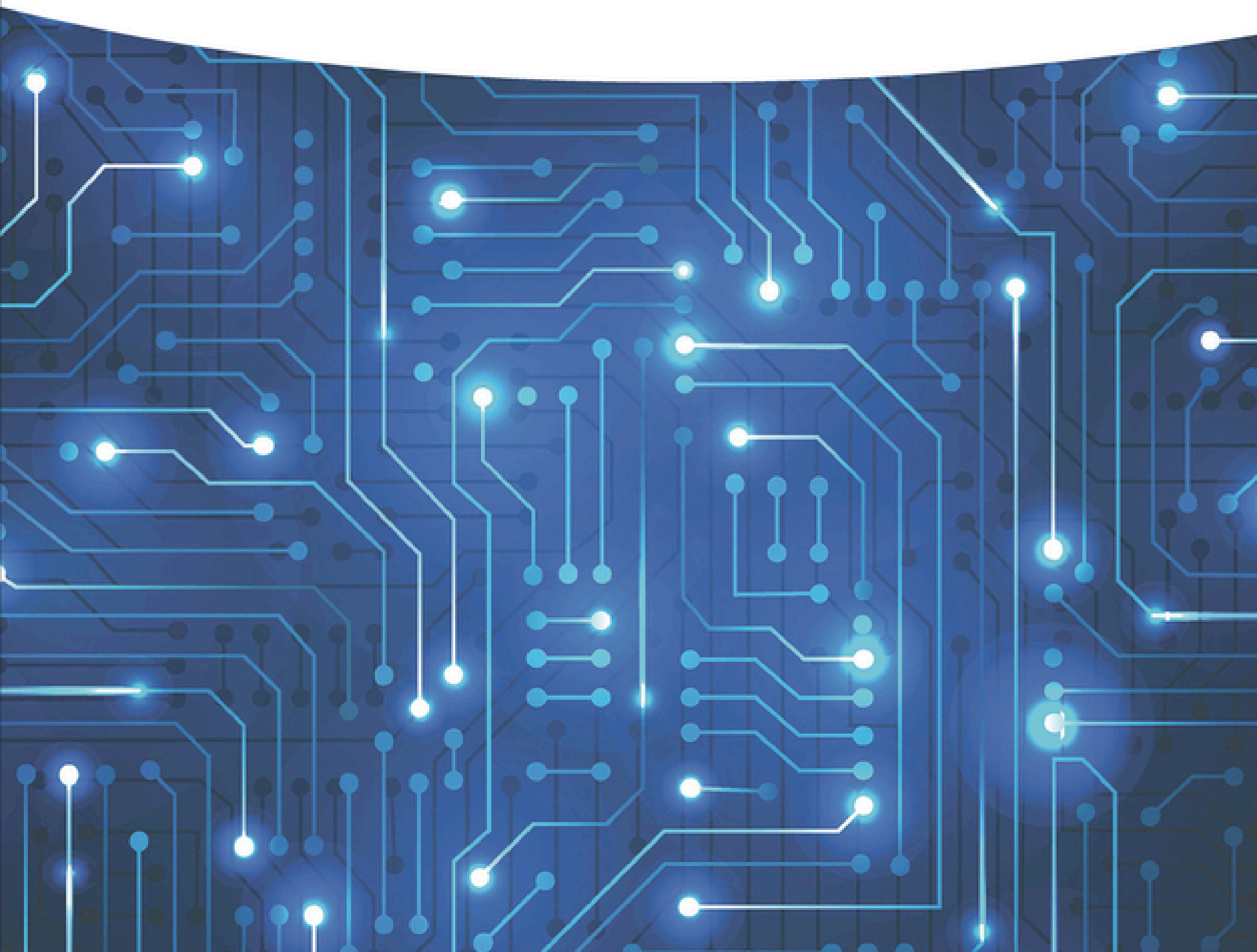
WILEY-VCH

Edited by
Muhammad Mustafa Hussain

# Advanced Nanoelectronics

## Post-Silicon Materials and Devices



**Advanced Nanoelectronics**

# Advanced Nanoelectronics

Post-Silicon Materials and Devices

*Edited by Muhammad Mustafa Hussain*

WILEY-VCH

**Editor**

***Prof. Muhammad Mustafa Hussain***
KAUST
Integrated Nanotechnology Lab.
Bldg. 3 (Ibn Sina Bldg.)
Level 3, Rm 3274
Makkah Province
23955-690 Thuwal
Saudi Arabia

**Cover image**: © iStock.com/Jackie2k

All books published by **Wiley-VCH** are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

**Library of Congress Card No.:**
applied for

**British Library Cataloguing-in-Publication Data**
A catalogue record for this book is available from the British Library.

**Bibliographic information published by the Deutsche Nationalbibliothek**
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

© 2019 Wiley-VCH Verlag GmbH & Co. KGaA, Boschstr. 12, 69469 Weinheim, Germany

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

**Print ISBN:** 978-3-527-34358-4
**ePDF ISBN:** 978-3-527-81183-0
**ePub ISBN:** 978-3-527-81185-4
**oBook ISBN:** 978-3-527-81186-1

**Cover Design** Wiley
**Typesetting** SPi Global, Chennai, India
**Printing and Binding**

Printed on acid-free paper

10 9 8 7 6 5 4 3 2 1

# Contents

**Preface** *xi*

**1 The Future of CMOS: More Moore or a New Disruptive Technology?** *1*
*Nazek El-Atab and Muhammad M. Hussain*
1.1 FinFET Technology *2*
1.1.1 State-of-the-Art FinFETs *3*
1.1.1.1 FinFET with Si Channel *3*
1.1.1.2 FinFET with High-Mobility Material Channel *4*
1.1.1.3 FinFET with TMD Channel *5*
1.1.1.4 SOI versus Bulk FinFET *5*
1.1.2 Industrial State *6*
1.1.3 Challenges and Limitations *7*
1.2 3D Integrated Circuit Technology *8*
1.2.1 Research State *9*
1.2.1.1 Thermal Management *9*
1.2.1.2 Through-silicon-vias *9*
1.2.1.3 Bonding in 3D IC *10*
1.2.1.4 Test and Yield *12*
1.2.2 Industrial State *12*
1.2.3 Challenges and Limitations *13*
1.3 Neuromorphic Computing Technology *13*
1.3.1 State-of-the-Art Nonvolatile Memory as a Synapse *14*
1.3.1.1 Phase Change Memory *15*
1.3.1.2 Conductive-Bridging RAM *16*
1.3.1.3 Filamentary RRAM *17*
1.3.2 Research Programs and Industrial State of Neuromorphic Computing *18*
1.4 Quantum Computing Technology *19*
1.4.1 Quantum Bit Requirement *20*
1.4.2 Research State *20*
1.4.2.1 Spin-Based Qubits *20*
1.4.3 Superconducting Circuits for Quantum Information *21*
1.4.4 Industry State *22*
1.4.5 Challenges and Limitations to Quantum Computing *23*
References *23*

**2 Nanowire Field-Effect Transistors** *33*
*Debarghya Sarkar, Ivan S. Esqueda, and Rehan Kapadia*
2.1 General Scaling Laws Leading to Nanowire Architectures *33*
2.1.1 Scaling of Planar Devices and Off-state Leakage Current *33*
2.1.2 FinFET and UTB Devices for Improved Electrostatics *35*
2.1.3 Nanowires as the Ultimate Limit of Electrostatic Control *37*
2.1.4 Quantum Effects *39*
2.1.5 Drive Current *43*
2.2 Nanowire Growth and Device Fabrication Approaches *43*
2.2.1 Bottom-up VLS Growth *43*
2.2.2 Top-down Oxidation *45*
2.3 State-of-the-Art Nanowire Devices *45*
2.3.1 Silicon Devices *45*
2.3.2 III–V Devices *46*
References *49*

**3 Two-dimensional Materials for Electronic Applications** *55*
*Haimeng Zhang and Han Wang*
3.1 2D Materials Transistor and Device Technology *56*
3.1.1 Operation and Characteristics of 2D-Materials-Based FETs *57*
3.1.2 Ambipolar Property of Graphene *57*
3.1.3 Important Figures of Merit *58*
3.1.3.1 $I_{on}/I_{off}$ Ratio *58*
3.1.3.2 Subthreshold Swing *59*
3.1.3.3 Cutoff Frequency and Maximum Frequency of Oscillation *59*
3.1.3.4 Minimum Noise Figure *60*
3.1.4 Device Optimization *61*
3.1.4.1 Mobility Engineering *61*
3.1.4.2 Current Saturation *62*
3.1.4.3 Metal Contact *63*
3.2 Graphene Electronics for Radiofrequency Applications *64*
3.2.1 Experimental Graphene RF Transistors *65*
3.2.2 Graphene-Based Integrated Circuits *67*
3.2.2.1 Graphene Ambipolar Devices *67*
3.2.2.2 Graphene Oscillators *73*
3.2.2.3 Graphene RF Receivers *73*
3.2.2.4 Graphene Electromechanical Devices: Resonators and RF Switches *74*
3.3 $MoS_2$ Devices for Digital Application *76*
3.3.1 Experimental $MoS_2$ Transistors *77*
3.3.2 $MoS_2$-Based Integrated Circuits *78*
3.3.2.1 Direct-Coupled FET Logic Circuits *78*
3.3.2.2 Logic Gates *79*
3.3.2.3 A Static Random Access Memory Cell based on $MoS_2$ *82*
3.3.2.4 Ring Oscillators based on $MoS_2$ *82*
3.3.2.5 Microprocessors based on $MoS_2$ *85*
References *87*

**4 Integration of Germanium into Modern CMOS: Challenges and Breakthroughs** *91*
*Wonil Chung, Heng Wu, and Peide D. Ye*
4.1 Introduction *91*
4.2 Junction Formation for Germanium MOS Devices *92*
4.2.1 Charge Neutrality Level and Fermi Level Pinning *92*
4.2.2 Metal/Ge Contacts *93*
4.2.2.1 Alleviation of FLP *93*
4.2.2.2 Metal/n-Ge Contact *93*
4.2.2.3 Recessed Contact Formation *94*
4.3 Process Integration for Ge MOS Devices *97*
4.3.1 Interface Engineering Issues *97*
4.3.2 Various Gate Stack Combinations for Ge MOSFET *97*
4.3.2.1 $GeO_x$-Free Gate Stack with ALD High-κ *98*
4.3.2.2 Silicon Interfacial Layer Passivation *99*
4.3.2.3 Germanium (Oxy)Nitridation *99*
4.3.2.4 $GeO_2$-Based Gate Stacks *99*
4.3.2.5 Rare-Earth Oxides Integrated into Germanium MOSFETs *100*
4.3.3 Stress and Relaxation of Ge Layer on an Si-Based Substrate *102*
4.4 State-of-the-Art Ge CMOS with Recessed Channel and S/D *102*
4.4.1 Germanium CMOS Devices *102*
4.4.2 Germanium CMOS Circuits *105*
4.5 Steep-Slope Device: NCFET *107*
4.6 Conclusion *111*
References *112*

**5 Carbon Nanotube Logic Technology** *119*
*Jianshi Tang and Shu-Jen Han*
5.1 Introduction – Silicon CMOS Scaling and the Challenges *119*
5.2 Fundamentals of Carbon Nanotube *122*
5.3 Complementary Logic and Device Scalability Demonstrations *124*
5.3.1 CNT NFET and Contact Engineering for CMOS Logic *124*
5.3.2 Channel Length Scaling in CNTFET *127*
5.3.3 Contact Length Scaling in CNTFET *132*
5.4 Perspective of CNT-Based Logic Technology *138*
5.4.1 CVD-Grown CNT versus Solution-Processed CNT *138*
5.4.2 Purity and Placement of Solution-Processed CNTs *140*
5.4.3 Variability in CNTFETs *140*
5.4.4 Circuit-Level Integration *142*
5.5 Summary and Outlook *142*
References *143*

**6 Tunnel Field-Effect Transistors** *151*
*Deblina Sarkar*
6.1 Introduction *151*
6.2 Tunnel Field-Effect Transistors: The Fundamentals *153*
6.2.1 Working Principle *153*

6.2.2 Single-Carrier Tunneling Barrier and Subthreshold Swing *154*
6.3 Modeling of TFETs *156*
6.4 Design and Fabrication of TFETs *161*
6.4.1 Design Considerations *161*
6.4.2 Current Status of Fabricated TFETs *162*
6.5 Beyond Low-Power Computation *166*
6.5.1 Ultrasensitive Biosensor Based on TFET *169*
6.5.2 Improvement in Biosensor Response Time *173*
References *175*

**7 Energy-Efficient Computing with Negative Capacitance** *179*
*Asif I. Khan*
7.1 Introduction *179*
7.2 How a Negative Capacitance Gate Oxide Leads to Sub-60-Millivolt/Decade Switching *181*
7.3 How a Ferroelectric Material Acts as a Negative Capacitor *182*
7.4 Direct Measurement of Negative Capacitance in Ferroelectric *186*
7.5 Properties of Negative Capacitance FETs: Modeling and Simulation *188*
7.6 Experimental Demonstration of Negative Capacitance FETs *190*
7.7 Speed of Negative Capacitance Transistors *195*
7.8 Conclusions *195*
References *196*

**8 Spin-Based Devices for Logic, Memory, and Non-Boolean Architectures** *201*
*Supriyo Bandyopadhyay*
8.1 Introduction *201*
8.2 Spin-Based Devices *203*
8.2.1 Spin Field-Effect Transistor (SPINFET) *205*
8.2.2 Single Spin Logic Devices and Circuits *209*
8.3 Nanomagnetic Devices: A Nanomagnet as a Giant Classical Spin *212*
8.3.1 Reading Magnetization States (or Stored Bit Information) in Nanomagnets *215*
8.3.2 Writing Magnetization States (or Storing Bit Information) in Nanomagnets *216*
8.3.2.1 Spin-Transfer Torque *216*
8.3.2.2 Spin-Transfer Torque Aided by Giant Spin Hall Effect *217*
8.3.2.3 Voltage-Controlled Magnetic Anisotropy *220*
8.3.2.4 Straintronics *224*
8.4 Conclusion *231*
Acknowledgments *232*
References *232*

**9 Terahertz Properties and Applications of GaN** *237*
*Berardi Sensale-Rodriguez*
9.1 Introduction *237*
9.1.1 Applications of Terahertz Technology *237*
9.1.2 Terahertz Devices and Challenges *239*
9.2 GaN: Properties and Transport Mechanisms Relevant to THz Applications *242*
9.2.1 Mobility and Injection Velocity *242*
9.2.2 Drift Velocity and Negative Differential Resistance *245*
9.2.3 Transport due to Electron Plasma Waves *246*
9.3 GaN-based Terahertz Devices: State of the Art *249*
9.3.1 High-Electron Mobility Transistors *249*
9.3.2 NDR and Resonant Tunneling Devices *252*
9.3.3 Quantum Cascade Lasers *254*
9.3.4 Electron-Plasma-Wave-based Devices *254*
References *256*

**Index** *265*

## Preface

We live in the age of information where electronics, especially transistors, play a critical enabling role. Nearly 60 years ago, when Jack Kilby built the first integrated circuit (IC) that was also the beginning of today's complementary metal-oxide-semiconductor (CMOS) technology whose arts and science of miniaturization has enabled Moore's Law to physically scale transistors for the past 40 years. The dominant material vehicle for transistor technology has been silicon. However, as we are approaching the scaling limit in silicon CMOS transistor technology, we need to find new ways to keep the momentum in CMOS electronics technology advances. Therefore, in this edited book, I have closely worked with leading authorities to assemble a comprehensive collection of chapters dedicated to emerging technologies, which will take the CMOS electronics technology forward. The objective of such technology is simple: faster, multifunctional, energy-efficient computing and infotainment applications.

This book has been organized in such a manner that it covers materials, physics, architecture, and integration aspects of future generation CMOS electronics technology. We have discussed about emerging materials and architectures such as alternate channel materials like germanium, gallium nitride, 1D nanowires/tubes, 2D graphene and other dichalcogenide materials, ferroelectrics; new physics such as spintronics, negative capacitance, and quantum computing; and, finally, 3D-IC technology.

In 2009 fall, when I transitioned from industry to academia, I realized that there is a certain gap in our knowledge and understanding, which needs a bridge for emerging concepts and technological advances in the area of CMOS electronics technology. As a follow-up, I introduced a course on Advancement in CMOS Electronics Technology focusing on graduate and senior-level undergraduate students. I invited leading authorities to offer lectures through online and in person. Interestingly, we observed an increased number of class presence; especially, postdoctoral fellows and research scientists also started participating in the class. We kept enriching the course every year as new concepts and technological advancements were taking place in parallel and at a fast pace. In 2012 fall, United States National Academy of Engineering Member Prof. James Plummer (former Dean of Stanford Engineering) visited us and I discussed with him about the content of the course. He was vastly impressed and encouraged me to think about a book. Therefore, this current book was implanted as a seed in my mind in 2012 and today it has materialized to provide a comprehensive understanding about

the future generation CMOS electronics technology to especially senior-level undergraduate students, graduate students, and professionals interested in the relevant technology.

Although the book in its current shape is rich in content, it was not an easy task to meet the timelines and to work with such a busy cohort of leading authorities as contributors of individual chapters. I definitely acknowledge their contributions along with the students and postdocs who have worked together to make it a comprehensive insightful resource on emerging CMOS electronics technology. I thank the proofreading service and especially my postdoctoral fellow Dr. Nazek Elatab. Her contribution as integrator has been instrumental for timely completion of this project. I also thank all my mentors and colleagues because it has been a privilege for me to serve with them in this exciting area of electronics with boundless promise. Understandably, my family has contributed by sacrificing long hours that I spent on this book instead of with them! Finally, I thank Wiley-VCH for encouraging me to lead this project and to finally realize a dream I have had for many years.

I sincerely hope that the readership will find this book a resourceful knowledge base for future generation CMOS electronics technology, which we are calling Advanced Nanoelectronics: Post-silicon Materials and Devices.

# 1

# The Future of CMOS: More Moore or a New Disruptive Technology?

*Nazek El-Atab and Muhammad M. Hussain*

*King Abdullah University of Science and Technology, Integrated Nanotechnology Lab, Thuwal, 4700, Saudi Arabia*

For more than four decades, Moore's law has been driving the semiconductor industry where the number of transistors per chip roughly doubles every 18–24 months at a constant cost. Transistors have been relentlessly evolving from the first Ge transistor invented at Bell Labs in 1947 to planar Si metal-oxide semiconductor field-effect transistor (MOSFET), then to strained SiGe source/drain (S/D) in the 90- and 65-nm technology nodes and high-κ/metal gate stack introduced at the 45- and 32-nm nodes, then to the current 3D transistors (Fin field-effect transistors (FinFETs)) introduced at the 22-nm node in 2011 (Figure 1.1). In extremely scaled transistors, the parasitic and contact resistances greatly deteriorate the drive current and degrade the circuit speed. Thus, miniaturization of devices so far has been possible due to changes in dielectric, S/D, and contacts materials/processes, and innovations in lithography processes, in addition to changes in the device architecture [1, 2].

The gate length of current transistors has been scaled down to 14 nm and below, with over $10^9$ transistors in state-of-the-art microprocessors. Yet, the clock speed is limited to 3–4 GHz due to thermal constraints, and further scaling down the device dimensions is becoming extremely difficult due to lithography challenges. In addition, further scaling down the complementary metal-oxide semiconductor (CMOS) technology is leading to larger interconnect delay and higher power density [3]. The complexity of physical design is also increasing with higher density of devices. So, what is next?

A promising More-than-Moore technology is the 3D integrated circuits (ICs) which can improve the performance and reduce the intra-core wire length, and thereby enable high transfer bandwidth with reduced latencies and power consumption, while maintaining compact packing densities [4]. Alternative technologies that could be promising for new hardware accelerators include resistive computing, neuromorphic computing, and quantum computing.

Resistive computing could lead to non–von Neumann (VN) computing and enforce reconfigurable and data-centric paradigms due to its massive parallelism and low power consumption [5]. Moreover, humans can easily outperform current high-performance computers in tasks like auditory and pattern recognition

*Advanced Nanoelectronics: Post-Silicon Materials and Devices,*
First Edition. Edited by Muhammad Mustafa Hussain.
© 2019 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2019 by Wiley-VCH Verlag GmbH & Co. KGaA.

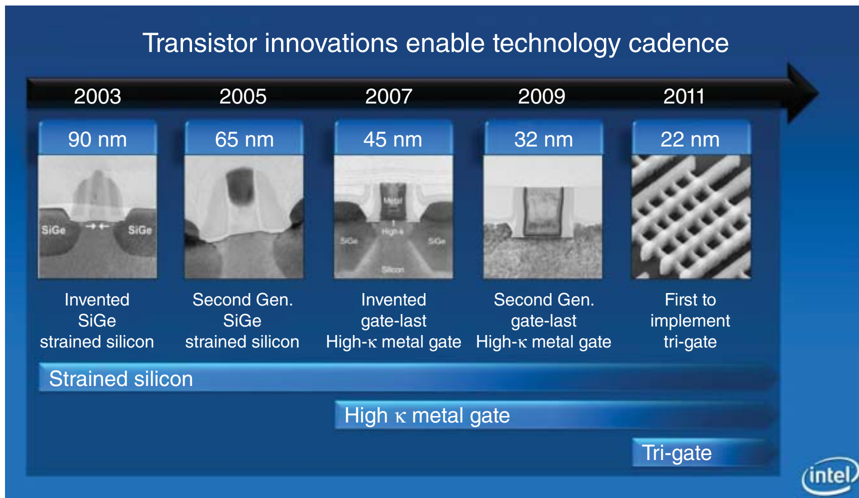



**Figure 1.1** Intel innovation in process technology for the past decade. Source: www.intel.in.

and sensory motor control. Thus, neuromorphic computing can be promising for emulating such tasks due to its energy and space efficiency in artificial neural network applications [6]. Quantum computing can solve tasks that are impossible by classical computers, with potential applications in encryptions and cryptography, quantum search, and a number of specific computing applications [7].

In this chapter, four main technologies are discussed: FinFET, 3D IC, neuromorphic computing, and quantum computing. The state-of-the-art findings and current industrial state in these fields are presented; in addition, the challenges and limitations facing these technologies are discussed.

## 1.1 FinFET Technology

Over the past four decades, the continuous scaling of planar MOSFETs has provided an improved performance and higher transistor density. However, further scaling down planar transistors in the nanometer regime is very difficult to achieve due to the severe increase in the leakage current $I_{off}$. In fact, as the channel length in planar MOSFETs is reduced, the drain potential starts to affect the electrostatics in the channel and, consequently, the gate starts to lose control over the channel, which leads to increased leakage current between the drain and source. A higher gate-channel capacitance can relieve this problem using thinner and high-κ gate oxides; however, the thickness of the gate oxide is fundamentally restricted by the increased gate leakage and the gate-induced-drain leakage effect [8–10].

An alternative to planar MOSFETs is the multiple-gate FETs (MuGFETs) which demonstrate better electrostatics and better screening of the drain from the gate due to the additional gates covering the channel [11–14]. As a result, MuGFETs show better performance in terms of subthreshold slope, threshold voltage ($V_t$) roll-off, and drain-induced barrier lowering (DIBL). Another alternative to planar

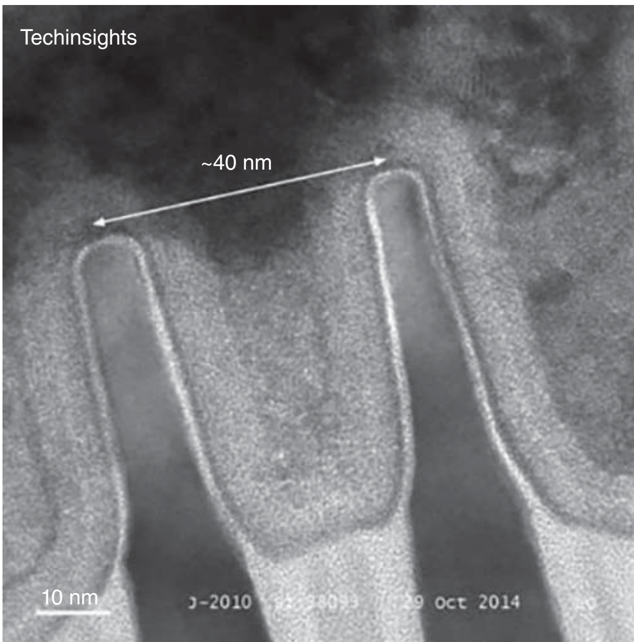



**Figure 1.2** TEM image of Intel's 14-nm transistors with sub-40-nm fin pitch. Source: www.techinsights.com.

bulk MOSFETs is fully depleted silicon on insulator (FDSOI) MOSFETs, which reduce leakage between drain and source due to the removal of the substrate right below the channel [15]. The performance of the FDSOI MOSFETs is comparable with the double-gate field-effect transistors (DGFETs) in terms of SS, low junction capacitance, and high $I_{on}/I_{off}$ ratio. Yet, the DGFETs have better scalability and can be manufactured on bulk Si wafers instead of silicon-on-insulator (SOI) wafers, which makes them more promising [16].

FinFETs or tri-gate FETs, which have three gates, have been found to be the most promising alternatives to MOSFETs due to their enhanced performance and simplicity of the fabrication process, which is compatible with and can be easily integrated into standard CMOS fabrication process (Figure 1.2) [17, 18]. In fact, an additional selective etch step is required in the FinFET fabrication process in order to create the third gate on top of the channel. FinFET devices have been explored carefully in the past decade. A large number of research articles that confirmed the improved short-channel behavior using different materials and processes have been published, as is shown in the following section. Next, the industrial state of FinFETs, their challenges, and limitations are discussed.

### 1.1.1 State-of-the-Art FinFETs

#### 1.1.1.1 FinFET with Si Channel

In the semiconductor industry, silicon is the main channel material. The first FinFET technology (22-nm node) was produced by Intel in 2011. The second FinFET generation (14-nm node) published by Intel used strained Si channel [19].

The gate length was scaled from 26 to 20 nm in the second FinFET generation, which was possible due to new sub-fin doping and fin profile optimization. With a $V_{DD}$ of 0.7 V, the saturation drive current is 1.04 mA μm$^{-1}$ and the off current is 10 nA μm$^{-1}$ for both nMOSFET (NMOS) and pMOSFET (PMOS). The SS is ∼65 mV/decade, while the DIBL for N/PMOS is ∼60/75 mV V$^{-1}$. High-density static random access memory (SRAM) having 0.0588 μm$^2$ cell size are also reported and fabricated using the 14-nm node. More recently, a research group from Samsung published a 7-nm CMOS FinFET using extreme ultraviolet (EUV) lithography instead of multiple-patterning lithography. This resulted in a reduction of the needed mask steps by more than 25%, in addition to providing smaller critical dimension variability and higher fidelity. The FinFET presented in this work consumes 45% less power and provides 20% faster speed than in the previous 10-nm technology. The reported SS is 65 and 70 mV/decade, and the DIBL is 30 and 45 mV V$^{-1}$ for NMOS and PMOS, respectively. A 6T high-density and high-current SRAM memory has also been demonstrated using the 7-nm FinFET, and the results show a reduction in the bit line capacitance by 20% as a result of the reduction in the parasitic capacitance.

#### 1.1.1.2 FinFET with High-Mobility Material Channel

The III–V materials gained growing attention for adoption as the channel material due to their promising characteristics such as high mobility, small effective mass, and, therefore, high injection velocity, in addition to near-ballistic performance. The first InGaSb pFET was demonstrated by Lu et al. [20], where a fin-dry etch technique was developed to obtain 15-nm narrow fins with vertical sidewalls. An equivalent oxide thickness (EOT) of 1.8 nm of $Al_2O_3$ was used as the gate oxide. The authors also demonstrated Si-compatible ohmic contacts that yielded an ultralow contact resistivity of $3.5\times10^{-8}$ Ω cm$^2$. Devices with $L_g = 100$ nm and different fin widths ($W_f$) were demonstrated. The results show that with $W_f = 100$ nm, $g_m$ of 122 μS μm$^{-1}$ is achieved; while with $W_f = 30$ nm, $g_m$ of 78 μS μm$^{-1}$ is obtained.

Moreover, FinFETs with strained SiGe have lately attracted much interest due to their potential advantages such as higher mobility, built-in strain, and improved reliability with respect to conventional Si-based FETs. Very recently, a group of researchers at IBM demonstrated high-Ge-content strained SiGe FinFETs with replacement high-κ (HK)/metal gate (RMG). A long-channel subthreshold swing (SS) as low as ∼68 mV/decade was reported [21]. This value is very competitive with other SiGe or Ge FinFETs with RMG process flow, where the reported SS values are in the range of 80–100 mV/decade [22]. In addition, a very high pFET hole mobility $\mu_{eff} = 235$ cm$^2$ V$^{-1}$s$^{-1}$ was shown in a multi-fin device with average fin width of 4.6 nm and EOT of 7 Å which could be very promising for the sub-5-nm node FinFETs. Finally, in the same work, SiGe FinFETs with gate lengths $L_g = 25$ nm were fabricated using a gate-first flow. At a $V_{DD} = 0.5$ V, the devices showed DIBL = 40 mV, $SS_{lin}/SS_{sat} = 77/86$ mV/decade and $I_{on} = 430$ μA μm$^{-1}$ at target high performance $I_{off} = 100$ nA μm$^{-1}$, which are among the largest reported values at such gate lengths (Figure 1.3).

In another work by Lei et al. [23], conducted in collaboration with Taiwan Semiconductor Manufacturing Co. (TSMC), the first GeSn FinFET device on

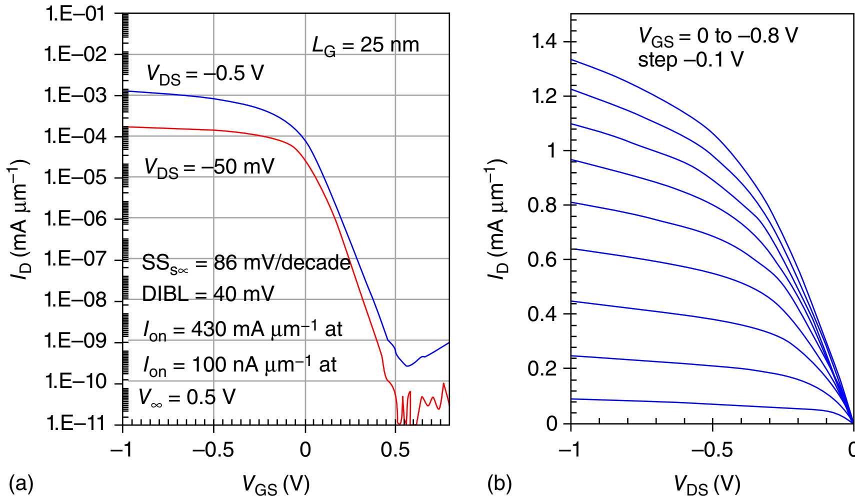



**Figure 1.3** Transfer and output characteristics of high-Ge-content SiGe FinFETs with $L_G$ 25 nm with gate first flow. Source: Hashemi et al. 2017 [21]. Reused with permision of IEEE.

a GeSnOI substrate was demonstrated with a channel length of 50 nm and $W_{Fin}$ = 20 nm, and 4 nm $HfO_2$ was used as the gate oxide. The novel substrate was fabricated by the growth of high-quality GeSn by chemical vapor deposition (CVD) followed by a low-temperature process flow to get the GeSnOI. The GeSn pFET yielded the lowest SS of 79 mV/decade, the highest transconductance $g_m$ of 807 µS µm$^{-1}$, and the highest hole mobility of 208 cm$^2$ V$^{-1}$ s$^{-1}$ ($N_{inv}$ of $8 \times 10^{12}$ cm$^{-2}$).

#### 1.1.1.3 FinFET with TMD Channel

For sub-5-nm nodes, a body with sub-3-nm thickness is required to maintain good channel control. Most channel materials like the conventional Si or III–V face limitations in terms of mobility, quantum capacitance, or process at such ultrathin body (UTB) thickness. Advanced two-dimensional transition-metal dichalcogenide (TMD) is very promising in UTB thickness due to its sub-nanometer monolayer UTB thickness potential in addition to its good transport characteristics in nanometer thickness [24]. Chen et al. demonstrated the first 4-nm-thick TMD body FinFET with back gate control [25]. The main processes in the fabrication of the TMD FinFET is the compatibility of the CVD growth of TMD with CMOS processing, in addition to the reduction of the contact resistance by hydrogen plasma treatment of $MoS_2$. The $V_t$ of this FinFET device can be adjusted dynamically by applying bias on the back gate. The front gate device showed an on/off current ratio over $10^5$ with $I_{on}$ of 200 µA µm$^{-1}$ for $V_{dd}$ = 1 V.

#### 1.1.1.4 SOI versus Bulk FinFET

Bulk FinFETs are built on bulk-Si wafers, which are less expensive and have a lower defect density than do SOI wafers, while maintaining a better heat transfer rate to the Si substrate with respect to SOI FinFETs. The first Intel FinFET was a

bulk FinFET. Lee [1] studied the 14-nm node FinFET technology and compared bulk and SOI FinFET in terms of scalability, heat dissipation, and parasitic capacitance. Lee showed that both 14-nm FinFETs with bulk and SOI substrates have the same $I-V$ characteristics when the same geometry and doping concentration are used. Therefore, both devices have similar scalability. Moreover, the fins in bulk FinFETs are easily depleted, which allows for the reduction of the S/D to fin body junction capacitance to values that are lower than in the case of SOI FinFETs. Finally, to increase the heat transfer rate in SOI FinFETs, the buried oxide should be made thinner than 20 nm, which could have a negative impact on the device performance such as an increase in the parasitic capacitance.

Finally, it is worth mentioning that there are many other factors that affect the performance and reliability of FinFET devices, such as the materials used for metal gate/gate oxide, the shape of the fins (trapezoidal versus rectangular), the spacing between the fins, the fin edge roughness, choice of FET structure (lateral, vertical), and so on, which are not discussed in this chapter.

### 1.1.2 Industrial State

In 2011, Intel was the first company to use the 22-nm bulk FinFETs in mass production of central processing units (CPUs), which is 18% and 37% faster at 1 and 0.7 V, respectively, than Intel's 32-nm transistors [26]. Intel reported at the International Electron Devices Meeting (IEDM) that these 3D tri-gate transistors have a saturation current that exceeds 2 mA $\mu m^{-1}$. Several companies then followed Intel and announced the production of 3D transistors such as Samsung, TSMC, and Global Foundries. In 2015, Samsung announced the first production of the 14-nm FinFET-based transistors for mobile applications followed by the first mass production of the 10-nm FinFET (10LPE) in October 2016. Samsung was able to show improvements in power (40% lower power consumption than their 14-nm FinFETs), performance (27% higher performance), and scalability of the 3D tri-gate transistors (30% higher area efficiency).

However, at the 10-nm node, only three companies were capable of manufacturing such transistors: Samsung, Intel, and TSMC (Global Foundries excluded). Moreover, the geometries of the transistors produced at the leading manufacturers are different. For instance, the 10-nm FinFETs produced at TSMC and Samsung are denser than Intel's 14-nm FinFETs; however, they are closer to Intel's 14-nm FinFETs than they are to the Intel's 10-nm (the metal pitch in the Samsung's 10-nm is just 1 nm shorter than Intel's 14-nm).

In addition, some foundries use a hybrid node while others execute full node shrinking, which results in different geometries. In hybrid node shrinking, a new structure for the transistor (or a smaller transistor) is used (front end of line (FEOL)) but employing a set of design rules established previously for connecting transistors together (back end of line (BEOL)). In full node shrinking, both FEOL and BEOL are shrinking. In fact, TSMC and Samsung used the hybrid nodes at 16/14 nm where they introduced the new FinFET structure, while Intel is the only company executing full node shrinking with every new technology. It is worth noting that hybridized nodes allow the foundries to tackle a single set of challenges since the whole design process is not fully scaled down at once.

During Intel's Technology and Manufacturing Day 2017, Intel announced the mass production of its 10-nm process which used self-aligned quad patterning (SAQP) for the first time. Intel's 10-nm technology showed 45% less power consumption and 25% better performance than their 14-nm transistors with a minimum gate pitch of 54 nm (versus 70 nm for Intel's 14 nm) and a metal pitch of 36 nm (versus 52 nm for Intel's 14 nm). Also, Intel's 10-nm density is 2.7 × higher than the previous node (new density of 100.8 mega transistors $mm^{-2}$), with 25% taller (53-nm fin height) and more closely spaced fins (34-nm fin pitch).

Saumsung's 10 nm uses triple-patterning technology with a 68-nm contacted gate pitch, 51-nm metal pitch, dual-depth shallow trench isolation (STI) with a single dummy gate (ref Common Platform Alliance Paper which was presented in 2016), while TSMC's 10-nm used quad-patterning technology which allows a double increase in density compared to their 16-nm technology. TSMC claimed a poly pitch of 64 nm and a metal pitch of 42 nm with 35% less power consumption and 15% higher performance than their 16-nm technology.

In June 2017, Global Foundries announced the mass production of its 7-nm FinFET technology which offers 40% improvement in performance with volume production ramping in the second half of 2018. The initial production ramp of the 7-nm technology employs triple and quadruple patterning technology using a 193-nm excimer laser. Global Foundries will introduce EUV to its manufacturing process to accelerate the production ramp and improve the yield.

TSMC announced recently that its 7-nm FinFET will offer around 25% speed enhancement or a 35% power reduction over its 10-nm FinFETs, while Samsung announced the addition of the 8- and 6-nm process technologies to its current process roadmap with an aim of improving the cost competitiveness over its 10- and 7-nm technologies. It is also worth noting that Samsung's 7-nm will be its first technology to use EUV lithography.

### 1.1.3 Challenges and Limitations

The introduction of the FinFET technology has enabled the gate length scaling down to 7 nm with a 48-nm contacted poly pitch (CPP) due to improved device electrostatics [27]. The improved performance has been achieved through the "Fin Effect" boost (effective fin width/fin pitch) which increased the drive current for a certain capacitive load. However, the restrictions on the fin thickness are being rapidly approached, which would lead to a faster scaling in S/D sizes versus the contacted gate pitch. An increasing "Fin Effect" will thus result, which in combination with a plateau in the gate length would put pressure on the conduction path from contacts to S/D. In a work conducted by a group of researchers from Global Foundries and IBM, current contact resistivity of $\sim 2 \times 10^{-9}\ \Omega\,cm^2$ [28] will significantly deteriorate the performance of FinFETs below 40-nm CPP, while fully ohmic contacts with resistivity of $\sim 1 \times 10^{-10}\Omega\,cm^2$ [29] might push the CPP to below 30 nm. The work concluded that in order to further improve the performance and power consumption in future CMOS in the 30–40 nm CPP, industry will face pressure to use new device architectures or scaling choices [2].

Another challenge is that sub-5-nm nodes would need sub-3-nm body thickness for maintaining good channel control [25]. However, most of the channel

materials such as Si, Ge, and other III–V materials face fabrication, mobility, and quantum capacitance challenges at such small body thicknesses [30]. In addition, a group of researchers from IBM have fabricated test structures to unambiguously observe quantum confinement effects. The structures included fins with 40-nm fin pitch, 20-nm $L_g$, and 4- to 30-nm $W_{Fin}$. The measurements showed performance/mobility degradation, increase in series resistance, increase in variability, DIBL, and in $V_t$ of NMOS/PMOS as the $W_{Fin}$ is reduced [31], which confirms the challenges to be faced when further scaling down the FinFET technology.

Wavy FinFET has been proposed by Fahad et al. [32] as a promising structure for the high-performance technology node. The wavy transistor integrates 2D UTBs with the fin structure which maximizes the chip area utilization resulting in higher density, higher gain, and back bias capability. The structure was simulated using the 2013 International Technology Roadmap for Semiconductors (ITRS) specifications for the 7-nm node with UTB thickness of 2.5 nm and fin thickness of 6.8 nm. The authors reported an improved SS and DIBL performance of the wavy channel with 109% higher non-normalized ON-state drive performance as opposed to conventional FinFETs.

## 1.2 3D Integrated Circuit Technology

3D integration technology can denote either 3D packaging or 3D IC, which can be defined in different ways. In general, in 3D packaging, the vertical stacks are achieved via traditional methods of interconnects such as wire bonding and flip chip [33, 34]. However, in 3D IC, interconnections between different stacking layers are formed via through-silicon-vias (TSVs) [35]. Die stacking can be achieved by connecting separately manufactured dies or wafers vertically through one of three integration schemes: die-to-die, die-to-wafer, and wafer-to-wafer. The contacts (mechanical and electrical) can be achieved using either microbumps or by wire bonding as used in system-in-package (SIP) and package-on-package (POP) devices. Even though SIP is sometimes referred to as a 3D stacking technology, it is better referred to as a 2.5D technology. Another approach is to integrate dies horizontally on a silicon substrate using interposers. The benefits of using interposers are several: (i) lower communication power consumption due to the short communication distance between dies, (ii) the possibility of stacking separately manufactured dies from heterogeneous technologies to get the best out of all technologies, and (iii) enhanced yield and cost of the system due to the ability of fabricating and testing the smaller dies separately before integrating them into a silicon substrate instead of fabricating very large dies with much lower yield. The most promising approach of 3D integration is the monolithic approach, where active layers are vertically grown on top of each other and interconnects are made through TSVs which provide the densest connectivity.

There are several topics related to 3D integration that have recently gained a lot of attention in research. In the following, the main research topics with corresponding state-of-the-art technology are presented, followed by the industrial state and the main challenges of this technology.

### 1.2.1 Research State

#### 1.2.1.1 Thermal Management

The biggest obstacle to the commercialization of 3D IC is the thermal management problem. As a matter of fact, the very thin thickness of chips in the 3D IC (<50 μm) in addition to the very high density of devices results in an increase in the temperature of the dies which are not close to the heat sink, and thereby deteriorating the performance of the system. In the past few years, research addressing thermal problems in 3D IC has gained growing attention. Goplen et al. reported that TSVs can act as a vertical path for heat flow [36]; therefore, thermal TSVs in addition to signal TSVs can be used to vertically transfer the heat and thereby reduce the die temperatures [37, 38]. Another study done by Lee et al. [39] showed that the heat transfer is directly proportional to the size of the via islands. In addition, it was found that a large number of TSVs can lead to routing congestion in the 3D ICs; thus, in addition to being expensive to fabricate, an optimization algorithm is needed to find the needed number of TSVs and their locations in order to be able to reduce the temperature of the dies. Moreover, Furumi et al. [40] proposed new cooling architectures for 3D ICs based on thermal sidewalls, interchip plates, and a bottom plate (thermal SIB). The experimental results conducted using a 3D thermal solver show that the thermal SIB can reduce the temperature in a 3D IC by over 40% when compared with structures that used a conventional heat sink only.

#### 1.2.1.2 Through-silicon-vias

Using TSVs in 3D ICs and 3D packaging is very promising since it allows higher integration density, higher clock rate, and lower power dissipation [41]. In addition, TSVs are used in the 2.5D through-silicon interposers which enable the integration of heterogeneous dies on a silicon substrate. However, the fabrication of TSVs can be challenging: the etch process of the high-aspect-ratio TSVs should lead to scallop-free Si [26] and the Cu-filled TSV should be void-free [42]. This is in addition to challenges related to Cu protrusion affecting the BEOL reliability [43], thinning of TSV wafer [44], revealing of the backside of the TSV, and the bonding process [45]. In general, TSV fabrication requires the following steps: patterning of the via, etching the via, depositing the dielectric liner, metallization, and, finally, chemical-mechanical planarization (CMP) for planarization [46].

Currently, scaling down the TSVs is driven by the need to lower the thermal-mechanical stress in addition to its effect on the BEOL performance. The depth of the TSVs is limited, constrained by the wafer thinning (usually fixed at 50 μm). For a higher aspect ratio TSV (beyond 10 : 1), using the physical vapor deposition (PVD) barrier and seed process might lead to non-conformal films. IMEC and Lam Research Corp developed a low-cost process for getting conformal deposition of a very thin barrier and seed layer in high-aspect-ratio TSVs. The process consists of depositing a highly conformal thin oxide liner using atomic layer deposition (ALD), followed by the ALD deposition of the WN barrier, electroless plating NiB seed, and, finally, filling the TSV with copper using electrochemical deposition (ECD) [47]. Tokyo Electron Limited also reported another method to deposit highly conformal barrier and seed layers using electroless plating of Cu on

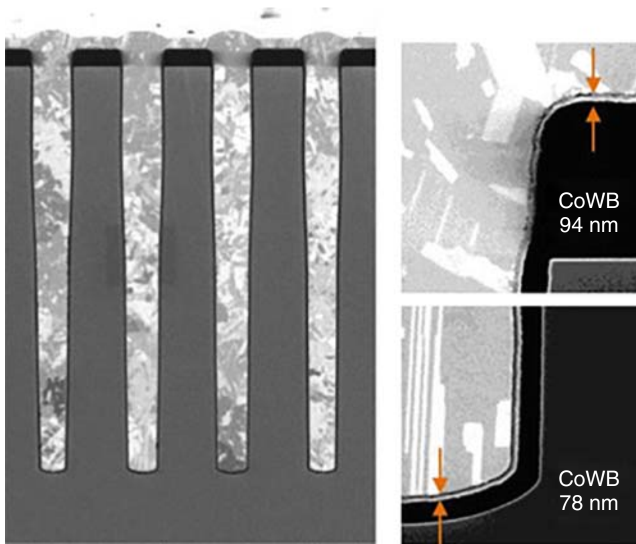



**Figure 1.4** FIB-SEM after ECD-Cu filling on Eless-Cu/CoWB layers of 5 × 50 μm TSV. Source: Tanaka et al. 2015 [48]. Reused with permission of IEEE.

CoWB followed by Cu filling the TSV using ECD. Figure 1.4 shows the 5 × 50 μm TSV reported by Tokyo Electron Limited [48].

Another innovative metallization process was developed by Aveni (previously known as Alchimer). This metallization method is based on molecular engineering, where the film is grown molecule by molecule and can be applied in industry. First, a barrier layer is deposited by grafting and the NiB compound is used to make a Cu diffusion barrier which maintains the resistivity levels such that Cu can fill the high-aspect-ratio TSV using electrografting without the need for a cupper seed layer. The final fill process results in large, uniform, and high-purity grains of Cu, which could increase the yield due to eliminated voids, shorts, and opens [49].

#### 1.2.1.3 Bonding in 3D IC

As already mentioned, the most important aspect of 3D IC is the ability to integrate heterogeneous dies fabricated at different foundries without performance degradation. The integration can be achieved either through wafer-on-wafer (WoW) bonding, chip-on-wafer (CoW) bonding, or chip-on-chip (CoC) bonding (Figure 1.5) [50]. WoW is the most preferred bonding due to its precise alignment [51]; more specifically, Cu metal-to-metal thermocompression bonding is the most favored among all bonding methods as it provides excellent electrical conductivity and mechanical strength after bonding [52]. During the thermocompression of Cu—Cu bonding, interdiffusion of Cu atom and grain growth across the bonding interface takes place. However, the main challenge to this process is to achieve it at low pressure and low temperature in order to avoid damaging the devices underneath or cause any reliability issues. But the Cu—Cu bonding process requires high temperature and pressure (or either of them) as native

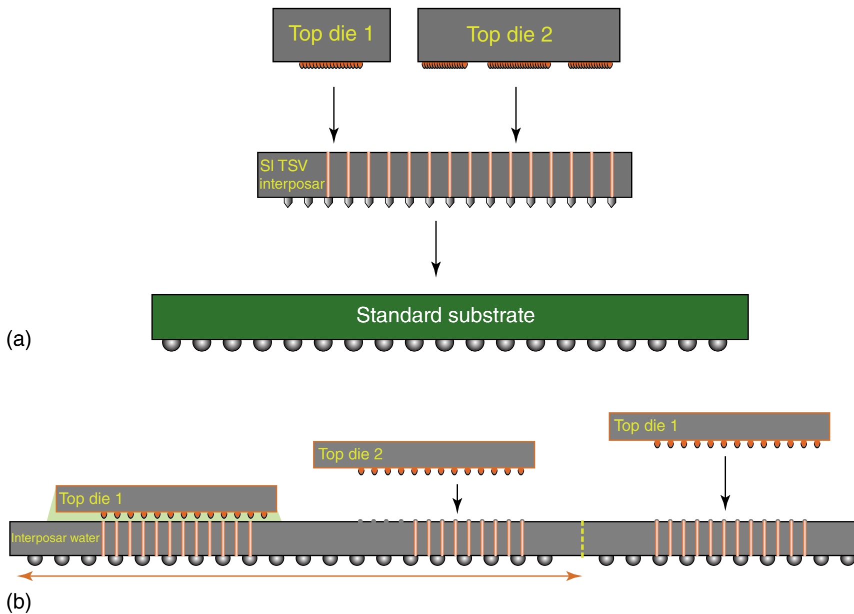



**Figure 1.5** Assembly die stacking process flow: (a) CoS and (b) CoW. Source: https://amkor.com/.

oxide can be easily grown on the Cu surface, which inhibits the Cu interdiffusion and degrades the bonding quality [53]. Therefore, achieving high-quality Cu—Cu bonding at low temperature and pressure is needed in 3D IC.

Researchers have worked on several ways to avoid the surface oxidation of Cu and to remove the already grown native oxide. Shigetou et al. reported a bonding method based on surface-activated bonding (SAB). First, the native oxide is removed using argon bombardment at ultra-high vacuum (UHV), and then the bonding takes place at room temperature using a SAB flip-chip bonder [54]. However, the need for UHV increases the complexity of the process, and as a result becomes unattractive for manufacturing. Also, different chemistries have been proposed to do wet etching/cleaning of the native oxide such as hydrochloric acid [55], citric acid [56], sulfuric acid [57], and acetic acid [58]. Although some wet etching chemistries succeeded in removing the native oxide, immersing the wafer in such chemistries for a prolonged time might lead to etching the Cu and deteriorating the performance of the devices underneath. In another work conducted by Tan et al., a self-assembled monolayer (SAM) of alkane thiol, an organic monolayer, is deposited on the Cu surface to passivate it; and then the SAM is desorbed before the Cu—Cu bonding [59]. The use of the SAM passivation layer protected the Cu from growing native oxide, and the SAM removal process can be done at 250 °C [60]. However, all passivation based on SAM are not CMOS compatible.

In a work addressing this problem, a 3-nm Ti layer was used instead to passivate the Cu surface at 160 °C and 2.5 bar [61]. However, the Ti materials are challenging to be used in the damascene process; in addition, Ti can oxidize if exposed to air for more than two days, which is not favorable for the 3D IC process. In

a continuation to this work, it has been reported that a 3-nm of manganin alloy passivation layer deposited at 150 °C and low pressure led to a strong Cu—Cu bonding of 5 kN force, in addition to being damascene compatible.

#### 1.2.1.4 Test and Yield

Every additional manufacturing step introduces a risk for defects and complicates the testing of the system. Yield is based on test results, and the cost is based on the test, yield, and throughput. 3D high yield is challenging to achieve, which is why the wire bonding of "known good dies" in 3D packages first found its application in mobile devices.

Any 3D IC process would be considered feasible only if its manufacturability yield is high. A group from Xilinx Inc. reported the key challenges faced during fabricating a 28-nm 3D IC with chip-on-wafer-on-substrate process [62]. During the initial ramp stage, most of the observed failures were related to the assembly at the interposer level such as open microbumps, opens and shorts in the interposer metal line, and TSV opens. Another failure mode is the deterioration of the transistors during the assembly of the 3D IC. The group developed a failure analysis technique based on a closed loop feedback, which resulted in improved yields.

### 1.2.2 Industrial State

Samsung is already using the monolithic approach to die stacking in 3D flash memory and smart sensors. The first commercial prototype of 3D IC (microcontroller) dates back to 2004 when Tezzaron released it [63]. In 2006, Intel assessed 3D chip stacking in Pentium 4 [64]. In 2011, IBM announced the introduction of the 3D chip production process [65]. Also, in 2012, Tezzaron released a prototype for its multicore design, which includes 64 core 3D-MAPS (*MA*ssively *P*arallel processor with *S*tacked memory) (http://arch.ece.gatech.edu/research/3dmaps/3dmaps.html)[66]. In 2013, a 128-Gb 3D NAND chip was introduced by Samsung which has 2× transistor density, 50% lower power consumption, 2× data storage speed, and 10× better retention characteristics compared to the planar version.

In 2015, Intel also introduced the 3D XPoint memory with 10× higher capacity than DRAM and 1000× faster than NAND flash [67]. Moreover, NVIDIA and AMD manufactured a high bandwidth memory (HBM) using 3D stacked memories, which is already used in the AMD GPU based on the Fiji architecture since 2015 (https://images.nvidia.com/content/pdf/tesla/whitepaper/pascal-architecture-whitepaper.pdf). A high-performance RAM competing with HBM is the hybrid memory cube (HMC), which was introduced in 2011 by Micron and is based on DRAM stacked using TSVs (https://www.micron.com/products/hybrid-memory-cube). SanDisk and Toshiba announced in 2015 the production of the world's first 3D NAND with 48 layers and using BiCS (Bit-Cost Scalable) technology. The 3D NAND achieved 32 GB capacity with a storage of 3 bits per transistor. The latest version is called BiCS3, which will have 64 layers and will show a 40% larger capacity than the BiCS2, according to Toshiba.

Moreover, Micron reported the mass production of its 64-layer 3D NAND by the end of 2017, while Western Digital began mass production of its 64-layer

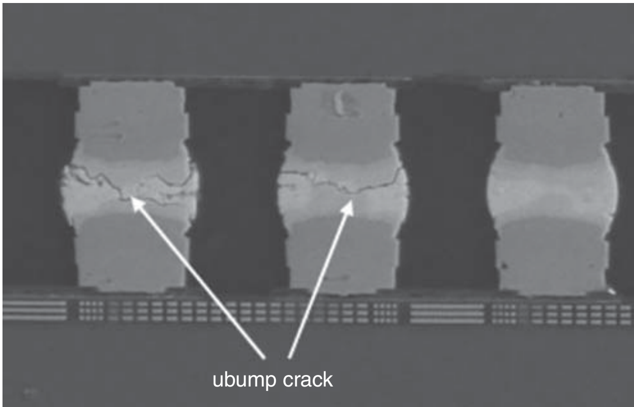



**Figure 1.6** Picture of a microbump crack. Source: Yip et al. 2017 [68]. Reused with permission of IEEE.

3D NAND flash chips in 2017 (https://www.anandtech.com/show/10274/the-crucial-mx300-750gb-ssd-review-microns-3d-nand-arrives). Also, in early 2017, Intel announced the world's first commercial solid-state drive (SSD) based on 64-layer 3D NAND with a capacity of 512 GB (https://www.anandtech.com/show/11571/the-intel-ssd-545s-512gb-review-64layer-3d-tlc-nand-hits-retail).

### 1.2.3 Challenges and Limitations

Several challenges face the commercialization of 3D IC or 3D packaging. In a work done by Intel, it was found that yield estimates modeled using traditional methods can be pessimistic by as much as 50%. New analytical models have to be established to take into consideration other effects such as defect clustering and systematic defects introduced by equipment and handling issues during manufacturing. Moreover, it has been reported that the electrical performance of 3D IC with TSVs is affected due to the structure of the TSV with microbumps. TSV and microbump structures lead to local mechanical stress and strain due to the mismatch in the coefficient of thermal expansion (CTE) of Si, Cu TSV, and microbump (Figure 1.6) [68]. Moreover, crystal defects and stress can be induced in the Si chip as it is thinned down to less than a couple of tens of micrometers. Also, the gettering layers used to avoid the contamination of the metal and crystal defects can be removed from the Si chip as it is thinned down.

## 1.3 Neuromorphic Computing Technology

The flexibility of the VN architecture for "stored program" has led to enormous improvements in system performance for more than five years. However, since miniaturizing devices have slowed down in the past years, the energy and time used to transport data between memory and processor has become difficult, especially for data-centric applications such as real-time pattern recognition where state-of-the-art VN systems try hard to meet the performance of a human

being. The human brain outperforms advanced processors on many tasks such as unstructured data classification due to its parallel architecture connecting low-power neurons and synapses which act as computing and adaptive memory elements, respectively. The human brain performance is actually inspiring for novel non–VN computing models needed in future computing systems.

Even though designing neural circuits using electronic components dates back to the implementation of retinas [69] and perceptrons [70], modern research about very-large-scale integration (VLSI) technology using the nonlinear current characteristics began in the mid-1980s through collaboration between Richard Feynman, Carver Mead, Max Delbrück, and John Hopfield [71]. In fact, Mead tried to imitate the gradual synaptic transmission in the retina using the analog properties of transistors rather than operating them as digital switches. Mead was able to demonstrate that neuromorphic circuits using analog transistors instead of digital ones can match the physical properties of the proteic channels in neurons [72], leading to the need for a much smaller number of transistors to emulate neural systems.

In the neural system, neurons are connected to many other neurons, and they pass electrical and chemical signals to each other via synapses. These connections are either strengthened or weakened through a process called spike-timing-dependent plasticity (STDP), which is biologically observed [73, 74]. STDP changes depending on the timing between spikes (action potentials) within the input neuron (presynaptic) and output neuron (postsynaptic). In long-term potentiation (LTP), causal spiking strengthens synapses; while in long-term depression (LTD), the synaptic strength is weakened by causal spiking [75]. The change in the weight of synapses, also called synaptic plasticity, explains how the brain learns and memorizes [76].

Neuromorphic computing technology is considered a promising candidate for implementing applications such as self-learning, recognition of patterns, gestures, and speech using energy-efficient/low-power spiking networks. However, the progress in this technology faces two main challenges: (i) the lack of a full understanding of how the brain works and (ii) the lack of agreement on which technology can achieve synaptic and neural circuits with the best balance between cost, performance, and power consumption. Currently, a great deal of research is being conducted on different technologies for neuromorphic computing including mathematical and machine learning algorithms, neuromorphic datasets, field programmable gate array (FPGA) codes, photonic neuromorphic signal processing, nonvolatile memory (NVM) solutions, and so on. In this chapter, NVM for neuromorphic computing is discussed. In addition, the current industry state of neuromorphic computing, its challenges, and limitations are discussed.

### 1.3.1 State-of-the-Art Nonvolatile Memory as a Synapse

Around $10^{11}$ neurons and $10^{14}$ synapses exist in the human brain. In order to be able to implement brain-like processing architectures without using large and expensive areas on the silicon wafer, highly scalable and low-power memory devices are needed.

Different NVM devices have different physical properties and switching behaviors, and thus can be used to emulate synapses in different ways. For instance, when synapses are connected or not, an on/off NVM response would be sufficient; and this can be achieved using conductive-bridging random access memory (CBRAM). In other cases, synaptic weights are needed; therefore, an NVM with adjustable conductance would be required and this can be achieved using phase change memory (PCM) or memristor/resistive-random access memory (RRAM). In the following, the different types of NVM used to emulate synapses are briefly explained with state-of-the-art examples from the literature.

#### 1.3.1.1 Phase Change Memory

In PCM, the state of the memory, whether programmed/SET or erased/RESET, depends on the difference in electrical resistivity between the amorphous and crystalline phases of the "phase change materials" leading to low (RESET) and high conductance (SET), respectively [77, 78] (Figure 1.7a).

PCM is attractive for neuromorphic applications where "device history" is needed, since the SET state can be achieved gradually by applying repetitive pulses to crystallize the phase of the plug in the device, resulting in a high-resistance state [84]. However, the RESET process can be only done sharply, since it involves melt and quench. The STDP can be implemented using a two-PCM approach: when an input neuron spikes, it outputs a signal (read pulse) and enters the LTP mode for a period of time $t_{\mathrm{LTP}}$. If the postsynaptic neuron spikes during this period, a SET pulse is then sent to the LTP synapse. If not, then the LTD synapse is programmed, as shown in Figure 1.8a,b.

Suri et al. demonstrated that by adding a thin $HfO_2$ layer to the $Ge_2Sb_2T_5$ (GST)–based PCM, their synaptic performance can be improved [85, 86]. The addition of the interface layer affects the nucleation and growth activation energies, and thereby the crystallization kinetics, resulting in an increased dynamic range. In a later work, the authors developed a circuit model including

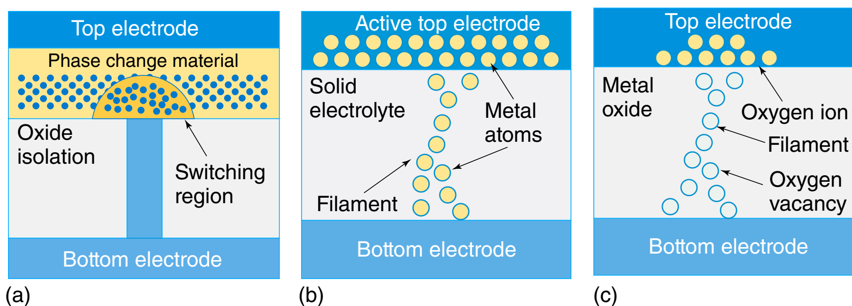



**Figure 1.7** (a) Phase change memory (PCM) depends on the large difference in electrical resistivity between the amorphous (low-conductance) and crystalline (high-conductance) phases of so-called phase change materials [79, 80]. (b) Conductive-bridging RAM is based on the electrochemical formation of conductive metallic filaments through an insulating solid electrolyte or oxide [81]. (c) The conductive filaments in a filamentary RRAM are chains of defects through an otherwise insulating thin-film oxide [82]. Source: Reused with permission from [83].