optogenetics, MEMS 2014, San Francisco, 26–30 Jan
- Lee JY, Bashur CA, Goldstein AS, Schmidt CE (2009) Polypyrrole-coated electrospun PLGA nanofibers for neural tissue applications. *Biomaterials* 30(26):4325–4335
- Luo X, Weaver CL, Zhou DD, Greenberg R, Cui XT (2011) Highly stable carbon nanotube doped poly (3, 4-ethylenedioxythiophene) for chronic neural stimulation. *Biomaterials* 32(24):5551–5557
- Martins PM, Ribeiro S, Ribeiro C, Sencadas V, Gomes AC, Gama FM, Lanceros-Mendez S (2013) Effect of poling state and morphology of piezoelectric poly(vinylidene fluoride) membranes for skeletal muscle tissue engineering. *RSC Adv* 3(39):17938–17944
- Memberg WD, Stage TG, Kirsch RF (2014) A fully implanted intramuscular bipolar Myoelectric signal recording electrode. *Neuromodulation* 17(8):794–799
- Metz S, Bertsch A, Bertrand D, Renaud P (2004) Flexible polyimide probes with microelectrodes and embedded microfluidic channels for simultaneous drug delivery and multi-channel monitoring of bioelectric activity. *Biosens Bioelectron* 19(10):1309–1318
- Midrio M (2006) The denervated muscle: facts and hypotheses. A historical review. *Eur J Appl Physiol* 98(1):1–21

- Mitch WE, Goldberg AL (1996) Mechanisms of disease: mechanisms of muscle wasting: the role of the ubiquitin-proteasome pathway. *New Engl J Med* 335(25):1897–1905
- Navarro X, Krueger TB, Lago N, Micera S, Stieglitz T, Dario P (2005) A critical review of interfaces with the peripheral nervous system for the control of neuroprostheses and hybrid bionic systems. *J Peripher Nerv Syst* 10(3):229
- Ortiz-Catalan M, Branemark R, Hakansson B, Delbeke J (2012) On the viability of implantable electrodes for the natural control of artificial limbs: review and discussion. *Biomed Eng Online* 11(1):33
- Park SY, Park J, Sim SH, Sung MG, Kim KS, Hong BH, Hong S (2011) Enhanced differentiation of human neural stem cells into neurons on Graphene. *Adv Mater* 23(36):H263–H267
- Plesse C, Vidal F, Teyssié D, Chevrot C (2010) Conducting polymer artificial muscle fibres: toward an open air linear actuation. *Chem Commun* 46(17):2910–2912
- Pongrácz ZF, Márton G, Bérces Z, Ulbert I, Fürjes P (2013) Deep-brain silicon multielectrodes for simultaneous in vivo neural recording and drug delivery. *Sensor Actuat B-Chem* 189:97–105
- Poole-Warren L, Lovell N, Baek S, Green R (2010) Development of bioactive conducting polymers for neural interfaces. *Expert Rev Med Devices* 7(1):35–49
- Quigley F, Razal JM, Kita M, Jalili R, Gelmi A, Penington A, Ovalle-Robles R, Baughman RH, Clark GM, Wallace GG (2012) Electrical stimulation of myoblast proliferation and differentiation on aligned nanostructured conductive polymer platforms. *Adv Healthc Mater* 1(6):801–808
- Radisic M, Park H, Shing H, Consi T, Schoen FJ, Langer R, Freed LE, Vunjak-Novakovic G (2004) From the cover: functional assembly of engineered myocardium by electrical stimulation of cardiac myocytes cultured on scaffolds. *Proc Natl Acad Sci U S A* 52:18129–18134
- Receveur RAM, Lindemans FW, de Rooij NF (2007) Microsystem technologies for implantable applications. *J Micromech Microeng* 17(5):R50–R80
- Robblee LS, Mchardy J, Agnew WF, Bullara LA (1983) Electrical stimulation with pt electrodes. Vii. Dissolution of pt electrodes during electrical stimulation of the cat cerebral cortex. *J Neurosci Meth* 9(4):301–308
- Rui YF, Liu JQ, Wang YJ, Yang CS (2011) Parylene-based implantable pt-black coated flexible 3-D hemispherical microelectrode arrays for improved neural interfaces. *Microsyst Technol* 17(3):437–442
- Rui YF, Liu JQ, Yang B, Li KY, Yang CS (2012) Parylene-based implantable platinum-black coated wire microelectrode for orbicularis oculi muscle electrical stimulation. *Biomed Microdevices* 14(2):367–373
- Simon T, Kurup S, Larsson KC, Hori R, Tybrandt K, Goiny M, Jager EW, Berggren M, Canlon B, Richter-Dahlfors A (2009) Organic electronics for precise delivery of neurotransmitters to modulate mammalian sensory function. *Nat Mater* 8(9):742–746
- Svennersten K, Berggren M, Richter-Dahlfors A, Jager EWH (2011) Mechanical stimulation of epithelial cells using polypyrrole microactuators. *Lab Chip* 11(19):3287–3293
- Takeuchi S, Ziegler D, Yoshida Y, Mabuchi K, Suzuki T (2005) Parylene flexible neural probes integrated with microfluidic channels. *Lab Chip* 5(5):519–523
- Tandon N, Cannizzaro C, Chao PHG, Maidhof R, Marsano A, Au HTH, Radisic M, Vunjak-Novakovic G (2009) Electrical stimulation systems for cardiac tissue engineering. *Nat Protoc* 4(2):155–173
- Thomas CK, Zaidner EY, Calancie B, Broton JG, Bigland-Ritchie BR (1997) Muscle weakness, paralysis, and atrophy after human cervical spinal cord injury. *Exp Neurol* 148(2):414–423
- Tian HC, Liu JQ, Du JC, Kang XY, Zhang C, Yang B, Chen X, Yang CS (2014a) Flexible intramuscular micro tube electrode combining electrical and chemical Interface. *IEEE Eng Med Biol*:6949–6952
- Tian HC, Liu JQ, Wei DX, Kang XY, Zhang C, Du JC, Yang B, Chen X, Zhu HY, NuLi YN, Yang CS (2014b) Graphene oxide doped conducting polymer nanocomposite film for electrode-tissue interface. *Biomaterials* 35(7):2120–2129
- Tian HC, Liu JQ, Kang XY, Wei DX, Zhang C, Du JC, Yang B, Chen X, Yang CS (2014c) Biotic and abiotic molecule dopants determining the electrochemical performance, stability and fibroblast behavior of conducting polymer for tissue interface. *RSC Adv* 4(88):47461–47471

- Tian HC, Liu JQ, Kang XY, Wei DX, Zhang C, Du JC, Yang B, Chen X, Yang CS (2014d) Poly(3,4-ethylenedioxythiophene)/Graphene oxide composite coating for electrode-tissue Interface. *IEEE Eng Med Biol*:1571–1574
- Tian HC, Liu JQ, Kang XY, He Q, Yang B, Chen X, Yang CS (2015) Flexible multi-channel microelectrode with fluidic paths for intramuscular stimulation and recording. *Sensor Actuat A-Phys* 228:28–39
- Vallejo-Giraldo AK, Biggs MJP (2014) Biofunctionalisation of electrically conducting polymers. *Drug Discov Today* 19(1):88–94
- Wang MH, Nikaido K, Kim Y, Ji BW, Tian HC, Kang XY, Yang CS, Yang B, Chen X, Wang XL, Zhang Y, Liu JQ (2017) Flexible cylindrical neural probe with graphene enched conductive polymer for multi-mode BCI applications, MEMS 2017, Las Vegas, 22–26 Jan
- Warden MR, Cardin JA, Deisseroth K (2014) Optical neural interfaces. *Annu Rev Biomed Eng* 16(16):103–129
- Wells J, Kao C, Mariappan K, Albea J, Jansen ED, Konrad P, Mahadevan-Jansen A (2005) Optical stimulation of neural tissue in vivo. *Opt Lett* 30(5):504–506
- Wise KD, Sodagar AM, Yao Y, Gulari MN, Perlin GE, Najafi K (2008) Microelectrodes, microelectronics, and implantable neural microsystems. *Proc IEEE* 96(7):1184–1202
- Yang SY, Kim BN, Zakhidov AA, Taylor PG, Lee JK, Ober CK, Lindau M, Malliaras GG (2011) Detection of transmitter release from single living cells using conducting polymer microelectrodes. *Adv Mater* 23(24):H184–H188
- Yang Z, Gao RG, Hu NT, Chai J, Cheng YW, Zhang LY, Wei H, Kong ESW, Zhang YF (2012) The prospective two-dimensional Graphene Nanosheets: preparation, Functionalization, and applications. *Nano-Micro Lett* 4(1):1–9
- Yang Z, Zhang Y, Itoh T, Maeda R (2014) Flexible implantable microtemperature sensor fabricated on polymer capillary by programmable UV lithography with multilayer alignment for biomedical applications. *J Microelectromech Syst* 20:21–29
- Yoon H, Jang J (2009) Conducting-polymer Nanomaterials for high-performance sensor applications: issues and challenges. *Adv Funct Mater* 19(10):1567–1576
- Yoshida K, Farina D, Akay M, Jensen W (2010) Multichannel Intraneural and intramuscular techniques for multiunit recording and use in active prostheses. *Proc IEEE* 98(3):432–449
- Yu L, Wu H, Wu B, Wang Z, Cao H, Fu C, Jia N (2014) Magnetic Fe<sub>3</sub>O<sub>4</sub>-reduced graphene oxide nanocomposites-based electrochemical biosensing. *Nano-Micro Lett* 6(3):258–267
- Zhang AMA, Adamantidis A, De LL, Deisseroth K (2007) Circuit-breakers: optical technologies for probing neural signals and systems. *Nat Rev Neurosci* 8(8):577
- Zhang Z, Li SW, Xue C, Yang S, Zhang W (2014) A bionic fish cilia median-low frequency three-dimensional piezoresistive MEMS vector hydrophone. *Nano-Micro Lett* 6(2):136–142
- Zhao Y (2009) Investigating electrical field-affected skeletal myogenesis using a microfabricated electrode array. *Sensor Actuat A-Phys* 154(2):281–287
- Ziegler TS, Takeuchi S (2006) Fabrication of flexible neural probes with built-in microfluidic channels by thermal bonding of Parylene. *J. Microelectromech Syst* 15(6):1477–1482





# Electrode Array for Neural Interfaces

Weihua Pei and Hongda Chen

## Contents

Neural Interface .....	1438
Neuroelectricity .....	1439
Neural Microelectrode and Neural Signal Sensing .....	1440
Equivalent Circuit and Impedance of Electrode .....	1441
Electrode Noise .....	1442
Classification of Neural Electrodes .....	1444
Surface Neural Electrodes .....	1444
Implantable Neural Electrodes .....	1445
Neural Electrode Fabricated by MEMS Techniques .....	1446
Skin Electrodes Based on Microneedle Array .....	1446
Metal Wire Implantable Electrodes .....	1449
Processing of Bulk Silicon Microelectrodes .....	1451
Processing of Microelectrodes Based on Thin-Film Silicon .....	1453
Advantages and Development of Silicon-Based MEMS Microelectrodes .....	1455
Process of Flexible Polymer Microelectrodes .....	1456
Interfacial Modification of Neural Electrodes .....	1459
Challenges Neural Microelectrodes Facing With .....	1464
References .....	1465

---

W. Pei (✉)

State Key Laboratory of Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

CAS Center for Excellence in Brain Science and Intelligence Technology, Beijing, China

University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

e-mail: [peiwh@semi.ac.cn](mailto:peiwh@semi.ac.cn)

H. Chen

State Key Laboratory of Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

© Springer Nature Singapore Pte Ltd. 2018

Q.-A. Huang (ed.), *Micro Electro Mechanical Systems*, Micro/Nano Technologies,

[https://doi.org/10.1007/978-981-10-5945-2\\_42](https://doi.org/10.1007/978-981-10-5945-2_42)

1437

---

**Abstract**

A neural interface is a kind of device or system which is used to connect neural system with external equipment, through recording electroneurographic (EEG) signals, a neural interface monitors or stimulates and regulates neural activities. To reduce as much as possible the interference of the device in neural system, neural interfaces are usually fabricated with MEMs technology, which helps to minimize the dimensions of neural interfaces. Apart from meeting the dimension requirements, neural interfaces should also be biocompatible in terms of biochemical characteristics, electrical properties, and mechanical properties. These requirements or limits mean more challenges upon neural interface materials and processing technologies. In this chapter, we will take implantable neural microelectrode array as an example, and introduce the development of existing microfabrication technology including working principle, material selection, structure, and manufacturing process of neural microelectrode device. At last, we will sum up the problems and challenges microelectrode devices are facing.

---

**Keywords**

Impedance · Implant · Microelectrode array · Neural interface · Recording site · Single unite · Stimulation

---

## Neural Interface

Neural interfaces refer to devices that can sense neuron's activity information or regulate and control the ways of neuron activity. Electrical activity is one of the most basic ways of neuron activity. Also, it is the main way for nerves to express, transfer, and receive signals. If we can record neural electrical signals and interpret the meaning of them, we will be able to control or command artificial equipment with the help of proper devices. In this way, paraplegics can control various mechanical equipment with EEG to complete various actions in daily life. In addition, using electrical stimulation to interfere in abnormal physiological activities of nerves, we can control epileptic fits and the onset of Parkinson's disease. Using coded signals modulated visually or acoustically to stimulate the remaining optic nerve or auditory nerve, the damaged hearing or sight will be restored (Humayun et al. 1999; Jung et al. 2015; Pandarinath et al. 2017). This kind of technology, which transfers information between nerve and artificial equipment, is called brain-computer interaction technology and the sensors used to record neural signals or stimulate and evoke neural signals are called neural interfaces. Neural electrode is one of the most frequently used neural interface devices.

Generally, neural interfaces can be divided into two categories: noninvasive and invasive interfaces. Noninvasive neural interface is put outside the scalp to detect neural activities by obtaining EEG from around the scalp. Therefore, it is usually called scalp electrode or brain electrode. Invasive neural interface is planted into cerebral cortex by surgery to record local field potential with high spatial resolution and even the electrical activities of a single neuron. Apart from EEG, other methods

such as magnetoencephalography (MEG), thermal imaging, functional magnetic resonance imaging (fMRI), and infrared spectrum (IR spectrum) are also used to obtain brain neural activity informations. Still, EEG signal is more frequently used in brain-computer interface technology.

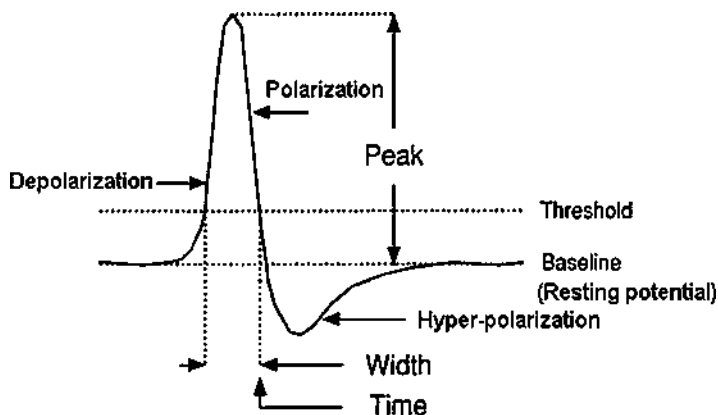
In a narrow sense, neural interface devices usually mean invasive neural electrode. According to the location, invasive neural electrodes can be divided into two types: cortical electrode and intracortical electrode (also called extracellular electrode). The former is located under the skull, above the cortex and its film-shaped electrode is mainly used to record ECoG signals; the latter is usually needle-shaped and can be stuck into cortex, where it is closer to neurons.

As to the quality of EEG signals obtained, compared with noninvasive electrode, invasive electrode obtains signals with higher quality and resolution (from single neuron), which can be attributed to it being very close or even clinging to neurons which generate electrical activities. However, invasive electrodes will cause risks of hemorrhage or infection. Even if these problems are well controlled, there will be a problem of biocompatibility between the implanted electrode and surrounding tissues. In such cases, the electrode may be enwrapped by surrounding tissues, and the electrode will not work as it gets too far away from the active neurons. So far, this problem is still a major one to be solved for record-type neural interfaces.

## Neuroelectricity

Neuroelectricity are generated by the electrical activities of nerve cells. The electrical signals of single neurons mainly include resting potential and action potential. Potential of nerve cell refers to the difference between intracellular potential and average potential of extracellular fluids (separated by well insulated cell membrane), which is mainly caused by the concentration difference between anions and cations in and out of the cell membrane. Potential changes when the difference of ion concentrations in and out of the cell membrane is changed through the ion channels on the membrane. The ions mainly include  $K^+$ ,  $Na^+$ ,  $Cl^-$ , etc.

Resting potential is the potential difference in and out of the cell membrane when the cell is not excited. Generally, resting potential ranges from  $-30$  to  $90$  mV (extracellular potential is usually defined as 0). If the extracellular potential is positive and intracellular potential is negative, it is the status that we call polarization. And if the difference between extracellular potential and intracellular potential is shortening, we call the status depolarization. If the difference between extracellular potential and intracellular potential is enlarging, we call the status hyperpolarization. Action potential refers to the process where cell potential experiences depolarization, polarization, hyperpolarization before it returns to resting status after the nerve cell is stimulated, the permeability of ion channels on cell membrane changes, and the potential difference in and out of the membrane changes quickly as a result of quick ion exchanging between both sides of the membrane. The duration is usually around the magnitude of ms. Different types of neurons have their own action potential models, for example, different hyperpolarization extent and recovery



**Fig. 1** Diagram of the generation of action potential

time. Peak value of potential, also called spike, will appear along with action potential. The emission frequency of action potential is directly proportional to the stimulation intensity. Generally, a single neural signal is coded with emission frequency. Figure 1 is the diagrammatic sketch of resting potential and action potential.

Resting potential and action potential are electrical signals related to activities of a single neuron. In our brain, there is another kind of electrical signals that is related to ensemble neuronal activities. These signals are usually slow wave signals with lower frequency and known as EEG signals. According to the recorded positions of these signals, they can be classified as local field potential, ECoG, and EEG.

## Neural Microelectrode and Neural Signal Sensing

Neural microelectrode is the sensor converting bioelectric signals carried by ions into current or voltage signals carried by electrons. Electroneurographic (ENG) signals are current signals created by the flow of voltage or ions generated by the unbalanced distribution of cations and anions. The said voltage or current are usually very weak or located within a small area. Therefore, the sensing electrode must be highly sensitive and very compact in dimensions. Most electrical devices, circuit and electronic equipment transfer and process the current or voltage signals based on the flow and distribution of electrons. In order to interface ENG with the electric signal of external equipment, neural interface devices are required to convert the electric signals carried by ions into electronic signals. The converting occurs at the electrode/tissue interface, and problems of two-phase interface between electrolyte and metal or other electrode materials are involved in the electrochemical interface formed between electrode and tissue.

When the electrode comes into contact with the liquid containing ions, as the inner potential of electrode is different from that of the body fluid, there will be

contact potential between electrode and the solution. The value of the contact potential depends on the electrode material, the type and concentration of ions in the solution, as well as conditions such as temperature. When microelectrode contacts body fluid (or extracellular fluid) containing ions such as  $K^+$ ,  $Na^+$ ,  $Ca^{2+}$ , and  $Cl^-$ , these ions will redistribute near the electrode surface. At the same time, electrons in the metal will also redistribute.

Generally, as long as the chemical potential of metal is not equal to that of solution, an electric field will form at the metal-fluid interface. The value of the potential depends on the metal material used at the electrode recording site, ion concentration of the solution, and ambient temperature.

As the electrode is used to record or stimulate the neural signals, the area of electrode exposed to the body fluid must be limited (by coating an insulation layer on other parts or by other methods) to sense potential changes in partial area of body fluid, for example, near a certain neuron. The area exposed to body fluid decides the characteristics of recorded neural signals. In this article, the recording site is used to refer in particular to the exposed metal part of the electrode device. When the material of electrode is stable and concentration of body fluid remains unchanged, the potential, namely the potential difference between electrode and body fluid, is balanced and stable.

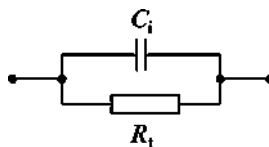
If the ion concentration of the body fluid changes, for example, when neuron discharges, the ion channels on cell membrane will open, there will instantaneously be a significant change of ion concentration near the cell. If at that time electrode recording site is located exactly at the area where concentration changes, the potential balance between electrode and body fluid will be tipped. Then the potential of electrode and body fluid needs to be raised or reduced to reach a new balance. On the premise that the impedance of electrode/body fluid interface is small enough, establishment of the new balance will be more quick than nervous discharge. Therefore, neural electrode is a device that can monitor real-time neural electrical activities.

## Equivalent Circuit and Impedance of Electrode

The electrical characteristics of body fluid/electrode interface equal to a circuit model with an electric capacity  $C_i$  and an electric resistance  $R_i$  connected in parallel. Generally speaking, electrode impedance is the impedance value of parallel  $R_i$  and  $C_i$  then connected with  $R_{met}$  in series.  $R_{met}$  is the impedance of the metal part of the electrode. Commonly,  $R_{met}$  is much less than the interface impedance. So, the equivalent circuit of a neural electrode can be simplified as Fig. 2.

In order to learn more about the equivalent circuit of body fluid/electrode interface, we have to think about how the electrical signals created by nerve cell are transferred to the electrode. Electrical activities of excitable cell generate micro-current. Carrying the current, ions in extracellular fluid move to the double layer near electrode from around the cell membrane through diffusion. Then, these ions may either enter the double layer on the side of solution and induce oppositely charged

**Fig. 2** Simplified equivalent circuit at a solution-electrode interface



electrons on the side of electrode or move through the double layer, reach the surface of electrode, and transfer electrons to electrode or obtain electrons from the electrode.

Electrode impedance is related to the surface area of electrode recording site: the larger the surface area is, the smaller the impedance will be. Electrode impedance is the function of frequency. As there exists double layer capacity  $C_i$ , electrode impedance will decrease as the frequency increases.

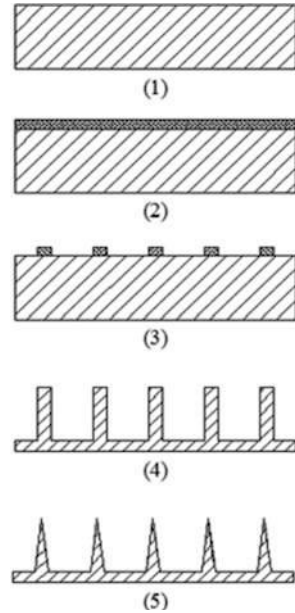
Electrode impedance is the most important electric characteristic of bioelectrode. Based on the above discussion, the value of electrode impedance is related to the material of electrode, recording site, and the surface area of electrode interface. The selection of recording site sizes is influenced by various factors. When the material of recording site is determined, electrode impedance will be decided by the size and surface area of recording site.

## Electrode Noise

When capturing electrophysiological signals, bioelectrode will unavoidably be influenced by various kinds of interference and noises. Among them, electrode thermal noise and the background noise in body fluid environment are the main noises. Electrode thermal noise is directly proportional to square root of electrode impedance and temperature. Thermal noise belongs to white noise (up to 100 THz) and is created by the irregular thermal motions of current carriers. The value of thermal noisy power spectral density induced by electrode impedance is calculated using  $S(f) = 4kTR(V^2/Hz)$ , wherein  $k$  is the Boltzmann's constant ( $=1.38 \times 10^{-23} \text{ J/K}$ ),  $T$  is the absolute temperature, and  $R$  is the electrode impedance, which includes not only the impedance of the model in Fig. 3, but also the impedances of all the electric connections between electrode and the recording amplifier. Taking an electrode with impedance of  $1 \text{ M}\Omega$  as an example, at body temperature of  $36 \text{ }^\circ\text{C}$ , suppose the interested frequency bandwidth ranges from 0 to 10 kHz, the mean square root of thermal noise is  $12 \text{ }\mu\text{V}$ , the influence of a noise source with such amplitude on the signal-to-noise ratio (SNR) of extracellular recorded signals should not be ignored. In order to raise the SNR, electrode impedance should be reduced as much as possible.

Background noises within a biosome refer to electrical signals created by discharge activities of other objects than the recorded neuro or other interfering signals sensed from the electrode. Its mean square root amplitude is up to mV magnitude. The electrode will definitely receive the noises apart from the electrical information from target neuro or tissue. When recording spike from single neuron, in order to

**Fig. 3** Fabrication process of dry electrode



reduce the influence of background noises, the size of the recording site should be as small as possible. However, the problem is that a small area leads to higher impedance. Therefore, the designers have to thrash out a compromise between the size of recording site and electrode impedance to achieve best recording effect. At the same time, the area of recording site has much to do with the frequency of the signal and spatial resolution. Generally, taking round electrode as an example, the recording site’s diameter for single neuron recording approximately ranges from 10 to 30  $\mu\text{m}$ ; electrode diameter used to record local field potential signals ranges from 50  $\mu\text{m}$  to several hundred  $\mu\text{m}$ ; the diameter of electrode used to record ECoG signals or EEG signals is usually more than 1 mm. With the development of electrode technology and the demand upon ENG signal application, electrode recording site are bound to be ultrasmall in size ultimately.

Electromagnetic interference (EMI) is another source of electrode noise. EMI can be controlled by external shielding measures. For example, neural signal extraction test with electrode can be done in an environment well protected by electromagnetic shielding measures. In addition, the following measures can also reduce the noise interference outside of the required signal frequency range: using shielded cables or optical cables to transfer signals; applying as far as possible DC power supply for amplifier and signal collection equipment; adopting measures such as filtering according to the features of required signals; putting electrode system and its recorded object into electromagnetic shielded environment.

With the development of material technology and improvement of processing techniques, for example, adopting nanotechnology or surface processing technology, surface area of electrode can be greatly increased without increasing the geometric

area, thus the impedance will be largely reduced and SNR will be improved. Also, the preamplifier circuit can be integrated around the electrode recording site, thus the influence of EMI on SNR will be minimized. Nanotechnology or MEMS technique can be used to improve corresponding performances of electrode.

---

## **Classification of Neural Electrodes**

According to their functions, neural electrodes can be divided into detecting electrodes and stimulating electrodes. Detecting electrodes are sensitive and specifically used for bioelectric potential measurement. Within a biosome, the physical and chemical changes lead to uneven distribution of positive and negative electric charges in various parts, and therefore the potentials of different parts of the biosome or in and out of a cell are different. We detect bioelectrical signals through measuring the potential differences of a pair of electrodes placed at proper location of a biosome. This kind of electrodes are called detecting electrodes. As bioelectrical signals are very weak and the background is noisy, the situation at the interface between biosome and electrodes is quite complex. Therefore, the polarization voltage of the recording (or measuring) electrodes should be lower, so as to reduce the influence of polarization to the records. Stimulating electrodes are used to apply electric current to biosome and act as executive components. Compared with recording site, stimulating electrodes require more electrochemically stable materials.

According to the application locations, electrodes can be divided into surface electrodes and internal electrodes. With larger sizes, the former are usually attached to the surface of our bodies and used for EEG measurement, anesthesia depth monitoring during surgical operations, or sleep quality detection. Internal electrodes must be planted into our bodies through surgical operations or be pushed into our body through our skin. They are used to record local signals in our bodies, such as local field potentials or action potentials of action cells. Usually, to control the background noises, the recording site will be limited. Most part of the electrode body will be wrapped by insulation materials and only a small tip is exposed for electrical activity detection.

## **Surface Neural Electrodes**

The major advantage of surface electrodes are noninvasive. And the disadvantages are the extracted EEG or neural electrical signals being weak in amplitude and low in resolution. The most frequently used surface electrodes are EEG (or scalp) electrodes. Generally, stratum corneum of our skin is a layer of poor conductor with high impedance. In order to achieve good electrical contact between electrodes and skin, surface electrodes are used in conjunction with electrographic gel. On one hand, electrographic gel is able to wet stratum corneum; on the other hand, it is able to fill in the gap and form good contact with low impedance between skin and electrodes.



Frequently used scalp electrodes are wet electrodes made of Ag/AgCl. Although wet electrodes feature high-quality signals and strong anti-interference ability, they must be assisted by electrographic gel, which brings some inconvenience: when using the gel, especially at scalp covered with hair, it will take longer time to make preparations and apply more gel when the previously applied gel gets dry. Finally, when the operation is completed, it is hard to get rid of the remaining gel. In order to make it more convenient when using surface electrode, researchers come up with many kinds of dry electrodes which do not need gel. These dry electrodes can generally be divided into two types: surface electrodes suitable for area without hair (such as forehead and root part of auricle) and comb-shaped electrodes suitable for area covered with hair. According to the principle of sensing, dry electrodes can be divided into capacitive electrodes and resistance electrodes; according to the structure, dry electrodes can be divided into minimally invasive electrodes based on microneedle array and surface contact electrodes; according to materials, dry electrodes can be divided into rigid dry electrodes and flexible dry electrodes. Viewing the development trend of portable system for health monitoring and brain-computer interaction equipment for application, in the future, skin electrodes featuring dry-contact and high SNR dry electrode devices, which are comfortable to wear, will be most popular.

## **Implantable Neural Electrodes**

Implantable neural microelectrode is a very important kind of neural interface. They need to work in body fluid environment, which is quite different from surface electrodes. On one hand, as they contact or connect nerve cells, they must meet relevant requirements of biocompatibility to maintain normal physiological functions of the part in contact; on the other hand, they should have good mechanical and electrical characteristics so as to meet the requirements of subsequent amplifying and signal processing circuits on SNR, frequency bandwidth, and sample collecting speed.

In order to reduce the damage to implant location, the lateral dimensions of implantable electrodes are generally quite small (around scores of micrometers) and can be implanted into interested cerebral cortex, where it is quite close or even cling to the neurons. Good signal coupling will be achieved between the location of electrode recording site and nerve cell. Therefore, neuron's individual electrical activities can be well recorded or action potential will be created by stimulating the cell and changing membrane potential. When the dimensions of electrode recording site is small enough, as small as the neuron is, it will be able to record action potential of a single or several neurons, and provide more refined neural electrical activity information with more dimensions. Especially, the accuracy, spatial resolution, and information flux provided by array cortical microelectrodes with several channels are far beyond those provided by scalp electrodes.

With the improvement of electrode materials, processing techniques, and implanting operation methods, implantable electrodes are causing less and less

damage to our brain. Some new applications have been developed to help paraplegia or other disabilities. A 51-year-old woman who was diagnosed with amyotrophic lateral sclerosis (ALS) can type up to 24 characters per minute (Pandarinath et al. 2017) with a Utah electrode array implanted in the cerebral cortex to record her neural activity and a special designed system to decode her intentions.

To fulfill these requirements, many research institutes at home and abroad have fabricated implantable neural electrodes of various materials and structures. According to the number of channels, they can be divided into single electrodes and array electrodes; according to the locations, they can be divided into surface-contact electrodes which are put under scalp or on the surface of cortex, needle-shaped electrodes which stick into cortex or nerve tract, as well as cuff electrodes put around nerve tracts. According to the materials of electrodes, they can be divided into glass tube electrodes, metal wire electrodes, silicon electrodes, and polymer flexible electrodes. Now, more and more materials are used to make microelectrodes, including SiC, diamond, sapphire, SU-8, etc. Apart from glass tube electrodes and metal wire electrodes, other electrodes, typically silicon and polyimide electrodes, are made with microprocess techniques.

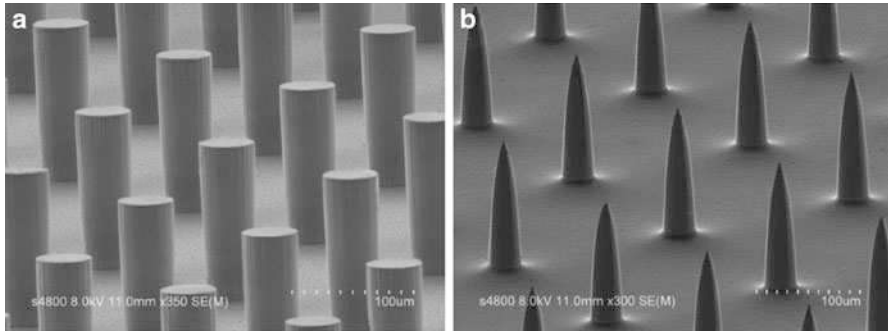
---

## Neural Electrode Fabricated by MEMS Techniques

Neural electrodes are sensors used to detect biosome EEG or nerve electrical activity signals. In order to reduce the interference in biosome's daily activities, electrodes are processed with MEMS technique, which is able to minimize the volume and weight of electrodes, improve the comfort of wearing, reduce the damage caused by implant operation, and increase the consistency and reliability of electrode devices. Following is an introduction of microprocess and fabrication techniques of several kinds of electrode devices.

## Skin Electrodes Based on Microneedle Array

Facility is important for daily use of an EEG signal recording system. To replace traditional Ag/AgCl electrode, which is inconvenient for long-time or daily use, a kind of microneedle array-based dry electrode is developed (Yu et al. 2011). Design of the microneedle is based on the structure of skin and the impedance needed for EEG recording. Skin consists of three layers: epidermis, dermis, and hypodermis (Fig. 1). The thickness of epidermis is varying from 50 to 100  $\mu\text{m}$ . The outermost layer of epidermis is SC which has a thickness of 10–15  $\mu\text{m}$ . It has high electrical resistivity. Other parts of epidermis under SC have low resistivity. Few blood vessels and nerves grow at dermis layer. The length of the microneedles is designed at a value that can penetrate the SC and stay its point at the dermis layer. Commonly, the length is around 50–100  $\mu\text{m}$ . The diameter of the microneedle is less than 100  $\mu\text{m}$  to minimize hurt to skin. The density of the microneedle array is 100/mm<sup>2</sup>.



**Fig. 4** SEM images (a) side view of pillars (b) side view of the silicon microneedle array

Dry electrodes were fabricated in three steps. The first step produces microneedle arrays. The second step deposits metal coating. The third step dices the wafer into appropriate size.

Four-inch double-sided polished, n-type silicon wafer,  $\langle 100 \rangle$  crystallographic orientation, 500  $\mu\text{m}$  thickness with 0.02  $\Omega/\text{cm}$  resistance was used to fabricate the dry electrode. Figure 3 shows the fabrication process. (1) All wafers were immersed in Piranha solution ( $\text{H}_2\text{SO}_4\text{:H}_2\text{O}_2 = 2\text{:}1$ ) at 125  $^\circ\text{C}$  for about 30 min followed by thoroughly rinsing with DI water. (2) 0.3  $\mu\text{m}$   $\text{SiO}_2$  was grown on the wafers in a thermal oxidation furnace at 1050  $^\circ\text{C}$ . (3) Then the  $\text{SiO}_2$  layer was patterned by hydrofluoric acid buffer. (4) Pillars were formed by deep reactive ion etching (DRIE) process. “BOSCH” process is used to form the pillars on the substrate. The selectivity between silicon and silicon oxide can be as high as 300:1 by the process. The height and diameter of the completed pillars were 50  $\mu\text{m}$ . The SEM image of pillars is shown in Fig. 4a. After the dry etching, an isotropic etching solution ( $\text{HF:HNO}_3\text{:CH}_3\text{COOH} = 1\text{:}3\text{:}8$ ) was adopted to sharp the tops of the pillars. The procedure of wet etching must be carried out in a vibration-free environment; otherwise sharp tips will not be formed (Normann et al. 1999). The image of microneedles are shown in Fig. 4b. Finally, 50/300 nm Ti/Au was deposited by sputtering on the microneedles and the backside of the wafers. Then the silicon wafers were diced into  $5 \times 5 \text{ mm}^2$ . The photo of the completed microneedle-based dry electrode is shown in Fig. 5.

To assess the efficiency of the dry electrode for EEG recording, impedance measurement and EEG-recording test are done. The measurement setup for evaluating electrode-skin-electrode impedance is shown in Fig. 6. The result is shown in Fig. 7. In order to compare impedances of the dry electrode with that of commercial electrode, the same test was carried out with commercial electrode. The result shows that the dry electrodes have similar impedances value to that of the commercial electrode at high frequency. While at low impedance, for example, at 10 Hz, the impedance value of dry electrode is 32  $\Omega$ , higher than that of wet electrode (18  $\Omega$ ). This value is acceptable for EEG recording.

EEG signals were recorded using dry electrode compared with the commercial electrode. The dry electrode was positioned at occipital region of a volunteer’s scalp.

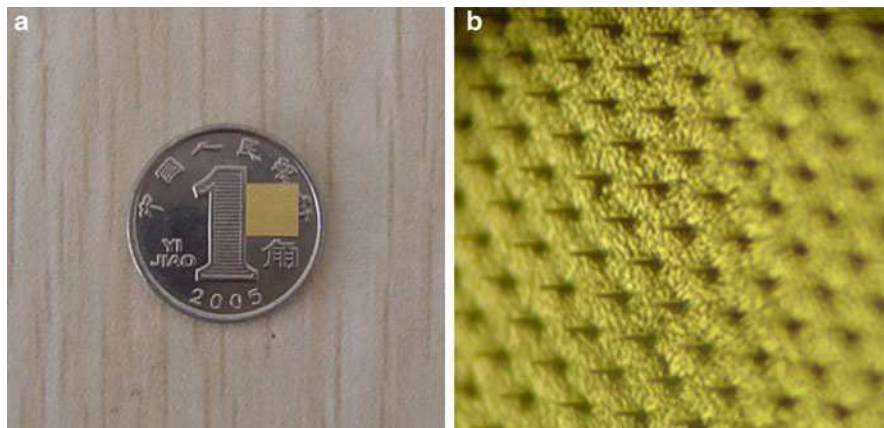


Fig. 5 (a) Photo of dry electrode. (b) Details of dry electrode

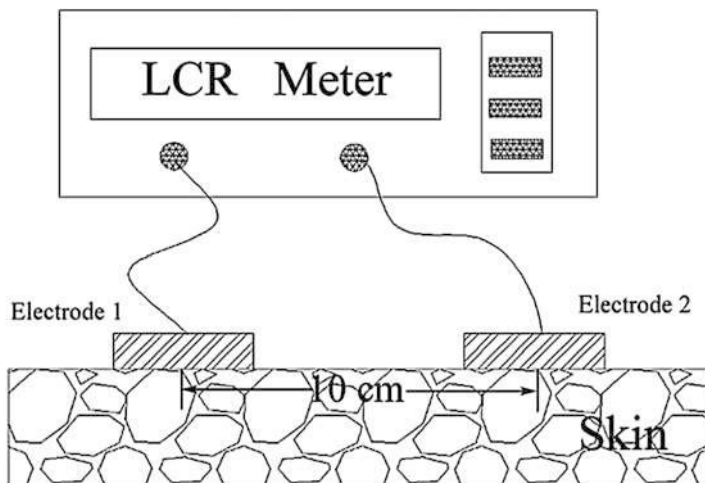
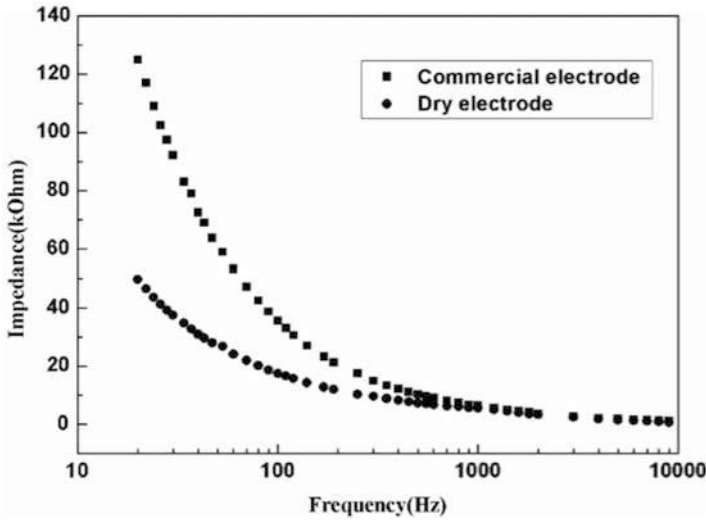


Fig. 6 Measurement setup for electrode-skin-electrode impedance

A commercial electrode was placed alongside the dry electrode as a control sample. Other two dry electrodes were put at the left and right ear lobes as reference electrodes.

Steady-state visual evoked potential (SSVEP) is recorded and analyzed to test and verify the quality of the EEG recorded by the dry electrode. When a person looks attentively at a continuously flashing light, at frequency range from 6 to 50 Hz, SSVEP can be evoked at the visual cortex and the flash frequency can be obtained by analyzing SSVEP. The dry electrode was placed at the  $O_1$  position (at visual cortex) for detection of SSVEP. A light-emitting diode (LED) modulated by a square wave was used as the flash light. The volunteer stared at the LED, which was flashing at a



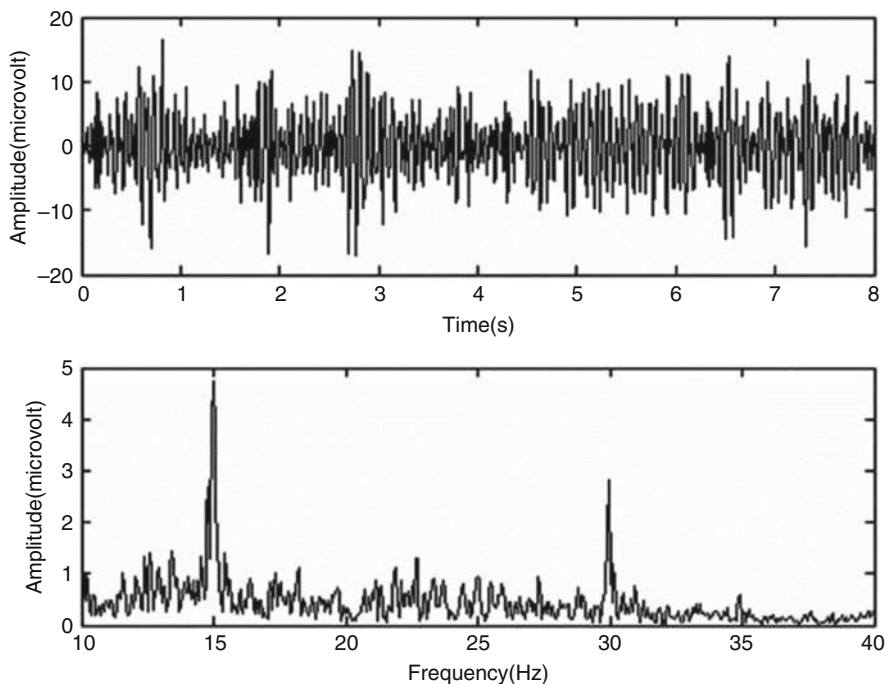
**Fig. 7** Impedance spectra of the dry electrode and commercial electrode according to the frequency changes

given frequency, for 30 s and the SSVEP were recorded during this time. Signals were preprocessed with a 50 Hz notch filter and a 4–50 Hz band-pass filter. Figure 8 shows a typical example of the temporal wave and frequency spectrum, the SSVEP induced by 15 Hz stimulation. The fundamental and second harmonics can be clearly identified at 15 and 30 Hz.

The test result verifies that the microneedle-based dry electrode performs well at EEG recording. However, this kind of silicon dice-based electrode wears uncomfortable. To improve wearing comfort and signal-to-noise ratio, especially at state of motion, flexible dry electrode is developed by replacing the silicon substrate with PDMS (Zhang et al. 2016). Based on conventional microneedles array, a layer of Parylene membrane is coated at the root of microneedles to insulate the electrode from the corneum layer. Only tips of the needles are exposed to make it contact with the stratum germinativum layer directly. As a result, skin potential will not be coupled into the recording electrode. Thus skin-potential variation will not influence the biopotential recording in the dynamic state. (Pei et al. 2017). The surface of the microneedle array was modified with PEDOT to reduce the impedance of the microneedle array-based dry electrode (Chen et al. 2013a).

## Metal Wire Implantable Electrodes

Insulation materials and conductive materials are two kinds of basic materials for electrode fabrication. Currently, metal wire electrode is widely used in extracellular recording. Metal wire microelectrode is the simplest extracellular recording



**Fig. 8** EEG recorded by microneedle array-based dry electrode. Temporal wave (*top*) and frequency spectrum (*bottom*) of SSVEP induced by a 15 Hz flash. The fundamental and second harmonics can be clearly identified at 15 and 30 Hz

microelectrode, and also one of the earliest used electrodes by scientists. It is made of a piece of metal wire covered with insulation materials. Only a tip is exposed as neural electrode recording site. Generally, glass, ceramic, polyimide (polydichlorstyrene), PTFE, or poly-p-xylyene are selected as insulation materials for electrodes, because they have good insulation characteristic and biocompatibility. Conductive materials are usually metal materials with stable chemical properties and will not be corroded in body fluid. Precious metals such as nichrome, tungsten, gold, platinum, and iridium are frequently used electrode materials. Multichannel recording requires arranging a bunch of metal wires into electrode array. In order to minimize damage, the diameter of metal wire should be as small as possible. However, the diameter of metal wire cannot be too small; certain solidity should be kept to ensure smooth implanting. Currently, the diameter of the thinnest metal wire in microelectrode array is around 12  $\mu\text{m}$ , which is generally made of nichrome or platinumiridium.

Although the metal wire electrode has simple structure and is convenient for use, establishing an electrode array with metal wire is a hard work, such as precise position of metal wires, controlling over the exposed area of recording site at the tip of metal wire, as well as insulation process of metal wire. Another disadvantage of metal wire electrode is every recording site requires a metal wire, and as a result,

too much wires are there in an electrode array. These disadvantages seem more obvious especially when manufacturing high-density electrode array.

As to manufacturing electrode array by metal wire, various research institutes adopt different arrangements of recording sites to meet their own requirements. Some electrodes arrange recording sites longitudinally along the electrode body (Barna et al. 1981). The other way of arrangement is horizontal juxtaposing. All the metal wires are parallelized with certain interval between adjacent wires. In order to control the distance between electrodes, various positioning devices are adopted, including spring, precast positioning grid, and even pattern dies (Takahashi et al. 2005). As most of the process have to be completed by hand, it is hard to ensure the accuracy of the distance between electrodes and consistency of electrodes performance. When the size of array gets large, it will be much harder to arrange electrodes by hand.

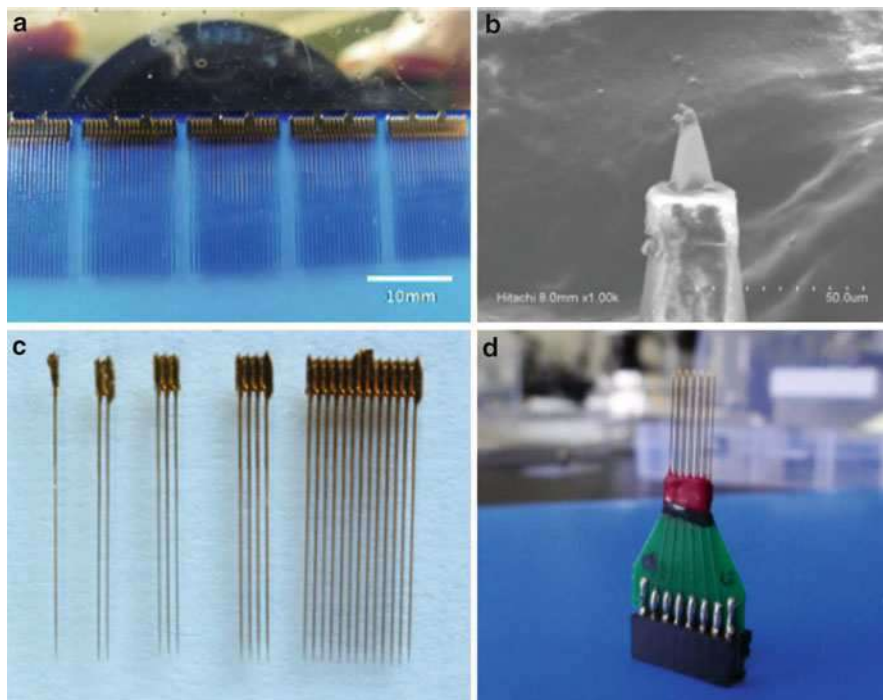
To simplify the assemble problem met in manufacturing metal wire microelectrodes, some research institutes carry out the method of processing electrodes with a metal block. Fofonoff and other coresearchers from MIT adopt spark discharge and wire cutting (Fofonoff et al. 2004 ) to process electrodes. They processed an electrode array of titanium needles on a titanium block and then sharpened the tips of electrodes through chemical corrosion. This kind of process makes metal wire electrode array manufacturing easier. However, the processes of electrode insulation and lead connection of the electrode array are comparatively complex. Our research team puts forward a new method: preparing linear arranged wire microelectrodes with semiconductor silicon material. We adopt photolithography and wet etching process to prepare silicon into silk-shaped microneedle electrodes. Different from metal wire electrodes, these electrodes fabricated with microprocessing methods feature good consistency and precise intervals between electrodes. According to actual requirements of different applications, users can select proper number of electrodes and assemble them into an electrode array. Then, to achieve insulation between electrodes, we precast breaking points at the roots of electrodes and break cleavage by taking advantage of the brittleness of silicon, as shown in Fig. 9.

Photolithography method helps define the dimensions of each silicon microneedle. Mature silicon wet etching process is used to prepare electrode array with more compact structure, better consistency and repeatability, as well as higher array density.

## Processing of Bulk Silicon Microelectrodes

Among all the electrodes fabricated on the basis of microprocess, silicon microelectrodes are the most representative ones. Early in 1970s, some international research institutes have begun to process microelectrode array with microprocess technology. The most successful one is Utah electrodes (Campbell et al. 1991) which is made of bulk silicon. Utah electrode is named after the research institute that developed it. Utah electrodes are 2D electrode array only with recording sites on the tips. Utah electrodes are processed in this way: first, work on the bulk silicon material and





**Fig. 9** Silicon-based precast microwire electrode array

process needles by adopting mechanical cutting and chemical corrosion methods; then achieve insulation and separation between needles by using semiconductor PN junction or glass. The length of electrode body is around 1 mm. Recording site is located at the tip and the shape of tip is formed through chemical corrosion method. Intervals between electrodes are around hundreds of microns. An electrode array may consist of up to 100 electrodes. The following are the steps: First manufacture the silicon needle array. Select N-type (100) crystalline silicon substrate with thickness of 1.7 mm, resistivity ranging from 6 to 20  $\Omega/\text{cm}$ . Pillars mixing with Al P+ are created on the N-type silicon wafer. Then, lay the aluminized side of the wafer down, separate each pillar mixing with Al P+ with a wafer scribe. The array consisted of these pillars will be arranged on the remained N-type substrate with thickness of around 0.2 mm. With wet etching process, all the pillars are etched into tapers (thick at the bottom end and thin at the upper end). Finally, plate metal on the needle tip and insulate the other part of the needles with polyimide.

Wire bonding is used to extract the electrodes. Polyimide-insulated metal wires (length: 6 cm, diameter: 36  $\mu\text{m}$ ) are bonded on the bottom of each needle electrode. Then cover the needles with polyimide. Thus, stable, reliable, and sealed electric connection between electrodes and lead is established. The other end of the lead is connected with the electrode cap which can be fixed on skull.

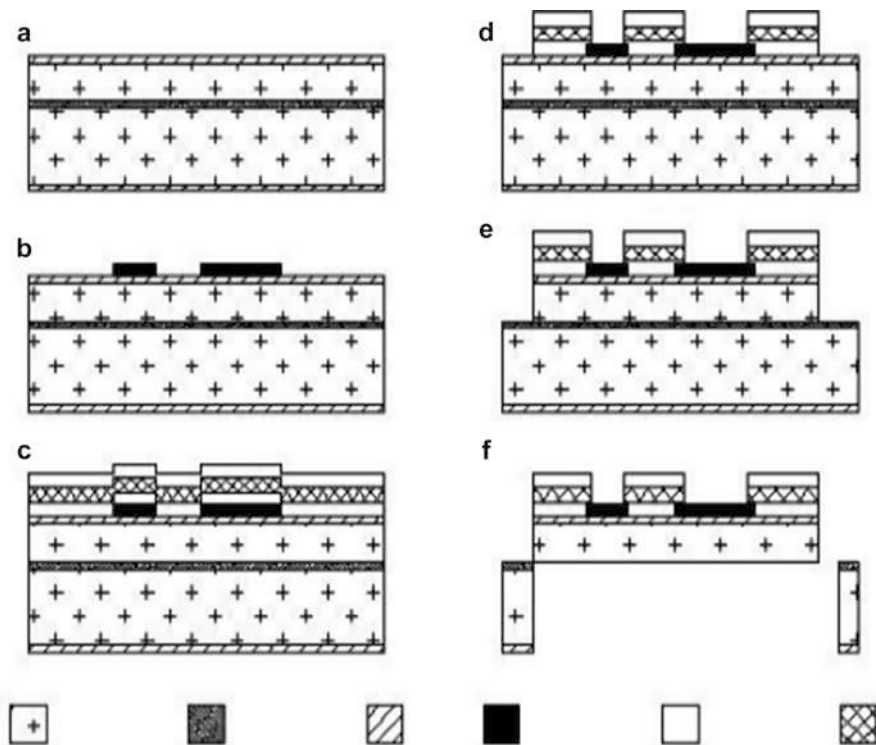


Compact in volume (accounting for only around 2% of the implanting area) and light in weight, and with a lead of proper flexibility, the electrodes can “float” on the surface of cerebral cortex and move with the motion of cortex, which reduces the relative motions between electrode fix device and the cortex. Putting the electrodes in this way will prolong the lifespan of electrodes working in brain. Today, Utah electrode is still one of a few kinds of electrodes that have been approved by FDA to be used to record human beings’ cortex signals.

## Processing of Microelectrodes Based on Thin-Film Silicon

Michigan electrode (Wise et al. 2004) is the most representative electrode manufactured with silicon planar process. Different from Utah array, Michigan electrode learns more from microelectronic technique. It looks like a thin sword with electrode recording sites arranged on the surface. Usually, the width of Michigan electrode ranges from 10 microns to over a 100 microns, and the thickness is only 10 microns or even less than 20 microns. Traditionally, the needle body of Michigan electrode is manufactured with boron diffusion plus selective etching. The steps include: first, thermally oxidize the  $\text{SiO}_2$  wafer; then define the diffusion area (for example, the shape of needle body) on the thermal oxide layer via photolithography method; after removing the thermal oxide layer of this area with HF acid, put the wafer into diffusion furnace for deep boron diffusion (concentration  $>10^{19}$ , diffusion depth equals the thickness of needle body). After the diffusion, remove the thermal oxide layer (used as masking) with HF acid. Then grow a set of stress-free dielectric film made of  $\text{SiO}_2/\text{SiN}/\text{SiO}_2$  by adopting the method of plasma-enhanced chemical vapor deposition (PECVD). Next, sputter or evaporate Ti/Au on this layer of dielectric film and etch metal wire. With the same method, grow another layer of dielectric film without stress, then well cover the metal wire under the dielectric film. Etch the dielectric film layer at where the recording sites and wire pad should be located; the metal layer will expose once the through-hole is formed; sputter Ti/Au or Ti/Ir on the second layer of metal, and create recording sites and bonding sites at through-hole locations through photolithography and etching process. Finally, once these steps are completed, put the silicon wafer into EDP solution (silicon anisotropic etching solution consisted of orthocatechol, ethidene diamine, and water) for etching. As the etching speed of silicon heavily doped by boron is much less than that of silicon lightly doped, the final etching-released structure is the part doped via boron diffusion at the beginning, which is the shape of the needle body. The feature of Michigan electrodes’ structure is that many recording sites are arranged on the same electrode needle body, which is quite suitable for high-density and high-flux recording. After being assembled, this kind of electrode array will be able to offer 256 channels and channel density will reach 12 channels/mm<sup>3</sup>.

Another widely used method to process Michigan electrode is using the silicon on insulator (SOI) (Chen et al. 2014). To adopt this method, customized SOI silicon wafer with certain thickness of top silicon layer is needed. Here, we take 15-micron-thick silicon electrode preparation as an example. First, customize SOI silicon

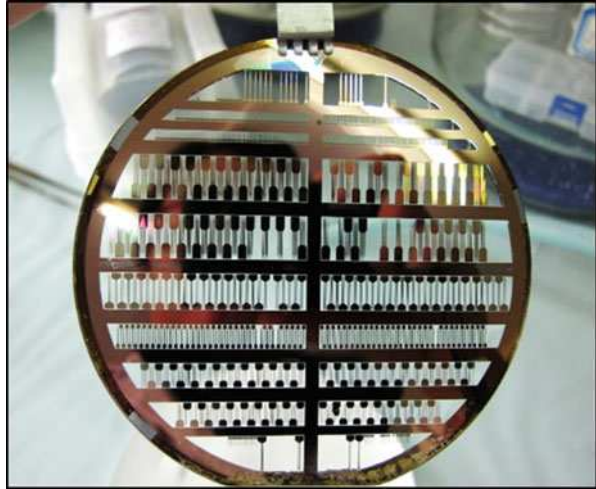


**Fig. 10** Schematic illustration of the fabrication process of microelectrodes (not to scale)

materials with 15-micron-thick top silicon layer, 1-micron-thick buried oxide layer, and 300-micron-thick handle wafer. The following process can be divided into two steps: front-side process and backside process, as shown in Fig. 10. Front-side process is quite similar with the preparation process of Michigan electrode: (a) deposit a dielectric film (consisted of SiO<sub>2</sub>/SiN/SiO<sub>2</sub>, thickness equals 1 μm) without stress on one side of SOI device layer; (b) deposit Ti/Au/Ti on it and create electrode wire through patterning; (c) deposit a layer of SiO<sub>2</sub>/SiN/SiO<sub>2</sub> dielectric film without stress and create top insulation layer; (d) mark on the top insulation layer, etch windows at where recording sites to be formed, and let the metal layer below expose; (e) taking thick resist or aluminum as marking, adopting ICP deep etching process, and etch through top and bottom insulation. Thus, the front-side process is completed.

The purpose of backside etching is to release the silicon electrode device defined by the device layer from the substrate. Such release can be divided into two kinds: wet and dry release; the latter is used more frequently. When adopting dry etching, thick resist or thermal oxide layer will be used to define a framework at the backside of SOI wafer. The framework is to connect the microneedles after release and provide convenience for the follow-up cleaning and equipment process. Then,

**Fig. 11** Silicon neural electrode array fabricated from SOI



remove the substrate silicon outside of the framework with ICP etching method and stop at buried oxide layer. Finally, remove the silicon dioxide from buried oxide layer via ICP etching. The completed film silicon neural electrodes are showed in Fig. 11.

### **Advantages and Development of Silicon-Based MEMS Microelectrodes**

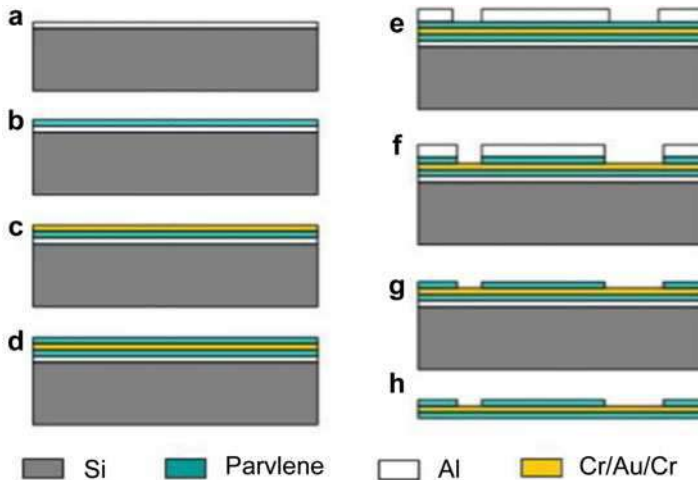
Adopt mature semiconductor materials and MEMS process technique; the above-mentioned two kinds of electrode arrays are far ahead of metal wire electrode process in the aspect of accuracy and consistency. Michigan electrode, especially, is able to assemble silicon needles with multirecord channels into 3D electrode array (Takahashi et al. 2005) as Utah electrode is. When such kind of electrode array is implanted into cortex, the recording sites distributed in three dimensions will help it collect neuron signals from certain area of brain in a denser way, which is quite important for brain function research and discovering neural coding mechanism. Because in our brain, the number of neurons reaches 10 billion magnitude, while currently the electrodes can only record activities of tens or hundreds of neurons. The record capacity is a far cry from being able to comprehensively reflect the actual activities of our brain, like the distance between several pixel and a high-definition movie. In the brain program carried out by USA in 2015, researchers set a target of recording a million of neurons. Imaging every spike from every neuro, as proposed by Scholvin (Scholvin et al. 2016) it will be quite helpful for both brain science research and brain-computer interface technologies, which are under development. However, such requirements mean great challenge to electrode technology. Apart from further improving process accuracy to reduce the sizes of devices, lead connection number arising from increase of electrode channels is another problem

to be solved when researching electrodes with high record density and high number of record channels.

Silicon is a good material for electrode. It not only features diversified means of process and low cost, but it also can achieve monolithic integration between electrode- and silicon-based amplifying circuit and signal process circuit. EEG signals are weak and the background noises are complex. If it is possible to integrate preamplifier and electrode, the SNR of neural signals will be largely improved. Such kind of integration simplifies the volume and complexity of the system, improves the efficiency, and reduces noises. Early in the design phase, researchers started trying coordinated design and process of Michigan electrodes and circuits and announced 3D high-density electrode array with up to 1,024 channels (Wise et al. 2004). In recent years, in order to increase the density of recording sites arranged on silicon electrodes and the number of recording channels, Jörg Scholvin and coresearchers from MIT, using electron beam lithography, have worked out high-density and high-flux electrodes with 1,000 recording sites (Scholvin et al. 2016). The electrode recording sites are connected by wires whose diameter is as small as 200 nm (and the interval between these wires is only 200 nm). These recording sites are distributed at five pieces of needles with interval of 500  $\mu\text{m}$ . The width of each needle is less than 100  $\mu\text{m}$  and the recording sites are squares with side length of 9  $\mu\text{m}$ . Center-to-center spacing between recording sites is 11  $\mu\text{m}$ . In 2016, Carolina Mora Lopez and coresearchers (Lopez et al. 2016) from IMEC reported an integrated electrode which includes 966 recording sites and integrates 384 amplifying circuit channels able to gate. The electrode is fabricated with SOICMOS process. The electrode tip is only 70  $\mu\text{m}$  wide and 20  $\mu\text{m}$  thick. Size of each recording site is 12  $\mu\text{m} \times 12 \mu\text{m}$ . In order to improve the SNR, a preamplifier is equipped under each recording site.

## Process of Flexible Polymer Microelectrodes

Neural microelectrode devices can be regarded as a bridge between artificial electronic equipment and biological neural system. Not only the sensing and transferring functions of the electrode devices should meet the requirements of signal process, the biocompatibility of electrode devices should also be good. Electrodes' materials are required with good biocompatibility, such as platinum, silicon, and  $\text{SiO}_2$ ; issues about biocompatibility of these materials will emerge in long-term service period because of solidity mismatch. So far, long-term effective implantable microelectrode for neural activity record has not been available. After a period of time (generally no more than 2 years), the record capability of implantable neural microelectrode will reduce. The most important reason is that the solidity of electrode materials does not match the neural tissue around the implants. The materials of implantable microelectrodes are usually hard materials, such as platinumiridium, stainless steel wire, silicon, and so on. Although these materials pose no toxic and will not trigger inflammation to surrounding tissues, solidity mismatch will cause wrapping and proliferation of surrounding neural glial cells as time goes by. Thus, the active neurons will get far away from the record electrodes, and the latter will not be able



**Fig. 12** Process flow chart of parylene-C flexible electrode

to record any signal. Experimental studies prove that through proper structure design, flexible electrodes made of flexible materials can relieve the wrapping reaction of tissues after being implanted and prolong the working hours of electrodes in our bodies. Electrodes with good flexibility can be fabricated with substrates that are made of polymers other than silicon.

Process of polymer microelectrode is similar to that of Michigan silicon microelectrode. Generally, it is hard to conduct photolithography and etching process on the surface of polymer film directly like what we do with silicon wafer. The reason is that the rigidity of polymer materials is not large enough. To be more convenient for operation, we attach polymer film on a flat substrate, usually it is a standard silicon wafer. When completed, we manage to peel the well-fabricated polymer microelectrode device off the wafer. Now, let us take chemical vapor-deposited parylene-C as an example to introduce the method of processing parylene-flexible electrode.

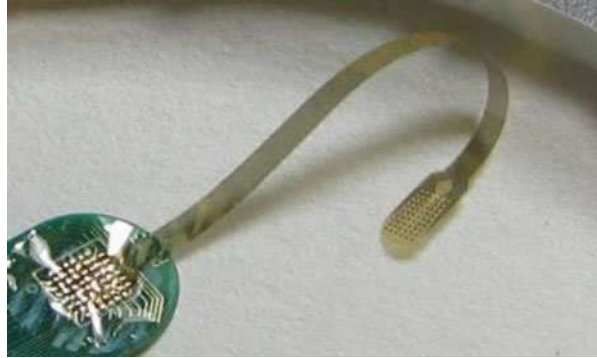
- (1) Clean the silicon wafer and vapor 300-nm-thick metal Al over it as sacrificial layer, as shown in Fig. 12a.
- (2) Using chemical vapor deposition (CVD) method, form around 10 μm parylene-C film on the Al-plated substrate as the bottom insulation layer of the microelectrode, as shown in Fig. 15b. If the prepolymer of polyimide is fluid, use spin coating plus baking film formation method to prepare bottom insulation layer on the Al-plated substrate.
- (3) Create the metal layer over the bottom polymer material via evaporation or sputtering, Cr/Au/Cr:70/200/70 nm. Create electric pole, metal wire, and welding points via photolithography, as shown in Fig. 12c. It should be noted that the surface energy of polyimide material is usually low and the adhesion formed between polyimide material and the metal plated on it is not so strong. Before deposition, it is better to launch silane treatment upon the parylene

surface. The purpose of silane treatment is to enhance the adhesion between parylene and other materials (such as Si, SiO<sub>2</sub>, and metal with oxide on the surface). It should be noted that parylene substrate surface after silane treatment should be plated with another layer of material as soon as possible, or the surface activity will get worse as time passes. In recent years, some researchers suggest to treat the surface of polyimide with oxygen plasma to improve the surface activity of polyimide, increase the adhesion between polyimide material and metal. The method works fine.

- (4) Form another layer of parylene film (10 μm thick) via CVD method as the top insulation layer of microelectrode, as shown in Fig. 12d. If liquid prepolymer of polyimide is used, the top insulation layer can also be prepared via spin coating and baking.
- (5) Etch electrode recording sites and Pad used to connect electrodes with oxygen plasma. With oxygen plasma etching, the etching speeds over gel and polyimide are almost the same. Photolithography resist of a little bit thicker than the top insulation layer is needed for masking. Or, we can evaporate another layer of Al with thickness of 500 nm and transfer the masking pattern on to the Al layer with the help of photolithography resist. Then make masking with Al and conduct oxygen plasma etching, as shown in Fig. 12e, f.
- (6) Etch the overall profile of flexible electrode body. Remove the electrode created when etching top insulation layer, or the remained resist of Pad, or Al mask. Adopt the method similar to Step 5: take Al as the mask, conduct oxygen plasma etching, and etch through the top and bottom insulation layers, as shown in Fig. 12g.
- (7) Finally, soak the silicon substrate for preparing flexible electrode into NaCl solution. Dissolve away the Al between the silicon substrate and flexible electrode via electrolysis method and release the final electrode, as shown in Fig. 12h. The well-fabricated electrode is shown in Fig. 13.

Because flexible electrodes cannot stick into tissues as silicon or metal wires do, this kind of film electrodes, for a long period, can only be implanted in the surface of retina, cortex or twine around nerve tract. With the help of additionally designed rigid implant tool, flexible electrodes can also be implanted. However, as the rigid implants have to be withdrawn, the implanted electrodes are not stable. Recently, through temperature transferring, material degradation, or dissolution, researchers have designed a transition for flexible electrodes, which is able to dissolution, researchers have designed example, turn the flexible electrode rigid through coating dissolvable maltose on the surface, then implant it into body; as the maltose dissolving, flexibility of the electrode restores gradually (Xiang et al. 2014). This method presents a new direction for further development of flexible electrodes. However, so far no long-term record by flexible electrode implanted via rigid transition has been reported. As to development of completely flexible electrodes, in 2015, *Nature Neuro Science* reported an ultrathin flexible electrode called Neuro Grid (Khodagholy et al. 2015). Over a 100 electrodes (10 μl<sup>2</sup>) are integrated in an area of several square millimeters. The device successfully recorded the action

**Fig. 13** Photo of a well-fabricated parylene-C flexible electrode



potentials emitted by neurons located at around  $200\ \mu\text{s}$  under the cortex. It means flexible electrode will be able to record the electrical activities of neurons located at certain depth within cortex, with no need to be stuck between cortices, as flexible electrodes can well couple with cortex. Early in 2017, MIT EECS reported in *Nature Neuroscience* a flexible electrode device (Park et al. 2017) which integrates light source, microfluidic channel, and microrecording electrode. The lateral dimension of the device integrating three functions is only  $200\ \mu\text{m}$  and the weight is only 0.5 g. Completely made of polymeric materials, the device has good flexibility, and is quite excellent in biocompatibility. These advantages present promising application prospect. Also at the beginning of 2017, Lan Luan and coresearchers from The University of Texas at Austin reported in *Science Advances* a kind of linear flexible electrodes (Luan et al. 2017) made of polymeric materials. The lateral dimension of the ultrathin flexible electrode almost equals to the size of neuron and the Young modulus is also similar to those of cells. After being implanted into biosome for 4 months, no tissue wrapping was found in the electrodes. Also, the electrodes featured good record characteristics and biocompatibility. The latest research results show that flexible electrodes become the key development orientation of multichannel recording microelectrodes, especially at the aspects of reducing tissue wrapping and ensuring effectiveness of long-term implanted neural electrodes.

---

## Interfacial Modification of Neural Electrodes

Electrode recording site/biosome interface is the key for coupling between neural electrodes and neural information. In order to make electrode recording signals more sensitive, and improve electrodes' capacity of signal transferring, researchers coat material layer with higher ion exchange capacity on the surface of recording sites or plate a nanometer interface layer with larger surface area. This method will greatly improve the SNR of electrodes, reduce electrode impedance, and lift the charge transferring capacity of stimulating electrodes.



Now let us illustrate the importance of interfacial modification in electrode processing with an example using interfacial modification to improve the characteristics of surface electrodes and implantable electrodes.

Since the electrochemical properties of an electrode are mainly dependent on the materials present on the interface, a common approach to bring down the impedance of electrodes with small sites is to modify the surface of the electrode site with suitable materials. Conductive polymer (CP) coatings have been widely used to be an enabling technology for smaller electrode designs (Kozai et al. 2012). The reason of high conductivity of conductive polymer lies in a conjugated backbone with a high degree of  $\pi$ -orbital overlap that can be subjected to oxidation or reduction by electron acceptors or donors, resulting in p-type or n-type doped materials, respectively. By changing dopant concentrations, electrical conductivity can be enhanced by as much as 15 orders of magnitude.

To improve the electrical performance and biocompatibility of extracellular electrode, PEDOT and MWCNT are deposited respectively or simultaneously on the recording sites of film silicon-based electrode (Chen et al. 2013b). Carbon nanotubes are doped into PEDOT polymerization and are electrodeposited on neural microelectrode arrays. Neural microelectrode arrays were fabricated in wafer scale on SOI substrate.

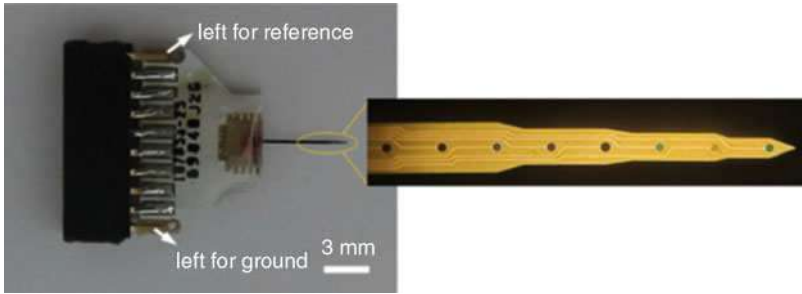
An aqueous solution containing 0.1%wt EDOT, 0.2%wt PSS, and 2 mg/mL CNTs in deionized water were dispersed under ultrasonic irradiation for 1 h to form a fully electropolymerized PEDOT. Electrochemical deposition of PEDOT/CNTs was accomplished on the microelectrode sites using electrochemical workstation (CHI 660D). In the galvanostatic experiments, various current densities were applied to optimize electrodeposition condition, while the total amount of transferred charge was kept constant for different current densities. After the current density is decided, the film thickness was controlled by adjusting electrochemical deposition time. In all electrochemical experiments of this study, electrodes were mounted in a three-electrode cell with an Ag/AgCl electrode acting as a reference electrode and a large-area Pt electrode as a counter electrode.

Electrochemical impedance spectroscopy (EIS), cyclic voltammetry, and potential transient method were used to assess the electrode–electrolyte interface of the modified neural microelectrodes. Based on the results of the characterization, the electrical parameters of the deposition process are optimized.

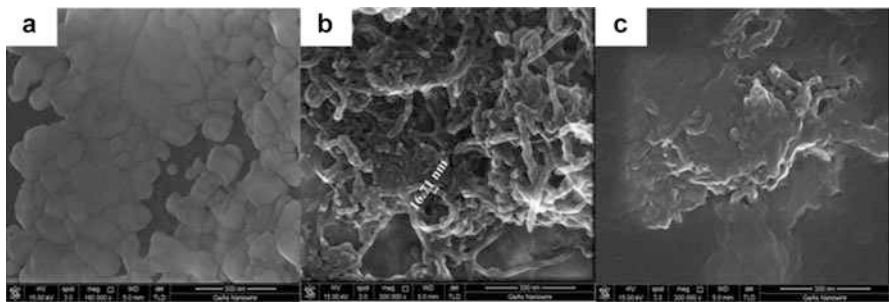
Optical micrographs show one of a well packaging microelectrode arrays coated with the composite film, as can be seen in Fig. 14. The color of recording sites is different with different material and thickness. The SEM image in Fig. 15 reveals that PEDOT/MWCNT hybrid surface (Fig. 15c) exhibits a different microstructure from that of either PEDOT (Fig. 15a) or MWCNT (Fig. 15b).

Electrochemical impedance spectra (EIS) were performed to characterize the electrode–electrolyte interface of the PEDOT/MWCNT composite microelectrode. EIS before and after surface modification with PEDOT/MWCNT film is shown in Fig. 16. The impedance of both bare gold and modified electrode sites decreased with increasing frequencies, while the magnitude of PEDOT/MWCNT film was significantly lower than that of bare gold electrode over the whole frequency range





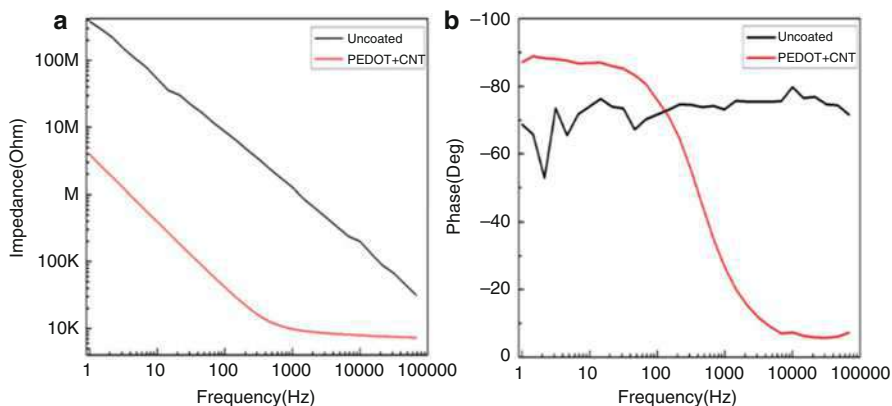
**Fig. 14** Photograph of silicon electrode modified with PEDOT/CNT



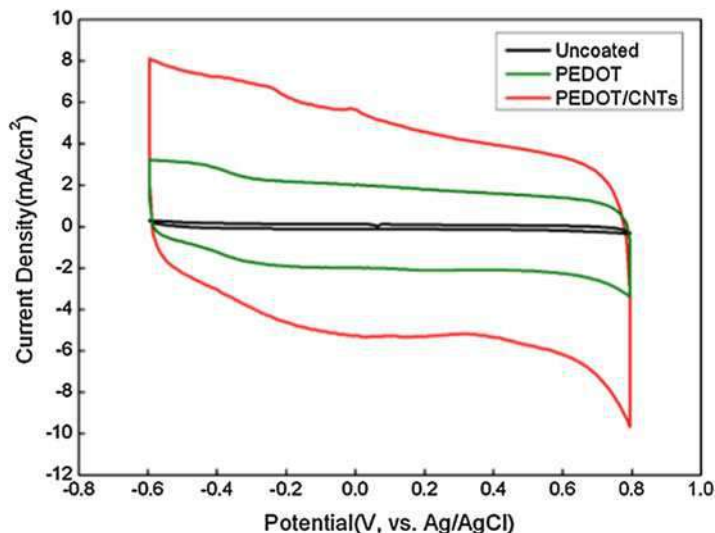
**Fig. 15** SEM images of different coated layer: PEDOT (a), CNTs (b), and PEDOT/CNTs composite film (c)

from 1 to 10 kHz. At 1 kHz, which is the typical frequency of neural biological activity, the magnitude of a coated electrode was only 12 k $\Omega$ . Compared to the impedance of bare gold (1.3 k $\Omega$ ), the impedance PEDOT/CNT reduced over two orders of magnitude. This phenomenon is due to the significant increase of the effective surface with the formation of PEDOT/MWCNT film. The phase angle for PEDOT/MWCNT-coated microelectrode at 1 kHz is 20°, while the phase angle of bare gold is about 75°.

Cyclic voltammetry (CV) was used to explore the capacity of charge transfer density for PEDOT that was deposited with the same charge density. Compared to the bare gold electrode, a significant hysteresis area can be observed. The capacity of charge transfer observably increased using PEDOT/MWCNT, as can be seen in Fig. 17. The integration of I(t) within the cycled region gives the charge storage capacity (CSC) of composite film. Obviously, the charge storage capacity of the PEDOT/MWCNT-coated microelectrode is significantly higher than that of bare gold electrode, and also higher than PEDOT-coated ones. When being used as stimulating electrode, high CSC means the electrode is able to transfer more electric charge to neuro under the safety voltage.



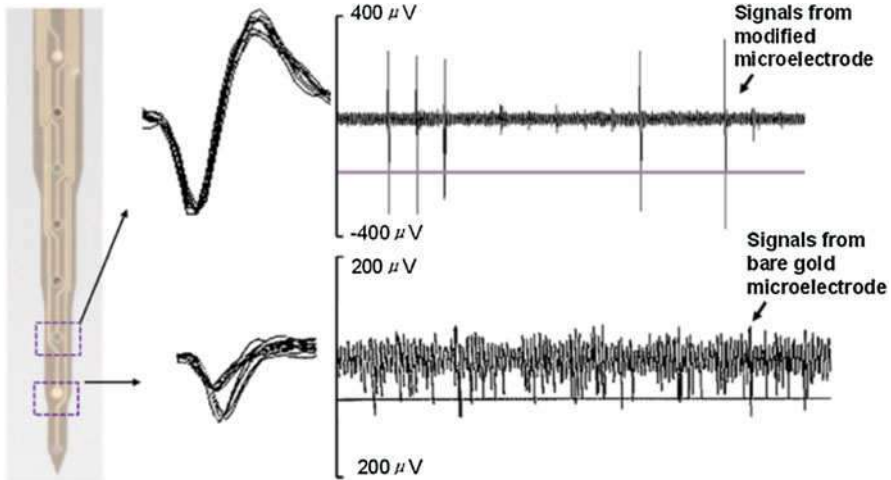
**Fig. 16** Electrochemical impedance (*left*) and phase (*right*) spectra of the bare gold electrode and PEDOT/CNT-modified electrode



**Fig. 17** Cyclic voltammetry curve of electrode site modified with different materials

A microelectrode with bare gold and PEDOT/MWCNT-coated microelectrode is implanted into Sprague-Dawley rats to evaluate the recording performance. Tucker Davis Technologies (TDT) multichannel neural acquisition processor is used to record the neural spike signals. Acoustic stimuli with different amplitude and frequency were delivered through an electrostatic speaker 25 cm far away from the animal's ear. All stimuli were delivered in an open field in an anechoic chamber. Neural signals recorded were amplified (5000 $\times$ ), filtered (0.3–5 kHz), and digitized at 25 kHz.

The microelectrode array is advanced into auditory cortex by a hydraulic micro-drive (FHC, Bowdoin, ME, USA). When penetrating into proper region, all the eight



**Fig. 18** Signal-to-noise ratio of PEDOT/CNT-modified electrode and bare electrode in vivo

channels of microelectrode arrays can simultaneously record the neural signals response to the stimuli. The signal-to-noise ratio can be expressed as below:

$$SNR = 10\log_{10} (V_s^2 / V_n^2)$$

Compared with the spike signals recorded using bare gold microelectrode arrays, PEDOT/CNTs-coated microelectrode arrays can achieve a higher signal-to-noise ratio >15 dB, while the bare gold microelectrode showed a relatively poor SNR of 6 dB, which demonstrated that the modified microelectrode arrays showed an obvious improvement in signal-to-noise ratio of recording, as shown in Fig. 18.

Iridium oxide is another material which can decrease the impedance dramatically when being used as a thin layer of interface material on the electrode. To fabricate the iridium oxide film compatible with the wafer-level process of film silicon neural electrode, the quality of iridium oxide film must be able to withstand the mechanical and chemical impact of postprocess and simultaneously achieve good performance as a neural electrode. Iridium oxide-fabricating process combined with fabrication flow of silicon electrode (as introduced in section “[Surface Neural Electrodes](#)”), at wafer level, is developed to produce film silicone neural electrode with iridium oxide-modified recording site. Compared with bare golden electrodes and poly(3,4-ethylenedioxythiophene)-coated electrodes, iridium oxide electrodes fabricated with this method exhibit a particular low open-circuit potential of 0.012 V, low impedance of 386 kΩ at 1 kHz, high safe charge storage capacity of 3.2 mC/cm<sup>2</sup>, as well as good impedance consistency of less than 25% fluctuation.

Modification method also can be used to decrease the impedance of skin electrode based on microneedle array (Chen et al. 2013a). PEDOT is coated at the surface of the microneedle array to improve its electrochemical performance. The mean impedance magnitudes (measured from skin) of PEDOT dry electrode can be decreased to lower than 10 kΩ at 100 Hz with a contact area of 6 mm × 6 mm.

## Challenges Neural Microelectrodes Facing With

Although neural interface devices have been applied in clinical medicine, the performance of existing devices is far from meeting the requirements of actual application. Taking brain-computer interaction as an example, many aspects of surface electrode devices need to be improved, including the electrodes being inconvenient for wearing, too much interference, and the spatial and time resolution of control signals being limited. As to implantable electrode devices, the performances of recording implantable neural electrode devices are also far away from meeting the actual requirements, especially when it goes to the number of implantable electrodes and biocompatibility.

Requirements of implantable neural microelectrodes upon high flux and the significance of relevant research: First, viewing from the perspective of neuroscience research, a huge amount of neurons should be recorded at the same time if we want to learn the way of operation of hundreds of billions of neuron networks (the number of neurons in human beings' brain is approximately in the  $10^{12}$  order of magnitude) in our brain, as well as the mode of operation and work of neural circuits. Among the Innovative Brain Research Technology programmed by the US Human Brain Project, one of the research objects is "recording (the activities of) a million neurons at the same time." To record simultaneously the activities of a huge amount of nerves is helpful for neuroscience researchers to learn the mechanism and internal relationship of a neurological cluster, and find the relationship between neural activity mode and behaviors in a comprehensive and detailed way, and finally discover the mechanism of higher nervous activities such as cognition, memory, and emotion. Secondly, viewing from the perspective of requirements of neural engineering technologies, taking implantable brain-computer interface application as an example, more electrodes need to be implanted to improve the dimension and robustness of control signals. In view of the existing conditions, with help of micro-nanofabrication and IC fabrication, some research institutes have been increasing the number of recording channels, and reducing the volume needed for implanting and signal output. However, the number of recording channels of a single electrode is still in the 1000 order of magnitude.

Viewing from the perspective of biocompatibility, even slight wrapping will disable the electrode recording as the electrode recording sites become far away from the neuron emission. Although the biocompatibility of electrode will be effectively improved and electrode wrapping will be relieved when the dimensions of electrode are reduced and flexibility of electrode is increased, electrode used to record emissions of single neuron must keep certain dimension; if it is too small, the SNR of recorded point activity signals will be too small to distinguish. As to electrodes made of flexible materials, although they have good biocompatibility in short term, the ultrathin flexible electrodes made of polymeric materials cannot ensure long-term sealing reliability. Therefore, we should dedicate more to the research of materials, structures, and processes for neural interface devices.

## References

- Barna JS, Arezzo JC, Vaughan HG Jr (1981) A new multielectrode array for the simultaneous recording of field potentials and unit activity. *Electroencephalogr Clin Neurophysiol* 52:5494–5496
- Campbell PK, Jones KE, Huber RJ, Horch KW, Normann RA (1991) A silicon-based, three-dimensional neural interface: manufacturing processes for an intracortical electrode array. *IEEE Trans Biomed Eng* 38(8):758–768
- Chen Y, Pei W, Chen S et al (2013a) Poly(3,4-ethylenedioxythiophene) (PEDOT) as interface material for improving electrochemical performance of microneedles array-based dry electrode. *Sensors Actuators B Chem* 188(11):747–756
- Chen S, Pei W, Gui Q et al (2013b) PEDOT/MWCNT composite film coated microelectrode arrays for neural interface improvement[J]. *Sensors Actuators A Phys* 193(15):141–148
- Chen SY, Pei WH, Hui Z et al (2014) 32-site microelectrode modified with Pt black for neural recording fabricated with thin-film silicon membrane. *Science China Inf Sci* 57(5):1–7
- Fofonoff TA, Martel SM, Hatsopoulos NG et al (2004) Microelectrode array fabrication by electrical discharge machining and chemical etching. *IEEE Trans Biomed Eng* 51(6):890–895
- Humayun M, de Juan E Jr, Weiland J et al (1999) Pattern electrical stimulation of the human retina. *Vis Res* 39:2569–2576
- Jung Y, Kwak JH, Kang H, Kim WD, Hur S (2015) Mechanical and electrical characterization of piezoelectric artificial cochlear device and biocompatible packaging. *Sensors (Basel)* 15(8):18851–18864
- Khodagholy D, Gelines JN, Thesen T et al (2015) NeuroGrid: recording action potentials from the surface of the brain. *Nat Neurosci* 18(2):310–315
- Kozai TDY, Langhals NB, Patel PR et al (2012) Ultrasmall implantable composite microelectrodes with bioactive surfaces for chronic neural interfaces. *Nat Mater* 11(12):1065–1073
- Lopez CM, Mitra S, Putzeys J, et al (2016) 22.7 A 966-electrode neural probe with 384 configurable channels in 0.13  $\mu\text{m}$  SOI CMOS. In: *IEEE international solid-state circuits conference*. IEEE, pp 392–393
- Luan L, Wei X, Zhao Z, Siegel JJ, Potnis O et al (2017) Ultra-flexible nano-electronic probes form reliable, glial scar-free neural integration. *Sci Adv* 15:1–9
- Normann RA, Maynard EM, Rousche PJ, Warren DJ (1999) A neural interface for a cortical vision prosthesis. *Vis Res* 39:2577–2587
- Pandarinath C et al (2017) High performance communication by people with paralysis using an intracortical brain-computer interface. *elife* 6
- Park S et al (2017) One-step optogenetics with multifunctional flexible polymer fibers. *Nat Neurosci* 20(4):612–617
- Pei W, Zhang H, Wang Y et al (2017) Skin-potential variation insensitive dry electrodes for ECG recording. *IEEE Trans Biomed Eng* 64(2):463–470
- Scholvin J, Kinney JP, Bernstein JG et al (2016) Close-packed silicon microelectrodes for scalable spatially oversampled neural recording. *IEEE Trans Biomed Eng* 63(1):120–130
- Takahashi H, Suzurikawa J, Nakao M et al (2005) Easy-to-prepare assembly array of tungsten microelectrodes. *IEEE Trans Biomed Eng* 52(5):952–956
- Wise KD, Anderson DJ, Hetke JF et al (2004) Wireless implantable microsystems: high-density electronic interfaces to the nervous system. *Proc IEEE* 92(1):76–97
- Xiang Z, Yen SC, Xue N et al (2014) Ultra-thin flexible polyimide neural probe embedded in a dissolvable maltose-coated microneedle. *J Micromech Microeng* 24(6):0650152
- Yu W, WeiHua P, Kai G, Qiang G, XiaoQian L, HongDa C, JianHong Y (2011) Dry electrode for the measurement of biopotential signals. *Science China Inf Sci* 54(11):2435–2442
- Zhang H, Pei W, Chen Y et al (2016) A motion interference-insensitive flexible dry electrode. *IEEE Trans Biomed Eng* 63(6):1136–1144

# Index

## A

- AC inducted magnetic fields/forces, 629, 631
- Acoustic wave devices
  - bulk, 819–821, 823
  - FBAR (*see* Film bulk acoustic resonators (FBAR))
  - film bulk acoustic wave resonator, 819, 820
  - SAW devices (*see* Surface acoustic wave (SAW) devices)
- Activation energy, 5, 15–20, 872, 882
- Active MEMS, for optical modulation, 1369
- Actuating sensing, 632–633
- AGC+PLL control loop, 438–441
- AGC self-oscillation control loop, 435–438, 440
- Air-driven membrane-structured TENG array, 1355, 1357, 1358
- Amplitude changes, 633
- Amplitude enhancement, 758, 767
- Angular rate, open loop system, 451–452
- Anisotropic etching of silicon, *see* Silicon anisotropic etching
- Anodic bonding, 395–396, 399, 401, 404, 405, 408, 416
- Anodic stripping voltammetry (ASV), 849
- ANSYS Multi-Field Solver, 117, 134
- Automatic mode-matching control, 443–445

## B

- Beam-diaphragm coupled model, 336–337
- Bearing-structure based triboelectric nanogenerator (B-TENG), 1366
- Biochemical sensors, piezoresistive microcantilevers, *see* Piezoresistive microcantilevers
- Biological cells, 998, 999
- Bulge test, 304, 316, 320
- Bulk acoustic wave (BAW) devices

- classification, 819
  - filters, 820–821, 823
  - oscillators, 821, 823
  - resonators performance, 819, 820
- Bulk silicon microelectrode processing, 1451–1453

## C

- Calorimetry, TEC detection,
  - see* Microcalorimeters, TEC detection
- Cantilever, 198, 229, 233
  - multilayered, 224–232
  - numerical fitting, 232
  - parameters, 231
- Cantilever-based DC current MEMS sensor
  - amplitude changes, 633
  - ANSYS analysis, structural designs, 650
  - micro fabrication, 641–644
  - parallel connected PZT plates, 637–641
  - passive DC sensor, 644–647
  - piezoelectric single-side DC current sensor, 647–649
  - polarization values PZT plates, 642, 644
  - resonant frequency shift principle, 632, 649
  - sensing principle, 631–632
  - series connected PZT plates, 637–641
  - structural design, 639, 641
  - theory of cantilever bending vibration, 635–637
  - theory of Maxwell equation, 633–634
  - working principles, 631–633
- Cantilever-based micro/nano structures
  - with mode localized oscillation, 754–767
  - with synchronized oscillation, 767–783
- Capacitive humidity sensors, 801
- Capacitive micromachined ultrasonic transducers (CMUTs), 101

- Capacitive torsional accelerometers
    - bandwidth, 469–471
    - double differential (*see* Double differential capacitive torsional accelerometer)
    - dynamic equation and transfer function, 468
    - linear accelerometers, 466–467
    - mechanical sensitivity, 468–469
    - operational principle, 471–475
  - Catalytic combustion, 739, 740
  - Cellular automation method, 34–36, 91
  - Circular diaphragm deflection model, 332
  - Circular microplates, electrostatic actuation
    - capacitance calculation, 116–117
    - problem formulation, 109–110
    - pull-in voltage predication, 115–116
    - reduced-order model, 110–113
    - static deflection analysis, 114–115
    - validation of theoretical analysis, 117–128
  - Closed loop control schemes, drive mode
    - AGC+PLL control loop, 438–441
    - AGC self-oscillation control loop, 435–438, 440
    - drive velocity AGC and open loop, 441–442
  - Closed loop control schemes, sense mode
    - Allan variance, 457
    - angular rate, 451–452, 456
    - external force, 453
    - force rebalance control system, 449–451
    - resonant frequencies, 453
    - stability indices, 455
  - CMOS, *see* Complementary metal oxide semiconductor (CMOS)
  - Cobalt-based microelectrode, 851–852
  - Coefficient of thermal expansion (CTE), 240, 241
    - double-test-beam micro-rotating structure, 267–269
    - experiment, 281–292
    - H-shaped test microstructure, 270–272
    - $\Pi$ -shaped test microstructure, 272–274
    - pull-in approach, 291–300
    - simulation and verification, 275–283
    - single-test-beam micro-rotating structure, 269–270
  - Cold vapor atomic absorption spectroscopy (CVAAS), 849
  - Comb shape, 910, 927
  - Complementary metal oxide semiconductor (CMOS)
    - compatible, 591–593
    - MEMS, 798
    - sensor, 584, 588–591
  - Computational model
    - data representation, 39–41
    - implementation of cellular automation, 41–42
    - mask edge recognition, 44–45
    - normal vector, determination of, 42–44
  - Compute device architecture (CUDA), 53, 54, 56
    - SIMT, 53
    - SM, 53
    - SP, 53
  - Conducting polymers, 1418, 1427, 1428, 1430
  - Contact mode TENGs, 1349, 1351
  - Continuous cellular automaton (CCA), 29, 35, 45–52
  - Coupled array, 755–759, 761, 764, 765
  - Cover-sheet based triboelectric nanogenerator (CS-TENG), 1341, 1343
  - Current monitoring, 626
- D**
- Damping effect, 167–169
  - DC current sensor, cantilever-based, *see* Cantilever-based DC current MEMS sensor
  - DC electrical characterization
    - polysilicon resistor, 1139–1140
    - SPDT switch, 1139
  - DC inducted magnetic fields/forces, 630, 631
  - Deep OFF-state, 1066, 1072
  - Deep reactive ion etching (DRIE), 28, 425
    - CUDA, 53, 54, 56
    - deposition model, 33
    - etching model, 32–33
    - fitness function for, 38–39
    - framework of simulation system, 30–31
    - implementation of cellular automation, 41–43
    - lag effect, 33
    - parameter calibration for, 48–50
    - PSO-CCA method, 45–48
    - scalloping sidewall, 34
    - simulation of, 57–64
  - Deflection sensitivity, 693, 694, 696, 703, 704
  - Detection principle
    - slide-film combs, 431–432
    - squeeze-film combs, 433
  - Diaphragm deflection model
    - circular, 332–333
    - rectangular, 333–336
    - structured diaphragm, block deflection model, 336–337

Dicing free process, 1008

Differential pulse stripping voltammetry (DPSV), 886

Direct-current triboelectric nanogenerator (DC-TENG), 1345–1346

Direct methanol fuel cell (DMFC)

- anode current collectors, 1269, 1277
- anode flow rate effect, 1288–1291
- assembly method, 1282–1283
- assembly pressure effect, 1283–1285
- catalyst characterizations and measurements, 1291–1295
- catalyst preparation, 1274–1275
- cathode current collectors, 1269, 1270
- cell temperature effect, 1291
- current collector processing, 1277–1281
- design modeling, 1270
- encapsulation method, 1281–1282
- membrane electrode assembly, 1275–1277
- methanol concentration, 1287–1288
- testing system and methods, 1285–1287

Dirichlet boundary conditions, 170

Dispersive liquid-liquid microextraction (DLLME), 884, 885, 887

Double-clamped beams, 198, 210, 212, 216–224, 232–233

Double differential capacitive torsional accelerometer

- capacitance sensitivity, 478
- common beams, types of, 479–480
- cross-axis error, analysis of, 494
- electrode design, 490–491
- in-plane bending stiffness, 484–486
- mode analysis, 492–494
- operational principle, 475
- out-of-plane bending stiffness, 481–484
- readout circuit, 502–505
- shock resistance, analysis of, 497–498
- static mechanical sensitivity, analysis of, 491–492
- stress-released frame, design of, 488–490
- temperature robustness, analysis of, 494–497
- torsional stiffness of beams, 480–481
- unbalanced structure, design of, 486–488

Drive principle

- slide-film combs, 430–432
- squeeze-film combs, 432–433

Dynamic behavior analysis, 142–150

Dynamic performance indicators

- calibration, sensor, 381–382
- shock wave tube, 380–381

Dynamics theory, 425–430

**E**

Electrical stimulation, 1438, 1440

Electrocorticography (ECoG), 1383

Electrode impedance, 1441–1442

Electrode noise, 1442–1444

Electrode recording sites, 1442, 1443, 1459

- bulk silicon microelectrode, 1451–1453
- flexible polymer microelectrodes, 1456–1459
- metal wire implantable electrodes, 1449–1451
- silicon based MEMS microelectrodes, 1455–1456
- surface neural electrodes, 1445–1446
- thin-film silicon based, 1453–1455

Electromagnetic induction generator (EMIG), 1367

Electromagnetic interference (EMI), 1443

Electro-mechanical models

- circular microplates (*see* Circular microplates, electrostatic actuation)
- lumped electro-mechanical model, 105–108
- rectangular microplates (*see* Rectangular microplates, electrostatic actuation)

Electroplating, 953, 954, 959, 970, 978, 982, 984, 990, 991

Electrostatically actuated microplates

- circular microplates, 109–128
- dynamic behavior analysis, 142–150
- finite element modeling, 102–104
- lumped electro-mechanical model, 105–108
- rectangular microplates, 128–142

Electrostatic comb-drive actuator, MEMS relays/switches, *see* MEMS relays/switches

Electrostatic effect, 169

Encapsulation method, 1276, 1281

Environmental effect

- humidity effect, 572–573, 575
- temperature effect, 572, 574

Etching process, 642

**F**

Fabricated MEMS piezoelectric single-side DC current sensor, 647–649

Fabrication

- glass substrate and anodic bonding, 499
- silicon structure, 498–499
- trimming and packaging, 499–502

Fabrication, piezoresistive microcantilevers

- backside releasing technique, 697–698
- frontside releasing technique, 698–700

Faradaic reaction, 1382



- Fast marching method (FMM), 81–82, 89  
 and HFMM, 83–85  
 initialization, 81, 82
- FBAR, *see* Film bulk acoustic resonators (FBAR)
- FEM analysis, *see* Finite element method (FEM) analysis
- Film bulk acoustic resonators (FBAR), 1177–1179  
 basic structures, 1179–1180  
 characterizations, 1180–1183  
 fabrication process, 1180, 1184–1185  
 flexible FBAR, 1186–1190  
 integration strategy, 1184  
 S parameters of, 1184, 1187  
 system-level simulation, 1185
- Film bulk acoustic wave resonator, 819, 820
- Finite element method (FEM), 102–105, 117, 118, 122, 124, 125, 133, 134, 137, 140, 148, 151, 170–171, 721–722, 726, 731–732
- Flexibility, 1382
- Flexible MEMS microelectrode  
 fluidic channels, 1422–1424  
 future development, 1432  
 materials for, 1417–1418, 1426–1431  
 micro light-emitting diodes, 1424–1426  
 neural interface, 1418–1422  
 recording and stimulation, 1416–1417
- Flexible polymer-based microprobes, 1392–1393
- Flexible polymer microelectrodes, 1456–1459
- Flow sensor, micro thermal, *see* Micro thermal flow sensor
- Focal plate array (FPA)  
 Fresnel zone plate, 1200  
 impact resistance of, 1238  
 infrared, 1203  
 microcantilever pixel, 1210  
 monolayer, 1244  
 noises, 1213–1215  
 parameters of, 1220  
 stencil, 1205  
 structure, 1199
- Forced resonance, 1023
- Force rebalance control system, 449–451
- Fracture properties, MEMS/NEMS thin films  
 experimental apparatus, 312–313  
 fracture analysis, 309–312  
 load-deflection, 305–308  
 LPCVD silicon nitride films, 313–314  
 silicon carbide film, 314–316  
 SiO<sub>2</sub>/Si<sub>3</sub>N<sub>4</sub> composite diaphragm, 316–320  
 stress profiles, 308–309
- Freestanding triboelectric-layer mode, 1339
- Frequency and phase detection, 1079, 1109
- Frequency multiplication, 780
- Full-bridge technology, 557–558
- ## G
- GaAs substrate  
 design, 1051–1053  
 fabrication, 1051  
 measurement, 1053–1055
- Galerkin method, 110, 129
- Gas diffusion electrode (GDE) method, 1275
- Gas sensors  
 catalytic, 732–741  
 semiconductor, 741–749
- Gold nanoparticles (GNPs), 863–865
- ## H
- Hash fast marching method (HFMM), 83, 88  
 data structure of, 83, 85  
 etch rate array, reuse of, 84  
 Gaussian rate function, 87  
 initialization and propagation, 85–86  
 narrowband hash table, 85  
 precision of, 87
- Heavy metal ion microelectrodes  
 environment-friendly electrode, 863–865  
 ultramicroelectrodes, 865–868
- High driving force, 921
- High resolution, 767, 773
- High temperature piezoresistive chip  
 time drift and temperature drift, 369  
 zero output, 369
- High temperature pressure sensors (HTPS)  
 diaphragm deflection model, 332–344  
 direct measurement, 327  
 harsh environment packaging, 365–366  
 high frequency response packaging, 366–367  
 high temperature packaging, 361–363  
 indirect measurement, 327  
 lithography mask designing, 354–356  
 micromachining fabrication process, 356–360  
 performance test analysis (*see* Performance test analysis)  
 piezoresistive effect of, 328–332  
 structure designing, 349–354  
 theory verification, 344–348  
 ultra-high pressure packaging, 363–365  
 ultra-high sensitive pressure sensor chip, packaging for, 367–368

High yield, actuators, 939  
 Homogeneous liquid–liquid microextraction (HLLME), 885  
 Hot-film  
   fabrication of, 600–602  
   geometric design of, 602–604  
   micro air vehicles, 613  
   sensitivity of, 607  
   respiration monitoring, 620  
   temperature coefficient of resistance, 608  
 Humidity sensors  
   absolute humidity, 789  
   capacitive, 801–807  
   dew point/frost point, 790  
   classification of, 790–791  
   mixing ratio, 790  
   package of, 810–813  
   parts per million, 790  
   piezoresistive, 792–801  
   relative humidity, 789–790  
   resistive, 791–792  
   self-heating technology, 807–810  
 Hysteresis, 376

## I

Impedance, *see* Electrode impedance  
 Implantable neural microelectrode, 1445–1446  
 Inductively coupled plasma mass spectrometry (ICP-MS), 849  
 Inertial microswitch, 946  
   laterally-driven, 971–979  
   omni-directional sensitive, 983–988  
   physical model of, 947–950  
   triaxial, 980–983  
   vertically-driven, 954–971  
 Inertial sensor, 518  
 Inertial sensor, thermal gas  
   design and fabrication, 527–531  
   principle of, 519–527  
 Inertial sensors, behavior modeling and simulation  
   PMOR method, implicit moment matching (*see* Parametric model order reduction (PMOR))  
   TPWL method, 180–191  
 Inline capacitive MEMS power sensor  
   cantilever beam, 1107–1108  
   tethered beam, 1105–1107  
 Inline coupling MEMS power sensor  
   attenuation and phase mechanisms, 1097–1098

  lumped equivalent circuit and S-parameter, 1094–1096  
   MEMS beam, design of, 1096–1097  
   state conversion control, 1100–1103  
   transmission direction detection, 1103  
   wideband structure, 1099–1100  
 Inline inserting MEMS power sensor, 1108–1109  
 In-plane electrostatic MEMS switches  
   drive electrode unit, 913  
   optimization algorithm and segmented driving electrode, 916–920  
   optimization results analysis, 920–924  
   quasi-static process, 912–913  
 In-situ test, 199  
 Integrated microchip system, 889–891  
   buoy-based multi-parameter monitoring systems, 891–894  
 Integrated TENG array, 1359  
 Internet of Things (IoT), 818  
 Ionic liquids (ILs), 884–885  
   SI-DLLME and on-line system, 888  
 Ionization sensor, 582, 584, 590, 591  
 Iridium oxide modified recording site, 1463  
 IR image system, 1201, 1206, 1221, 1225

## L

Lamb wave resonator (LWR)  
   classification, 823–824  
   filter, 825–826  
    $Kt^2$  enhancement, 827  
   oscillator, 826  
   quality factor enhancement, 826–827  
   spurious mode suppression, 827–831  
 Large displacement, 930–932  
 Large stroke, 930  
 Lateral actuating RF MEMS switches, SOG  
   process, *see* Silicon on glass (SOG) process  
 Lateral sliding mode, 1338  
 Linear accelerometers, 466–467  
 Linear sweep voltammetry, 860, 861  
 Liquid crystal polymer (LCP) substrate  
   design, 1056–1057  
   fabrication, 1056  
   measurement, 1058–1059  
 Liquid-metal-based triboelectric nanogenerator (LM-TENG), 1344–1345  
 Liquid–solid contact triboelectric nanogenerator, 1344  
 Lithography mask designing, 354–356

- Lithography simulation, SU-8 photoresist, *see* SU-8 photoresist
- Load-deflection, 305–308, 313–317, 319
- Locking, 1001, 1014
- Lorentz force, 931
- Lower explosive limit (LEL), 732, 734, 737, 739
- Low-pressure chemical-vapor-deposition (LPCVD), 304, 313–314, 317, 319, 878
- Low pull-down voltage, 909
- Lumped electro-mechanical model, 105–108
- M**
- Magnetic actuator arrays, 1019
- Magnetic travelling wave micropump  
performance, 1031–1034  
structure and fabrication, 1026–1028
- Measurement principle, flow sensor, *see* Micro thermal flow sensor
- Mega-scale power source, 1351–1357
- Membrane electrode assembly (MEA), 1269, 1275
- MEMS, *see* Microelectromechanical system (MEMS)
- Metal-based microprobes, 1389–1392
- Metal oxide semiconductor (MOS), 741
- Metal wire implantable electrodes, 1449–1451
- Micro anemometer, 615–617
- Microcalorimeters, TEC detection  
device fabrication, 676–677  
heated reaction, 673–675  
output voltages of, 683, 684  
preconcentration, in-situ integrated carbon nanotubes for, 677–681  
qualitative model, 662–665  
static state, 668–671  
transduction principles, 665–668  
transient state, 671–673
- Microelectrode arrays (MEAs)  
biocompatibility, 1380–1381  
charge transfer capability, 1382–1383  
flexible planar, 1384  
flexible protruding, 1384  
impedance, 1381–1382  
for neural interface (*see* Neural interfaces)  
penetration implanted into brain, 1387–1393  
peripheral nervous system, 1393–1405  
placed on nerve, 1401–1406  
retinal prosthesis, 1394–1400
- Microelectromechanical system (MEMS), 584, 1018
- actuator  
applications, 943  
cycling test, 938, 942  
Lorentz force, 931  
optical photographs, 935  
scanning electron microscope, 934  
types, 936  
working principle, 931–934  
classification, 1079–1080  
conducting thin films, on-line test (*see* On-line test microstructures)  
development of, 1198–1201  
devices, 1298  
driving electrode optimization, experimental verification of, 923–927  
fabrication, 396, 400  
inertial switch  
bistable postbuckling structures, 952  
CMOS circuit, 953  
CNT-contact electrodes, 951  
contactless mechanism, 952–953  
data flip-flop chip, 988, 989  
electrostatic force, 952  
latching structure, 951  
liquid-metal droplet, 951–952  
non-silicon surface micromachining (*see* Non-silicon surface micromachining)  
PCB, 989  
physical model and working principle, 947–950  
squeeze film effect, 950–951  
triaxial sensitivity, 953  
inline capacitive, 1105–1108  
inline coupling, 1092–1104  
inline inserting, 1108–1109  
in-plane MEMS electrostatic actuator, 911–923  
microcalorimeters for TEC detection (*see* Microcalorimeters, TEC detection)  
microcantilever, 1205–1207  
microgripper (*see* Microgripper)  
microwave frequency detector, 1111–1112  
microwave phase detector, 1110–1111  
modeling and analysis of, 1201–1221  
power sensor (*see* Microelectromechanical system (MEMS) power sensor)  
process, 1303

- PVEH, 1314
- relays/switches, 909–911
- terminal indirect-heating, 1080–1091
- terminal self-heating, 1091–1092
- vibratory gyroscopes
  - AGC+PLL control loop, 438–441
  - AGC self-oscillation control loop, 435–438, 440
  - allan variance, 457
  - angular rate, open loop system, 451–452, 456
  - ANN temperature compensation system, 456–457
  - automatic mode-matching, 443–445
  - centrifugal effect, 170
  - Coriolis matrix, 171
  - damping effect, 167–169
  - drive velocity AGC and open loop, 441–442
  - dynamics theory, 430–432
  - electrostatic effect, 169
  - element mass matrix, 170
  - external force, open loop system, 452–453
  - fabrication, 423–425
  - finite element modeling, 170
  - force rebalance control system, 449–451
  - fuzzy algorithm based automatic mode-matching, 444–445
  - Laplace transform of equation, 172
  - mode matching (*see* Mode matching)
  - parameters of, 166
  - real-time mode-matching, 445–446
  - resonant frequencies, 453–456
  - schematic structure of, 165
  - simulation and experiment, 172–179
  - slide-film combs, 430–432
  - squeeze-film combs, 432–433
  - stability indices, 455
  - structure, 423–425
  - temperature effect, 170
  - trans-impedance amplifier detection, 433–435
  - zero bias varies, 457–458
  - wind sensor (*see* Micromachined wind sensor)
- Microfabrication, 660, 676, 1383
  - humidity sensors (*see* Humidity sensors)
- Microfluidics, 1019
- Microgripper, 998
  - characterization, 1011–1012
  - fabrication, 1007–1009
  - grasping structure, 1003–1004
  - pollen cells patterning, 1012–1013
  - ratchet mechanism, 1003–1005
  - rotary actuator, 1005–1006
  - silicon-on-insulator, 998
  - working principle, 999–1003
- Micro-heaters
  - FEM analysis, 721–722
  - materials, 720–721
  - structure design, 719–720
  - 3D, 727–732
  - 2D, 722–727
- Micro-hot-platform (MHP), 742, 743, 745, 746, 750
- Micromachined humidity sensors,
  - see* Humidity sensors
- Micromachined reconfigurable attenuator
  - DC electrical characterization, 1139–1140
  - fabrication, 1134–1137
  - microwave compensate structure design, 1130–1131
  - polysilicon resistive network, 1126–1127
  - reliability test, 1145
  - RF measurement, 1140–1144
  - RF MEMS switch design, 1131–1132
  - structural observation, 1138–1139
  - TaN STFR attenuation module, 1127–1130
  - 3-bit reconfigurable attenuator, simulation of, 1133–1134
  - topology design, 1124–1125
- Micromachined sensor, 518, 536
- Micromachined silicon resonant pressure sensor
  - anodic bonding, 395–396
  - BCB adhesive bonding, 394–395
  - design, 396–397, 404
  - fabrication, 398–400, 404–408
  - FEM simulation, 397–398
  - in-plane vibration mode, 403–404
  - out-of-plane vibration, 401–403
  - self-temperature compensation, 409–410
  - working principle, 389–392
- Micromachined wind sensor
  - calorimetric principle, 542–543
  - constant current mode, 544
  - constant power mode, 544
  - constant temperature difference mode, 544
  - constant voltage mode, 544
  - environmental effect (*see* Environmental effect)
  - full-bridge technology, 557–567
  - hot-wire principle, 541–542
  - low thermal conductivity, substrate with, 553–554

- Micromachined wind sensor (*cont.*)  
 1D wind sensor, 545–546  
 package (*see* Package)  
 substrate etching technology, 551–562  
 temperature balance mode, 544  
 time-of-flight principle, 543  
 2D wind sensors, 546–551, 561
- Micromachining fabrication process, 356–360,  
 579, 583, 584, 1179–1180
- Micromechanical accelerometers  
 applications, 463  
 bandwidth, test of, 508–509  
 bias stability testing, 509–510  
 challenge for, 465  
 cross-axis error, test of, 507–508  
 double differential capacitive torsional  
 accelerometer, mechanical design,  
 479–486  
 fabrication technique, 498–502  
 micromechanical capacitive torsional  
 accelerometers, fundamentals of,  
 466–479  
 noise of the accelerometer, 511  
 nonlinearity, test of, 507  
 piezoresistive accelerometers, 464  
 readout circuit, 502–513  
 resonant accelerometers, 464  
 resonant frequency and quality factor, 508  
 scale factor, test of, 506–507  
 temperature performance, test of,  
 511–513  
 thermal accelerometers, 464  
 tunneling accelerometers, 464
- Micro/nano fabrication, 745–748
- Micropumps  
 performance, 1028–1034  
 principle, 1019–1023  
 structure and fabrication, 1023–1028
- Micropumps nozzle/diffuser valveless, 1018
- Micro-scale power source, 1339–1350
- Microsensors, 661
- Micro thermal flow sensor  
 fabrication, 600–602  
 geometric design, 602–605  
 micro air vehicles, surface airflow detection  
 for, 610–614  
 principle of, 597–600  
 respiratory airflow monitor, 617–621  
 temperature compensation (*see* Temperature  
 compensation)  
 2D wind detection, micro anemometer for,  
 615–617  
 working operation, 606–607
- Microwave measurement, 1078, 1079, 1088,  
 1109, 1111
- Minimum detectable deflection (MDD), 694,  
 695, 704
- Modeling, silicon anisotropic etching,  
*see* Silicon anisotropic etching
- Modeling, SU-8 photoresist,  
*see* SU-8 photoresist
- Mode localization, 754, 755, 757, 758, 760,  
 764, 767, 784
- Mode matching  
 automatic mode-matching, 443–445  
 fuzzy algorithm based automatic mode-  
 matching, 444  
 real-time mode-matching, 445–446  
 resonant frequencies in ambient  
 temperature, 447–449  
 temperature characteristics analysis,  
 442–443
- Monolithic integration, 704–708
- Multilayered cylindrical TENG, 1346–1347
- Multilayered electret films based TENG, 1341
- Multilayer films, 214, 215, 224, 225, 232–233
- N**
- Nanomaterials, 805
- Nanotechnology, 1432
- Neural electrodes  
 challenges, 1438–1463  
 classification, 1444–1446  
 fabrication by MEMS techniques, 1446–1459  
 interfacial modification, 1459–1463
- Neural interfaces, 1418–1422  
 categories, 1438  
 description, 1438  
 electrode impedance, 1441–1442  
 electrode noise, 1442–1444  
 equivalent circuit, 1441–1442  
 (*see also* Neural electrodes)  
 neural micro electrode and signal sensing,  
 1440–1441  
 neuroelectricity, 1439, 1440
- Nitrate microelectrodes, 857  
 electrochemical measurement, 860–862
- Non-contact mode TENGs, 1349–1351
- Non-silicon surface micromachining  
 laterally-driven inertial microswitch, 971–979  
 multi-directional sensitive inertial  
 microswitch, 980–988  
 vertically-driven inertial microswitch,  
 954–966

**O**

- On-line test microstructures, 266
  - double-clamped beams, CTE, 295
  - double-clamped beams, thermal conductivity and thermal diffusivity, 260, 261
  - experiment, 253–256, 264–266
  - heat transfer, fundamentals of, 242
  - H-shape test microstructure, CTE, 270–272
  - Π-shaped test microstructure, thermal conductivity, 244–249, 272–275
  - simulation and verification, 251–253, 262–264
  - theoretical model, 260–261
  - thermal resistors, 259–260
  - transient-state thermal analysis, double-clamped beam, 256–259
  - T-shaped test microstructure, 249–251
- On-off switching, 646
- Optical-readable thermal imaging technology, 1198–1201

**P**

- Package
  - Au-Au bonding, 569–575
  - direct chip attach, 566
  - flip chip, 566–567
  - of humidity sensors, 810–813
- Parametric model order reduction (PMOR)
  - method
    - MEMS capacitive accelerometer, 157–165
    - MEMS vibratory gyroscope, 165–179
- Particle swarm optimization (PSO), 29, 36
  - principle of, 37–38
  - PSO-CCA method, 45–48
  - structure of, 38
- Passive DC sensor, 644–647
- Passive sensing, 631–632
- Performance test analysis
  - dynamic performance indicators, sensors, 379–380
  - high temperature piezoresistive chip, characteristic analysis of, 369–371
  - static performance, sensor, 372
- Peripheral nervous system (PNS), MEAs, 1393–1394
- Phase noise suppression, 768, 777, 783
- Phosphate microelectrodes, 850–851
  - molybdophosphate complex, 852–855
- Picogram-order mass sensors
  - with mode localized oscillation, 754–767
  - with synchronized oscillation, 767–783
- Piezoelectric actuator arrays, 1019, 1023
- Piezoelectric cantilevers, 631, 632, 641–644
- Piezoelectric effect, 1299–1300
- Piezoelectric equation (PE), 1300–1302
- Piezoelectric micro/nano mechanical sensors
  - biomolecule detection, 833–834
  - flexural mode vibration, 832–833
  - lab-on-a-chip system, 836–837
  - lateral extensional mode, 831–832
  - multi-mode sensing system, 834–836
  - signal enhancement strategy, 837–838
  - thickness extensional mode, 831
  - thickness shear mode, 831
  - VOC gas sensing, 838–841
- Piezoelectric travelling wave micropump
  - performance, 1029–1031
  - structure and fabrication, 1023–1026
- Piezoelectric vibration energy harvester (PVEH), 1299
  - fabrication of, 1317–1329
  - modeling of, 1303–1309
  - operation modes, 1299
  - optimization of, 1309–1317
- Piezoresistive accelerometers, 464
- Piezoresistive humidity sensors
  - basics, 793–794
  - cantilever structure, 797–801
  - membrane structure, 794–797
- Piezoresistive microcantilevers
  - biosensors, 712
  - chemical sensors, 708–711
  - design method, 695–696
  - fabrication process (*see* Fabrication, piezoresistive microcantilevers)
  - monolithic integration, 704–708
  - noise, 693–694, 702–703
  - readout methods, 692–693
  - resolution, 694–695
  - sensitivity, 693, 703–704
  - spring constant, 700–701
- Pirani sensor, 579, 581, 583, 584, 586, 587, 589, 592
- Polydimethylsiloxane (PDMS), 877, 878, 889
  - microchannel, 1024, 1028
- Polysilicon resistor, 1139–1140
- Power consumption, 541, 548, 550–573
- Power-managed triboelectric nanogenerator (PTM-TENG), 1359
- Power sensor, MEMS,
  - see* Microelectromechanical system (MEMS) power sensor

- Power-transformed-and-managed TENG (PTM-TENG), 1357
- Precision, 376–377
- Pressure-induced electrical spring softening (PIESS), 122, 123
- Pressure sensor chip
  - circular diaphragm deflection model, 332–333
  - harsh environment packaging, 365–366
  - high frequency response packaging, 366–367
  - high temperature packaging, 361–363
  - rectangular diaphragm deflection model, 333–336
  - structured diaphragm, block deflection model, 336–344
  - theory verification, 344–348
  - ultra-high pressure packaging, 363–365
  - ultra-high sensitive pressure sensor chip, packaging for, 367–368
- Printed circuit board (PCB) technology, 1348
- Proper orthogonal decomposition (POD), 157, 180
- Proportion-integration-differentiation (PID) controller, 435
- Proton exchange membrane (PEM), 1269, 1286
- Protruding, 1383
- Pt-based catalyst, 1293, 1294
- Pull-in approach
  - average temperature change, 258, 259, 297
  - double-clamped beam, 291–295
  - experiment, 297–300
  - pull-in analysis, 291–295
  - test microstructure and theoretical model, 295–296
  - thermal stress and effective Young's modulus, 297
- Pull-in voltage, 199, 205, 206, 208, 211–213, 233
- Push-pull type switch, 1064–1065, 1069, 1071, 1073
- R**
- Radial grating based high-performance TENG, 1348–1349
- Radio frequency (RF) measurement
  - SPDT switch, 1140–1141
  - 3-bit MEMS attenuator, 1142–1144
  - 20 dB polysilicon resistors, 1141–1142
- Radio frequency micro-electromechanical system (RF MEMS) switches
  - DC test, 1139
  - features of, 1122
  - modelling of, 1132
  - simulation results, 1132
  - 6-bit step attenuator, 1122
  - structure optimization, 1132
  - 3-bit reconfigurable attenuator (*see* 3-Bit reconfigurable attenuator)
- Ratchet mechanism, 998–999
  - design of, 1003–1004
  - microgripper with, 1000
  - schematic view, 1000
  - working process, 1001
- Reactive ion etch (RIE), 358, 359, 939
- Rectangular diaphragm deflection model, 333–336
- Rectangular microplates, electrostatic actuation
  - capacitance calculation, 133
  - problem formulation, 128–129
  - pull-in voltage prediction, 132–133
  - reduced-order model, 129–131
  - static deflection analysis, 131–132
  - validation of theories, 133–142
- Reliability test, 1145
- Residual stress, 198, 200, 205, 210, 213, 215, 217, 218, 222, 224, 227, 230, 232, 233
- Resistive humidity sensors, 791–792
- Resonance frequency, measurements, 214–233
- Resonant accelerometers, 464
- Resonant frequency shift principle, 632–633, 649, 650
- Resonant pressure sensor, micromachined silicon
  - electromagnetically driven and sensed, 393–400
  - electrostatically driven and capacitive sensed, 400–408
  - self-temperature compensation (*see* Self-temperature compensation)
  - working principle, 389–392
- Retinal prosthesis, MEAs, 1394–1395
- Reynolds's equation, 158
- RF MEMS switch
  - actuation mechanism, 1049–1050
  - GaAs substrate, 1051
  - LCP substrate, 1056–1059
  - series capacitive type, 1044–1047
  - series contact type, 1041–1043
  - shunt capacitive type, 1043–1044
  - shunt contact type, 1047–1049
  - SOG process, 1059–1064

- Rise-fall current switching
  - AC electrical appliances, 627
  - current monitoring, 626
  - DC electrical appliances, 627
- Rotary actuator, 999, 1005–1006
- Rotary disk based direct-current TENG, 1345–1346
  
- S**
- Salt induced dispersive liquid-liquid extraction (SI-DLLME), 888
- SAW devices, *see* Surface acoustic wave (SAW) devices
- Seebeck effect, 1079, 1081, 1083, 1092, 1093, 1108
- Self-heating technology, 807–810
- Self-powered active sensors, 1357–1365
- Self-powered vehicle collision acceleration sensor, 1362–1365
- Self-powered velocity and trajectory tracking sensor array, 1362, 1363
- Self-powered wireless sensing node, 1357–1361
- Self-temperature compensation, 409–410
  - algorithm, 410
  - calibration, 411
  - compensation results and analysis, 411–415
- Sensitivity, 541, 546, 551–565
- Sensor design, micromachined silicon resonant pressure, *see* Micromachined silicon resonant pressure sensor
- Sensors, SAW devices
  - DNA biosensor, 1176–1177
  - temperature sensor, 1171–1175
- Sensor system, 584, 586, 590, 592, 593
- Separation by implantation of oxygen (SIMOX), 327, 357, 382
- Series/parallel connections, 647, 648
- Shunt capacitive switch, 1043–1044, 1047, 1051
- Shunt contact switch, 1047–1048, 1056–1059
- Silicon anisotropic etching
  - microscopic activation energies, surface atoms, 15–20
  - microscopic etch rates, surface atoms, 5–14
  - simulations, 20–23
- Silicon based MEMS microelectrodes, 1455–1456
- Silicon-based microprobes, 1387–1389
- Silicon dry etching
  - cellular automation method, 34–36
  - computational model, 39–45
  - DRIE process simulation system, 30–34
  - parallel computation for DRIE simulation, 52–57
  - process parameter optimization, 36–39
  - PSO-CCA method, 45–52
  - simulation of DIRE applications, 57–63
- Silicon on glass (SOG) process
  - design and analysis, CPW lines, 1061–1064
  - fabrication, 1060–1061
  - mechanical characteristics, 1073–1074
  - mechanical design, 1067–1068
  - push-pull type switch, 1064–1065
  - RF characteristics, 1070–1073
  - three-state switch, rhombic beam deformation, 1065–1067
- Silicon-on-insulator, 998
- SIMOX, *see* Separation by implantation of oxygen (SIMOX)
- Simulation, silicon anisotropic etching, 4, 5, 14, 16, 20
- Single thin film resistor (STFR), 1127
  - equivalent circuit of, 1129
  - length and sheet resistance of, 1129
  - simulations results of, 1130
- Single-electrode mode, 1338–1339
- Single-pole double-throw (SPDT), 1139–1141
- Slide-film combs, 430–432
- Slide film damping, 167
- SOG process, *see* Silicon on glass (SOG) process
- Squeeze-film combs, 432–433
- Squeeze film damping, 157, 158, 161, 167
- Static performance, sensor, 372
  - full-range output and thermal sensitivity excursion, 379
  - hysteresis, 376
  - nonlinearity, 372–374
  - precision, 376–377
  - repeatability, 374–376
  - zero output, zero drift and thermal zero drift, 377–378
- Steady-state thermal analysis, test microstructures, *see* Thermal conductivity
- Stokes-type damping effect, 168
- Streaming multiprocessors (SM), 53
- Stream processors (SP), 53
- Stress concentration region (SCR), 349, 350, 353, 354
- Structural optimization, of TENG, 1355
- SU-8 photoresist
  - aerial image simulation models, 73–76
  - development simulation models, 79



- SU-8 photoresist (*cont.*)  
 exposure simulation models, 76–78  
 FMM (*see* Fast marching method (FMM))  
 PEB simulation models, 78–79  
 simulations, 90–92  
 UV lithography, 69–72
- Surface acoustic wave (SAW) devices, 818, 823, 831  
 AlN/diamond/Si multilayered structures, 1153–1154  
 DNA biosensor, 1176–1177  
 flexible pseudo-SAW devices, 1166–1171  
 LiNbO<sub>3</sub>/diamond/Si multilayered structures, 1154–1159  
 schematic structure, 1152  
 super high frequency, 1162–1166  
 temperature compensated structure Z-cut LN/SiO<sub>2</sub>/LN, 1159–1161  
 temperature sensor, 1171–1175
- Surface micromachining, non-silicon, *see* Non-silicon surface micromachining
- Surface neural electrodes, 1444–1445
- Synchronization  
 of coupled systems, 768  
 mutual, 770  
 region, 771–775  
 with and without, 775
- T**
- TaN single thin film resistor (TaN STFR)  
 attenuation module, 1127–1130
- Temperature compensation, 456–458  
 analogy compensation, basic principle for, 607–608  
 temperature sensors, 608–610
- Temperature controlled dispersive liquid phase microextraction (TC-DLPME), 886
- Temperature controlled ionic liquid dispersive liquid phase microextraction (TC-IL-DLPME), 885
- Temperature difference, 541, 542, 544, 550, 552, 553
- Terminal indirect-heating MEMS power sensor, 1080–1081  
 high dynamic range structure, 1089–1090  
 microwave performance, 1082–1083  
 optimized validation, 1084–1086  
 sandwich structure, 1090–1091  
 symmetric structure, 1086–1088  
 thermal performance, 1083–1084  
 three-channel structure, 1088–1089  
 wireless input structure, 1090
- Terminal self-heating MEMS power sensor, 1091–1092
- Test structure(s), 199, 206, 210, 214, 218, 220, 221, 224, 230, 232  
 thermophysical properties (*see* On-line test microstructures)
- Theory of cantilever bending vibration, 635–637
- Theory of Maxwell Equation, 633–635
- Thermal accelerometers, 464
- Thermal conductivity, 239, 240, 242–256, 543, 553–554  
 experiment, 253–256  
 heat transfer, fundamentals of, 242–243  
 Π-shaped test microstructure, 244–249  
 simulation and verification, 251–253  
 and thermal diffusivity (*see* Thermal diffusivity)  
 T-shaped test microstructure, 249–251
- Thermal convection, 518, 522, 523, 532
- Thermal diffusivity, 240, 255–268  
 experiment, 262, 264–268  
 simulation and verification, 262–265  
 test microstructure and theoretical model, 260–262  
 thermal resistors, 259–260  
 transient-state thermal analysis, double-clamped beam, 256–259
- Thermal expansion, 518–521, 525, 527, 531, 532
- Thin-film silicon based microelectrodes, 1453–1455
- 3-bit reconfigurable attenuator  
 fabrication process, 1135  
 simulation and measurement results, 1143  
 simulation performance, 1134  
 3D FEA model of, 1125, 1133
- Three-layer BP neural network, 445–446
- Three-state switch, 1065–1069, 1072
- Through-glass vias (TGV), 396, 401, 416
- Through-silicon vias (TSV), 416
- Torsional stiffness, 481
- Trace energetic chemicals (TEC) detection, microcalorimeters, *see* Microcalorimeters, TEC detection
- Trajectory piecewise-linear (TPWL) method, 180, 186–188  
 compact model extraction, 184–186  
 finite element discretization, 181  
 geometric and material parameters, 183–184  
 global projection matrix, 182–183

- local linearised models, 182
  - local reduced basis order and linearization point number, 188–190
  - thermoelectric governing differential equations, 180–181
  - TPWL compact model, expandability of, 190–191
  - Transient-state thermal analysis, *see* Thermal diffusivity
  - Trans-impedance amplifier detection method, 433–435
  - Tribo-controlled devices, 1365–1371
  - Triboelectric nanogenerator (TENG)
    - fundamental principle modes, 1337–1339
    - micro-scale power source, 1339–1357
    - self-powered active sensors, 1357–1365
    - tribo-controlled devices, 1365–1371
  - Triboelectrification, 1337
    - induced trace memorization device, 1372
    - self-powered memory system, 1371
    - UV light emission, 1370
  - Tunneling accelerometers, 464
  - Two-wire appliances cord
    - cross section, 645
    - fabricated device usage, 645
    - magnet positioning, 634
- U**
- Ultrahigh sensitivity, 767
  - Ultramicroelectrode arrays (UMEAs), 866, 867, 869
- V**
- Vacuum packaging, 393, 395, 396, 403, 405, 416
  - Vacuum sensor
    - CMOS Pirani, 584–590
    - CMOS platform technology for, 584
    - development of, 579
    - hot-filament, 590
    - ionization, 581–582
    - micro, 582–584
    - microfabricated, 579
    - Pirani, 580–581
    - thermal ionization, 591
  - Vertical actuating RF MEMS switches
    - GaAs substrate, shunt capacitive RF MEMS switch on, 1051–1056
    - LCP substrate, shunt contact RF MEMS switch on, 1056–1059
  - Vertical contact-separation mode, 1337
    - electret film based TENG, 1341, 1342
  - Vibration energy harvesting, 1298
- W**
- Water quality determination, 850, 855, 883
    - microsensors and systems, 862, 875, 884, 889
    - photocatalytic digestion, 873–875
    - TP digestion, 871–873
  - Wavy-structured robust triboelectric nanogenerators (WS-TENGs), 1354–1355
  - Weibull analysis, 310, 311
  - Wind sensor, micromachined thermal,
    - see* Micromachined wind sensor
  - Woven-structured triboelectric nanogenerator (W-TENG), 1339–1341
- Y**
- Young's modulus, 198, 200, 206, 207, 209, 210, 213, 215, 217, 218, 222, 226, 227, 230, 233
- Z**
- Zero bias varies, 457