SCIENCE CHINA
Information Sciences

October 2021, Vol. 64 201401:1–201401:33
https://doi.org/10.1007/s11432-021-3235-7

• REVIEW •

Recent progress of integrated circuits and optoelectronic chips

Yue HAO[1*], Shuiying XIANG[1,2*], Genquan HAN[1], Jincheng ZHANG[1], Xiaohua MA[1], Zhangming ZHU[1], Xingxing GUO[2], Yahui ZHANG[2], Yanan HAN[2], Ziwei SONG[2], Yan LIU[1], Ling YANG[1], Hong ZHOU[1], Jiangyi SHI[1], Wei ZHANG[3], Min XU[4], Weisheng ZHAO[5], Biao PAN[5], Yangqi HUANG[6], Qi LIU[7], Yimao CAI[8], Jian ZHU[9], Xin OU[10], Tiangui YOU[10], Huaqiang WU[11], Bin GAO[11], Zhiyong ZHANG[12], Guoping GUO[13,14,15], Yonghua CHEN[16], Yong LIU[17], Xiangfei CHEN[18], Chunlai XUE[19,20], Xingjun WANG[21], Lixia ZHAO[22], Xihua ZOU[23], Lianshan YAN[23] & Ming LI[19,20,24]

[1]*School of Microelectronics, Xidian University, Xi'an* 710071*, China;*
[2]*State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an* 710071*, China;*
[3]*National Integrated Circuit Innovation Center, Shanghai* 201203*, China;*
[4]*School of Microelectronics, Fudan University, Shanghai* 200433*, China;*
[5]*School of Integrated Circuit Science and Engineering, Beihang University, Beijing* 100191*, China;*
[6]*China Center for Information Industry Development, Beijing* 100017*, China;*
[7]*Frontier Institute of Chip and System, Fudan University, Shanghai* 200433*, China;*
[8]*Institute of Microelectronics, Peking University, Beijing* 100871*, China;*
[9]*Nanjing Electronic Devices Institute, Nanjing* 210016*, China;*
[10]*State Key Laboratory of Functional Material for Informatics, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai* 200050*, China;*
[11]*Institute of Microelectronics, Tsinghua University, Beijing* 100084*, China;*
[12]*Center for Carbon-based Electronics, Department of Electronics, Peking University, Beijing* 100871*, China;*
[13]*CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei* 230026*, China;*
[14]*CAS Center for Excellence and Synergetic Innovation Center in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei* 230026*, China;*
[15]*Origin Quantum Computing Company Limited, Hefei* 230026*, China;*
[16]*Key Laboratory of Flexible Electronics & Institute of Advanced Materials, Nanjing Tech University, Nanjing* 211816*, China;*
[17]*State Key Laboratory of Electronic Thin Films and Integrated Devices, School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu* 610054*, China;*
[18]*National Laboratory of Solid State Microstructures & College of Engineering and Applied Sciences, Nanjing University, Nanjing* 210093*, China;*
[19]*State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing* 100083*, China;*
[20]*Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing* 100049*, China;*
[21]*State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Electronics Engineering and Computer Science, Peking University, Beijing* 100871*, China;*
[22]*School of Electrical and Electronic Engineering, Tiangong University, Tianjin* 300387*, China;*
[23]*Center for Information Photonics and Communications, School of Information Science and Technology, Southwest Jiaotong University, Chengdu* 611756*, China;*
[24]*School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing* 100049*, China*

Received 9 February 2021/Revised 26 March 2021/Accepted 30 March 2021/Published online 27 May 2021

*Corresponding author (email: yhao@xidian.edu.cn, syxiang@xidian.edu.cn)

© Science China Press and Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract Integrated circuits (ICs) and optoelectronic chips are the foundation stones of the modern information society. The IC industry has been driven by the so-called "Moore's law" in the past 60 years, and now has entered the post Moore's law era. In this paper, we review the recent progress of ICs and optoelectronic chips. The research status, technical challenges and development trend of devices, chips and integrated technologies of typical IC and optoelectronic chips are focused on. The main contents include the development law of IC and optoelectronic chip technology, the IC design and processing technology, emerging memory and chip architecture, brain-like chip structure and its mechanism, heterogeneous integration, quantum chip technology, silicon photonics chip technology, integrated microwave photonic chip, and optoelectronic hybrid integrated chip.

Keywords integrated circuit, semiconductor science and technology, optoelectronic device and chip, photonic integrated circuit, wide bandgap semiconductors, silicon photonics, hybrid integration, quantum chip, integrated microwave photonic, photonic neural computing

Citation Hao Y, Xiang S Y, Han G Q, et al. Recent progress of integrated circuits and optoelectronic chips. Sci China Inf Sci, 2021, 64(10): 201401, https://doi.org/10.1007/s11432-021-3235-7

1 Introduction

One of the significant technological achievements of the last 60 years is the integrated circuit (IC). Nowadays, the IC industry is closely related to everyone's daily life, as shown in Figure 1. With the rapid development of the computer, Internet, mobile communication, big data, cloud computing, artificial intelligence (AI), 5G communication, optical communication, Internet of things (IoT), automotive electronics, satellite, quantum information, and other emerging technology applications, the demand for IC has been explosively increased in recent years. In particular, due to the COVID-2019, there are new consumption scenarios such as intelligent conferences, remote intelligent medical systems, and online education. These applications call for high-performance computing, low transmission delay, and low power consumption.

Semiconductor science and technology are the foundation stones of the IC industry [1]. The IC products, such as the information transmission and processing devices and chips, are made of semiconductor materials. The first-generation semiconductor materials are represented by Si and Ge. The silicon devices and ICs account for the majority of all semiconductor products, which lead to the rapid development of microcomputers. The second-generation semiconductor materials are represented by GaAs and InP. With these compound semiconductor materials, the semiconductor lasers with wavelengths at 850, 1310, and 1550 nm that match well with the low loss windows of optical fiber, are the foundation of the optical communication network. Thus, optical communication chips have become an indispensable part of the communication industry. Besides, the wide bandgap (WBG) third-generation semiconductor materials represented by SiC and GaN have broad application prospects in the fields of light display, light storage, and light illumination [2, 3]. The GaN is especially suitable for high frequency, high efficiency, high temperature, high voltage, and wideband high power microwave devices, which are promising candidates for radar detection, satellite communication, and 5G communication. Furthermore, the ultra-wide bandgap semiconductor materials represented by Ga_2O_3 and AlN have also attracted extensive attention in recent years [4, 5]. Various semiconductor devices and ICs promote the rapid development of modern information society. Every electronic product we touch may contain at least one semiconductor chip. An example is a mobile phone, which contains Si and GaAs components.

Since complementary metal-oxide semiconductor (CMOS) technology was born in the 1950s, the IC industry has been driven by the so-called "Moore's law". In 1965, Gordon Moore predicted that the number of components per chip would continue to increase by a factor of two every year, and conversely, the size of each transistor would decrease continuously [6]. In 1975, Moore updated his earlier prediction by forecasting that components per chip would double every two years [7]. In 1965, the minimum feature size of the IC chip was approximately 50 mm. The device dimension of the CMOS transistor has been scaled continuously [8]. The progression of semiconductor scaling is marked by technology nodes [9]. Traditional metal-oxide-semiconductor field-effect transistors (MOSFETs) scaling worked well up until the 130-nm generation in the early 2000s. After that, significant innovations in transistor materials and structure were introduced to continue scaling. An example was the introduction of strained silicon transistors on Intel's 90-nm technology in 2003. As another example, high-K metal gate transistors were introduced on Intel's 45-nm technology in 2007. The next major innovation was the introduction

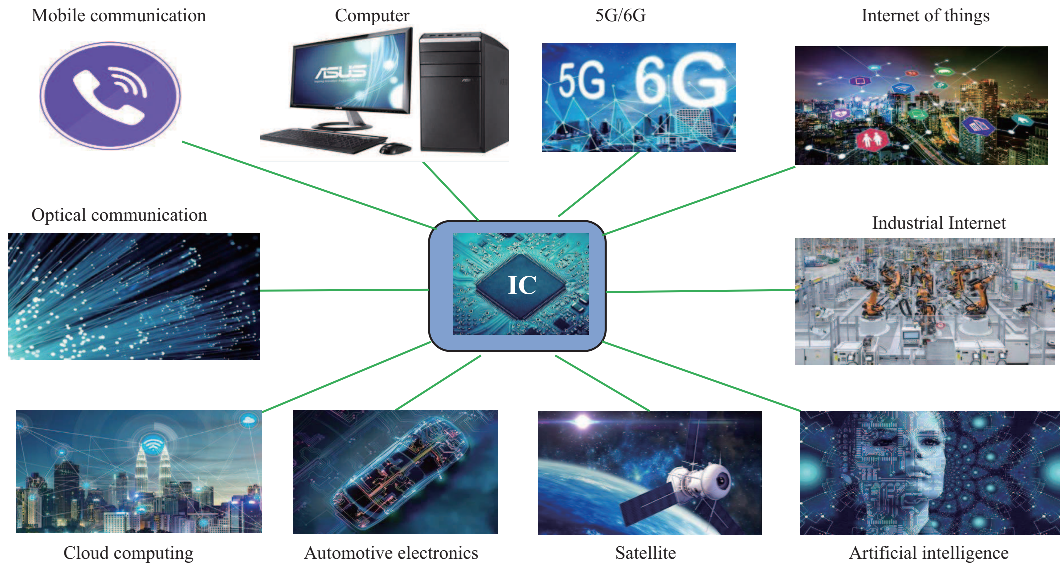

Figure 1 (Color online) Representative applications of ICs.

of FinFET transistors on Intel's 22-nm technology in 2011. Since then, FinFET has involved in five generations of evolvement as predicted by Moore's law and typically progresses as 22, 14, 10, 7, 5, 3 nm. These innovations have continuously improved the IC performance.

Unfortunately, the device downscaling trend is now much slowing down owing to the power limitations and quantum effects. The feature size has approached the technological and physical limits. Note that, as the feature size of the transistor becomes smaller and meanwhile as billions of transistors are integrated into one chip, both the leakage currents induced static power and Joule heat-induced dynamic power are challenging problems. Simultaneously, quantum tunneling effect and matter-wave characteristics will make the transistor extremely unreliable [10]. In addition, design and fabrication cost greatly block the technological evolution. Therefore, it is difficult to further improve the computing performance by relying solely on transistor miniaturization.

At present, IC has entered the post Moore's era from the "Moore's law era". As shown in Figure 2, "More Moore", "More than Moore", and "beyond-CMOS" are three main development directions [11,12]. For instance, FinFETs with new materials and new structures have been developed as a "More Moore" technology. The possibility to realize the diversification of functionalities of CMOS circuits by integration with other technologies has been referred to as "More than Moore". The "More than Moore" approach allows for incorporating digital and non-digital functionality, e.g., power control, analog circuits, radio frequency (RF) circuits, passive components, sensors, and actuators, into compact systems with the system on chip (SoC) or system in package (SiP) technologies. An example that has witnessed considerable commercial success is provided by CMOS imagers which can be found in any cell phone camera. The optical sensors based on Si photodetectors or phototransistors are monolithically integrated on a CMOS chip. Besides, the integration of different technologies, for instance, Si and III-V semiconductors, is also appealing. For example, optical passive components can be fabricated with Si technology, which could provide guiding, routing, and other optical functionality directly on chip. Light sources can be fabricated with compound semiconductors like GaAs and InP. These chips fabricated with different technologies can be further integrated with three-dimensional (3D) or heterogeneous integration technology, which can take advantage of the integration density of Si CMOS and the superior performance of non-Si devices or components.

Meanwhile, optoelectronic technology is an important branch of electronic information technology. Significant achievements have been made in optoelectronic technology and industry. Optoelectronic devices and integration technology have many obvious advantages, such as low power consumption, high speed, and wide bandwidth. They are widely used in optical communication, sensing, computing, displaying, and illumination. At present, photonic integrated circuit (PIC) mainly relies on different integrated material systems, such as III-V group, silicon-based, silicon nitride/silicon dioxide, lithium niobate (LN) material to realize chip integration. According to different functionalities, distinct material systems are required. The coexistence of multiple material systems will become the state of optoelectronic integration

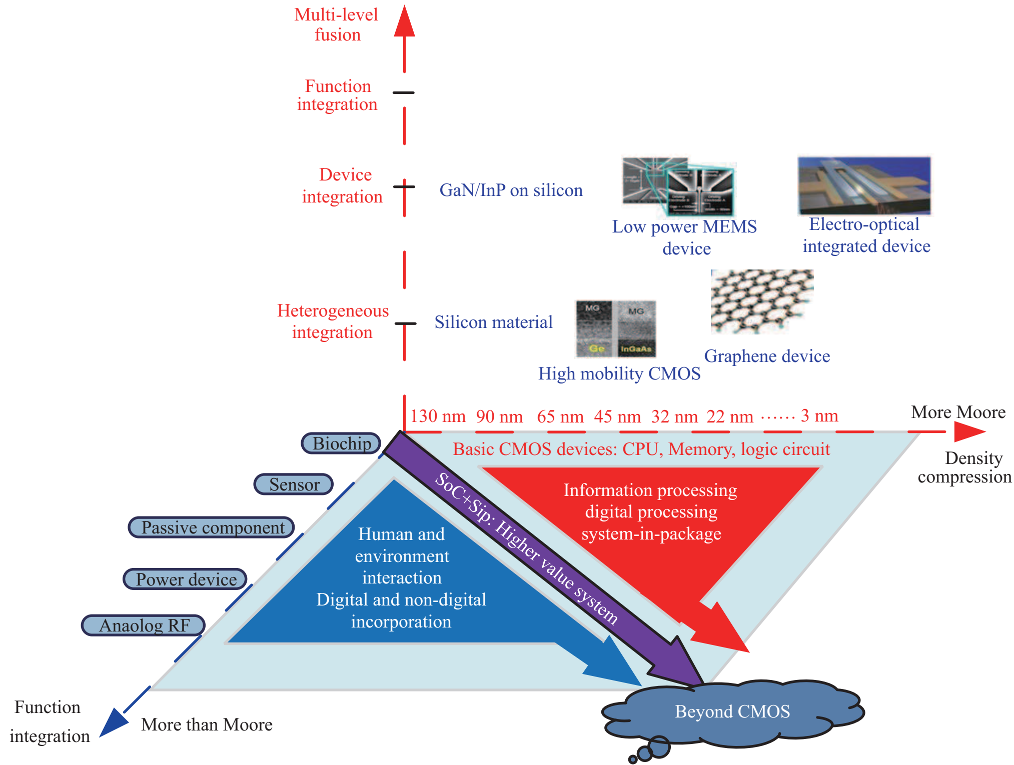

Figure 2 (Color online) A roadmap for IC development in the post Moore's era [11].

technology for a long time in the future. How to realize the system integration of multiple material systems and multiple functional devices needs to be solved, which is quite different from the microelectronic IC technology. The advanced IC and optoelectronic chips are the key foundations to further promote high-performance computing chips and transmission chips. In recent years, many innovative devices, chips, and integration technologies have been developed to further increase the IC performance. Herein, we review the recent progress of the IC and optoelectronic chips in China and abroad. The research status and development trend of devices, chips and integrated technologies of typical IC and optoelectronic chip are focused on. The organization of this paper is shown in Figure 3. In Section 2, we focus on the typical IC technologies and chips. The main contents include the advanced CMOS device and technology in the post Moore's era, emerging memory and chip architecture, IC design and electronic design automation (EDA) technology, cross-dimensional heterogeneous integration, AI chip and brain-like chip, carbon nanotube (CNT) electronic chip, wide and ultra-wide bandgap devices, quantum chip technology, and flexible electronics chip (FEC). In Section 3, we present the advances of the optoelectronic chip and integration technology. The hybrid photonic integration technology, silicon photonics, microwave photonic chip, photonic neural computing technology, and GaN-based optoelectronic integration are discussed. Finally, we summarize the challenges and opportunities faced by IC and optoelectronic chips, and propose promising solutions and perspectives.

2 Typical IC technologies and chips

2.1 Advanced CMOS device and technology

As predicted by "Moore's law", with the downsizing of the devices, the chip is expected to achieve better PPAC (performance, power, area, and cost) than the prior node. However, these aggressive scaling targets have become more and more difficult to achieve at advanced technology nodes with numerous new challenges arising [13–15]. From the transistors perspective, referred to as 'front end of the line' (FEOL), the challenge is to overcome the short channel effects (SCE), i.e., the difficulty to effectively turn off the transistors of nanometer-scale gate lengths, and to boost the transistor drive current simultaneously. Over the past 20 years, emerging technologies, such as strain Si technology, high-K metal gate technology,

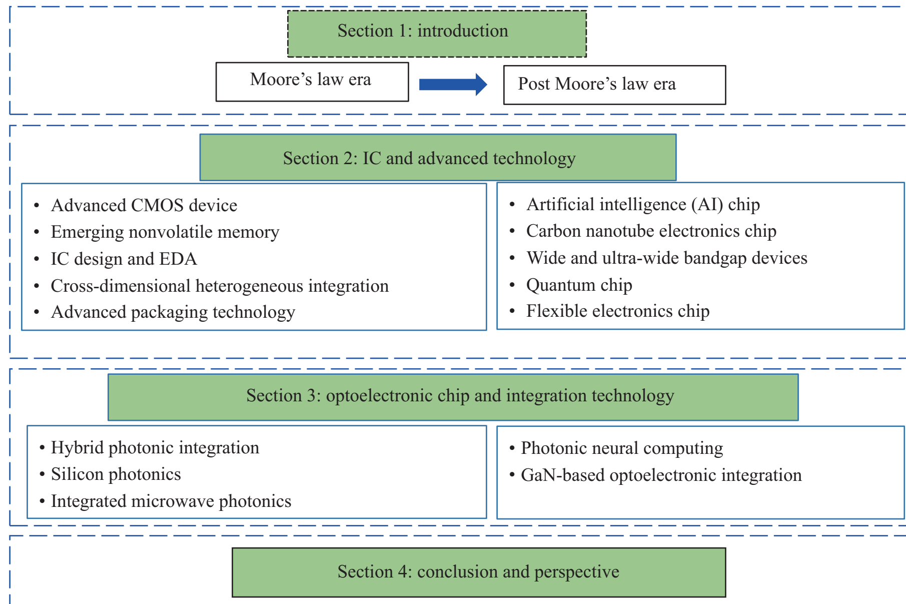

Figure 3 (Color online) Organization of this paper.

and FinFET technology have been widely adopted in mainstream CMOS manufacture to maintain the CMOS scaling. From the interconnect perspective, referred to as 'back end of the line' (BEOL), the challenge is to control the RC delay from the more and more sophisticated wiring. Technologies like advanced silicide, novel lower resistivity metal materials, and low-K dielectric are widely adopted into the advanced CMOS platform.

To maintain a profitable PPAC scaling, several key innovations should be explored and implemented in the advanced CMOS technology. With gate length down to 12–15 nm at 3 nm technology node and beyond, the SCE becomes out of control and FinFET does not have good capability to turn off the channel underneath the Fin, causing significant degradation of subthreshold swing (SS $>$ 75 mV/dec) [14]. Besides, beyond 5 nm technology node, the Fin aspect-ratio will be larger than 10:1, which can cause Fin bending and eventual collapse. For continuing CMOS technology scaling down, gate-all-around (GAA) transistor will replace FinFET as the mainstream [16]. The new device architecture requires a lot of innovations before putting into production, such as channel formation, inner-spacer, gate stacks, novel channel strain engineering, and some unique characterizations of channel integrity and channel strain evolution along with the process flow. Tremendous efforts have been spent to overcome all the above challenges for the production of GAA CMOS technology beyond 3 nm, as shown in Figure 4.

Furthermore, to consider the continuing development of the GAA technology, 3D integration and two-dimensional (2D) material logic technology have to be intensively explored [17]. As the IC technology continuously scales down to 3 nm node and beyond, vertically integrated devices such as vertical nanowire transistors and complementary FET will show the potential in high density, low cost, and low power after lateral GAA technology. Besides the device-level vertical integration, the monolithic 3D integration or 3D sequential integration (3DSI) is a more effective way to build high density and high energy-efficient computing systems [18]. The main challenge for 3DSI is the thermal budget which not only affects the reliability of bottom devices but also limits the quality of top devices. 3DSI has provided new opportunities in boosting the high-performance computing, energy-efficiency computing along the "More Moore" route while cost reduction and smart sensor along with the "More than Moore" road. 3DSI also provides a lot of room for scientific exploration in materials, algorithms, architectures, and even ecosystems.

Benefit from the reduction of spatial dimensions, 2D semiconductors can break the quantum confinement in bulk materials, and are expected to provide high-performance, low-power computing at the sub-3 nm technology node. Moreover, taking advantage of low temperature processes and van der Waals (vdW) integration of 2D materials, low-cost BEOL integration with current technology shows unparalleled advantages and potential of commercialization [14].

From the process perspective, the main challenges focus on advanced photolithography technology,

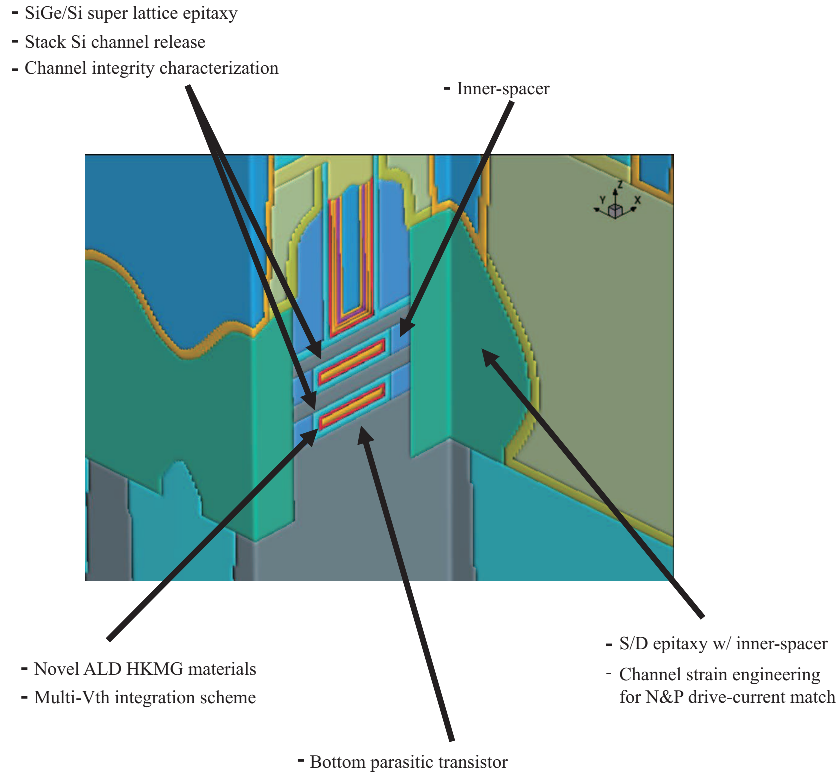

Figure 4 (Color online) The key technology challenges for GAA transistor implementation.

suppression of parasitic resistance and capacitance, and control of reliability and its modeling. For future advanced CMOS technology development, a joint design-technology co-optimization plays a pivot role in the pathfinding of appropriate device, layout, and process integration to meet the profitable PPAC scaling, with much better efficiency and lower cost [19].

2.2 Emerging nonvolatile memory: devices and chips

In conventional von Neumann computing architecture, the processing unit and the memory unit are physically separated, and data communication between the two components relies on a data bus with limited bandwidth. The data communication results in bottlenecks in computing performance and power consumption. Particularly, in the era of big data-driven AI, data communication is more frequent and the amount of data that needs to be stored and processed is far greater than before. The "von Neumann bottleneck" is becoming more prominent with the explosive growth of data volume.

As a result, the search for a good memory medium that supports such an integration of data storage and processing in a single die has been accelerated in the field of microelectronics. In this context, emerging nonvolatile memory (NVM) technologies, such as resistive RAM (RRAM), phase-change memory (PCM), and magnetic RAM (MRAM), have attracted considerable attention in academia and industry [20–22]. Mainstream IC companies, e.g., Intel, Samsung, TSMC, and IBM, and several startups, e.g., Crossbar and Avalanche are involved in the development of these nonvolatile memories. As can be seen from Figure 5(a) [23], RRAM devices comprise metal-insulator-metal (MIM) stacks, and the resistive switching process typically involves the creation and disruption of conductive filaments (CF) comprising a localized concentration of defects. A low resistance state (LRS) corresponds to CFs bridging the two metal layers. Even though the history of RRAM can be traced back to at least the 1960s [24], key technological demonstrations in the 2000s [25] gave significant impetus to this technology. PCM, which also dates back to the 1960s [26], is based on the property of certain types of materials, such as $Ge_2Sb_2Te_5$, to undergo a Joule heating-induced, rapid and reversible transition from a highly resistive amorphous phase to a highly conductive crystalline phase [27]. As shown in Figure 5(b), a typical PCM device has a mushroom shape where the bottom electrode confines heat and current. This results in a near-hemispherical shape of the amorphous region in the high resistance state (HRS). By crystallizing the amorphous region, the LRS state is obtained. MRAM is a product of spintronics, which manipulates the spin freedom in addition to

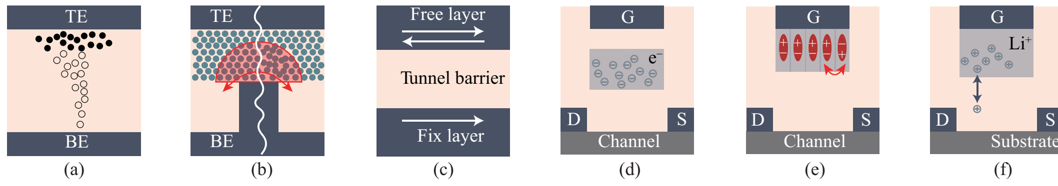

Figure 5 (Color online) (a)–(c) Two-terminal NVM devices. (a) An RRAM device in the LRS where the CF comprises a large concentration of defects for example oxygen vacancies in metal oxides or metallic ions injected from the electrodes. (b) A mushroom-type PCM device in the HRS state where the amorphous phase blocks the bottom electrode. (c) An STT-MRAM device with two ferromagnetic layers (pinned and free) separated by a tunnel oxide layer. (d)–(f) Three-terminal NVM devices: flash memory (d), FeRAM (e), and ECRAM (f). The FeRAM device (e) utilizes the partial polarization switching within the ferroelectric gate oxide to change conductance. The conductance tuning of an ECRAM device (f) is based on the motion of Li ions between the solid-state electrolyte and tungsten oxide. Reprinted with permission from [23] @Copyright 2020 Nature Publishing Group.

the charge of electrons to carry and process information [28]. An MRAM cell consists of a magnetic tunnel junction (MTJ) structure with two ferromagnetic metal layers (pinned and free), which could be made of the CoFeB alloy and thin tunnel oxide such as MgO (Figure 5(c)). In the pinned layer, the magnetic polarization is structurally fixed to act as a reference, whereas in the free layer it is free to change during the write operation. Voltage pulses of opposite polarities are applied to switch the polarization of the free layer. Depending on whether the two ferromagnetic polarizations are parallel or anti-parallel, the LRS and HRS states are obtained due to the tunnel magnetoresistive effect [29].

The abovementioned three NVMs are among the most promising candidates for mass production, especially the Spin transfer torque MRAM (STT-MRAM), which stores information in the magnetization of a nanoscale magnet. The STT-MRAM has attracted wide attention owing to its low power consumption and high performance, and STT-MRAM chip demonstrations have advanced significantly in recent years. Standalone STT-MRAM products became commercially available in 2015 with a 64-Mb part [30]. Subsequently, leading semiconductor industries and tool suppliers have aggressively launched their development programs for STT-MRAM in terms of either eFlash or SRAM replacement. For example, Intel announced NOR-Flash replacement, while Samsung and Everspin/GF announced the release of a 1 Gb STT-MRAM on the 28 nm node [31]. Besides STT-MRAM, RRAM, and PCM operate based on the rearrangement of atomic configurations and hence have worse access times (write speed) and cycling endurance than MRAM. However, they have substantially larger resistance windows that enable the storage of intermediate resistances even at an array level. RRAM has the advantage of using materials that are common in semiconductor manufacturing. Despite the simplicity of the device concept, a comprehensive understanding of the switching mechanism is still lacking compared to PCM and MRAM.

Other candidates that are approaching commercialization include Ferroelectric RAM (FeRAM) [32] and electrochemical random-access memory (ECRAM) [33] as shown in Figures 5(d)–(f). Three-terminal NVM devices provide better control of conductance tuning capability through the modulation from an additional gate terminal. In comparison, flash memory (Figure 5(d)) is the most mature NVM technology. The main challenge regarding flash memory technology for computing applications lies in its slow programming speed, which significantly increases the learning energy. FeRAM (Figure 5(e)) has gained attention due to its faster speed, lower programming voltage, and nearly symmetric channel conductance tuning ability using carefully designed programming schemes [34]. Very recently, a lot of emerging electrochemical synaptic devices have been developed (Figure 5(f)), showing excellent analog conductance tuning performance.

Novel NVM devices which store information using different physical mechanisms offer a unique opportunity to bring large amounts of memory closer to the computing elements, resulting in high-bandwidth, low-latency access. The in-memory computing is an appealing technique that can break through the "memory wall". With the ability to tightly integrate massive amounts of memory with logic, it is conceivable that future computing chips may be much more compact and energy-efficient.

2.3 IC design: AI and IC design, ADC, millimeter-wave and THz IC

The AI chip and high-performance computing, ultra-high speed analog-to-digital converter (ADC) and millimeter-wave (MMW) and terahertz (THz) ICs are hotspots and in short-needed technologies for high-tech developments as shown in Figure 6.

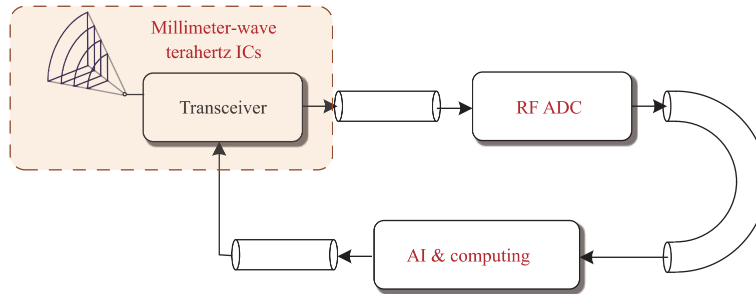

Figure 6 (Color online) The key ICs of high-performance electronic and communication systems.

2.3.1 *Chip design for AI high-performance computing*

In the era of AI, IC is experiencing rapid transformations prompted by the increasing amount of calculation and applications of AI in the field of the chip design. With the rapid development of AI, the computation power of traditional devices like CPU and GPU can hardly keep pace with the development of algorithms such as accelerated computing of deep neural networks demanded by AI applications. This brings challenges to the chip design. Meanwhile, the application of AI in the field of chip design also brings new opportunities for chip development.

For the purpose of eliminating the performance limit of the von Neumann architecture, one solution to realize a better AI accelerator chip design is to make it more brain-like [35]. One example would be the TrueNorth chip invented by IBM in 2015. It integrates 5.4 billion silicon transistors to form 4096 physical cores. The processing capacity of the TrueNorth is equivalent to a supercomputer [36]. Intel has also released a brain-like chip named "Intel Loihi", which applied a neural mimic calculation method imitating the principle of the human brain [37]. It has an asynchronous circuit that uses an asynchronous spiking neural network (SNN), and therefore eliminates the usage of a global clock signal. Intel claims that Loihi is up to 1000 times faster than traditional CPUs and is up to 10000 times more energy efficient in specific applications such as sparse coding, schema search, simultaneous positioning and mapping, and constraint satisfaction problems. Although there are some examples of brain-like chips, it still has a long way to go before putting the brain-like chips into practical applications.

In addition, to improve chip performance from an architectural perspective, the chiplet is also a good choice for improving performance in other aspects in recent years [38]. By integrating cores of different architectures, chiplet does not only lower the cost, but also helps to reduce the power density.

Overall, to meet the ever-increasing need for AI computing ability, the design of ICs has three directions to explore. Firstly, we can customize the neural network accelerator according to the different characteristics of the neural network, which can bring high performance to the specific algorithm. Secondly, we may give up the traditional von Neumann architecture, and turn the eyes to the brain-like chip architecture. Thirdly, a new packaging technology like chiplet can be adopted.

To get a promising performance under specific network accelerators, designers will need to do a full-stack optimization rather than keep eyes merely on the chip design or algorithm research. The optimization of each individual stage of AI chip design and manufacturing is gradually approaching its threshold, and requires joint optimization at all stages, such as the construction of the low-level AI development environment, the design concept of the fully self-developed AI computing framework, and the development of the full scene AI application. Furthermore, with the regularization of AI chips, full-stack optimization is now possible in the field of AI chips. To fully tap into the accelerator's potential, designers must take into account the software features while designing the hardware.

2.3.2 *AI application in chip design and architecture*

Not only is the chip designed to meet the requirements of the AI computation market, AI also plays a more and more important role in the domain of IC design. For instance, there are a lot of problems in the 5G system that are yet to be solved, including the network self-organization and self-optimization and time-frequency optimal allocation. These problems can be easily solved using AI techniques [39]. Furthermore, 6G can be transformative and will revolutionize the wireless from 5G's "connected things" to "connected intelligence". One key feature of 6G will be its ability to exploit a flexible subnetwork-wide evolution to effectively adapt to the local environments and user demands, thereby resulting in a

"network of subnetworks". In fact, in the white paper of Samsung, Samsung clearly proposed that each 6G component used the AI technology to optimize the overall system performance and network operation. Using AI modules inside a chip to obtain high performance and the low energy consumption is also a promising trend to develop AI chips.

2.3.3 *IC design law for ultra-high speed ADC*

Applications like wireless communications and Ethernet networks are in great demand for ultra-high-speed ADCs (or RF ADCs) [40–43]. Both academia and industry paid extensive efforts to promote the development of RF ADCs in the past decades from the following three aspects. The first one is to improve the ADC sampling speed by interleaving more channels. For example, Fujitsu released 256-channel time-interleaved ADC recently, which achieved 56–72 GS/s with an 8-bit resolution. The second one is to use compound track and hold amplifier (THA) or heterogeneous integration of ADC core to achieve high bandwidth. Third, the continuous forward finer CMOS process makes it possible to realize RF ADC with high resolution and low power, since it allows significant digital signal processing to be integrated on the chip. In 2020, ADI reported a 12-bit 18 GS/s RF ADC with 16 nm FinFET.

The main technical challenges of high-speed RF ADCs are as follows. (1) As the process scales down, the lower power supply decreases the voltage headroom for the analog circuit design. Thus, the signal-noise-ratio (SNR) of the RF ADC becomes smaller. (2) Although the smaller gate length and feature size of the CMOS transistor makes the digital gates get faster and smaller, it lowers the transistor intrinsic gain and results in a large error and precision degradation in the analog domain. (3) The increasing of device mismatch worsens the linearity and exacerbates the distortion. (4) The analog signal bandwidth is limited not only by devices but by metal interconnection lines. This issue becomes worsen in RF ADCs with several tens of GS/s. (5) Clock jitter should be even less than several femtoseconds in RF ADCs to achieve high SNR, which requires higher accuracy and reliability of clock generation systems.

For the ADC designers, there are two possible ways to further improve the RF ADC performance and break the technology limitation. One is optimizing the performance from the circuit level. For example, using a fast and low-power dynamic amplifier instead of the conventional power-hungry operational amplifier when designing pipeline ADCs. Other circuit design techniques such as high-order dynamic element matching and advanced background calibration can also help to improve the RF ADC performance. The other way is developing novel ADC architectures from a system level. Hybrid architectures like time-domain RF ADCs can take full advantage of technology scaling down and improve the ADC power efficiency. Besides, micro-system integration is a promising trend of RF ADCs, which combines compound high bandwidth THA and silicon ADC core to achieve both high sampling speed and high bandwidth.

2.3.4 *Millimeter-wave and THz integrated circuits*

MMW and THz frequency bands have a large number of potentially underutilized spectrum resources and have the potential to achieve ultra-high-speed communications, radars, electronic warfare, and other applications. Owing to the unique low cost and high integration of the silicon-based technology, researchers began to investigate the silicon-based MMW and THz ICs.

In 2008, Seok et al. [44] reported a 410 GHz voltage-controlled oscillator based on a 45 nm CMOS process demonstrated with a push-push structure. This work was the first reported silicon-based THz source in the THz frequency band, but its output power was very small, only −47 dBm (20 nW). In 2012, Sengupta et al. [45] demonstrated a 4×4 array of 280 GHz source based on the 45 nm CMOS process, with the output power reaching 1 dBm. After that, Afshari et al. [46, 47] did a series of research on silicon-based THz oscillators from 200 to 500 GHz. In 2017, Meng et al. [48] proposed an oscillator design method and realized a 193 GHz voltage controlled oscillator.

The earliest exploration of silicon-based THz transceiver systems mainly adopted on-off keying (OOK) modulation, which requires the lowest clock jitter performance. In 2012, Park et al. [49] adopted a quadrupler-based architecture to implement an OOK-modulated silicon-based THz transceiver that operated at 260 GHz based on harmonics, with a data rate of 10 Gbps, and the power consumption exceeded 1 W. Aiming at the problems of low gain, low output power and low efficiency of the harmonic-based THz transceiver system architecture, in 2013, Wang et al. [50] proposed an OOK-modulated silicon-based THz transceiver based on the fundamental frequency and working at 210 GHz. In 2015, Deng et al. [51] reported a 320 GHz OOK transmitter based on a 0.13 μm SiGe BiCMOS process.

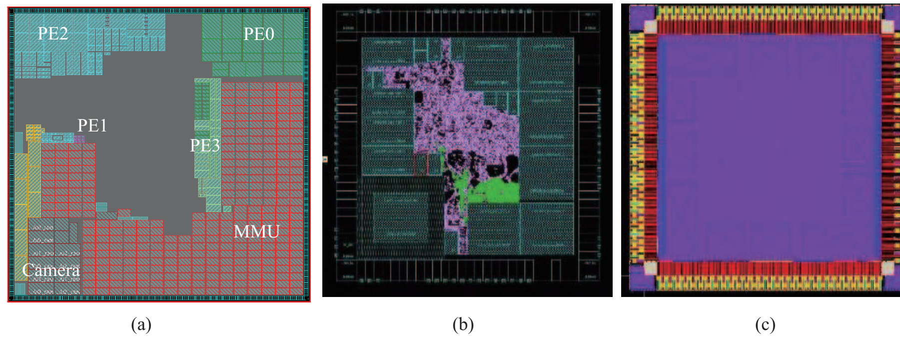

Figure 7 (Color online) (a) Aerospace level chip layout; (b) chip layout of RISCV; (c) signal processing chip layout.

The current progress on the MMW and THz ICs has demonstrated the great potential of silicon-based ICs at high frequency. However, there are still huge challenges for the design of MMW and THz ICs. (1) The silicon-based transistor model provided by the CMOS foundry is not suitable for the MMW and THz frequency band. (2) The output power and power-added efficiency of silicon-based THz front-end amplifiers are low. (3) The phase noise problem of the silicon-based THz local oscillator frequency source plays an important role in the transmission quality of the signal in terms of the error vector magnitude (EVM). The EVM will significantly deteriorate with the increase of carrier frequency and the modulation order. The above technical challenges make the current THz IC design very difficult, and traditional design methods are no longer applicable. Facing these difficulties and challenges, it is necessary to conduct in-depth research in terms of models, circuits, systems, in terms of THz transistor noise model/large-signal model, near-fmax design methodology, high-gain THz amplifier, low-jitter THz source, THz phased array, and low-loss THz chip interconnection technology to make breakthroughs to prompt the development of new spectrum technologies.

2.4 EDA technology

The EDA is the foundation and tool for the development of the IC industry [52–54]. More specifically, EDA aims to realize the automatic design of electronic products with a computer platform, by integrating electronic technologies, computer technologies, information processing technologies, and intelligent technologies. EDA has completely transformed the way of designing and manufacturing ICs. With EDA, the circuit design, performance analysis, IC layout design, printed circuit board (PCB) layout design, and back-end verification of electronic products are realized by using a computer software platform. For instance, EDA has been successfully applied in circuit design from the first microprocessor (Intel 4004), with 2250 transistors, to the latest multicore processor, with over a billion transistors. Figure 7 shows some chip layouts designed with EDA by researchers from Xidian University.

The development of EDA can be divided into four stages. In the computer aided design (CAD) stage in the 1970s, the computer-aided IC layout editing, PCB layout and wiring were adopted to replace the manual operation. The 1980s was referred to as the stage of computer aided engineering (CAE). Different from CAD, CAE not only has drawing function, but also can realize circuit function design and structure design. With the help of electrical connection netlist, the engineering design can be realized. The main functions of CAE include the schematic input, logic simulation, circuit analysis, automatic layout and wiring, and PCB post analysis. The 1990s corresponds to the EDA stage. The design abstract level was promoted to the language level. The language description can be automatically transformed into circuit through EDA tools, which can greatly improve the design efficiency. The key process is synthesis, namely, transforming hardware description language (HDL) into circuit netlist and layout. Since the 21st century, EDA technology has been further abstracted and systematized, and gradually evolved to electronic system level (ESL) [1].

With the rapid development of the AI technology, the IC design methodology has undergone a subversive change, from the traditional “aided” design methodology based on the analysis and optimization technology to the “intelligent” design methodology based on data-driven machine learning. It is believed that, the realization of intelligent and agile design of ICs through AI and machine learning is the future development trend of IC design methods and tools.

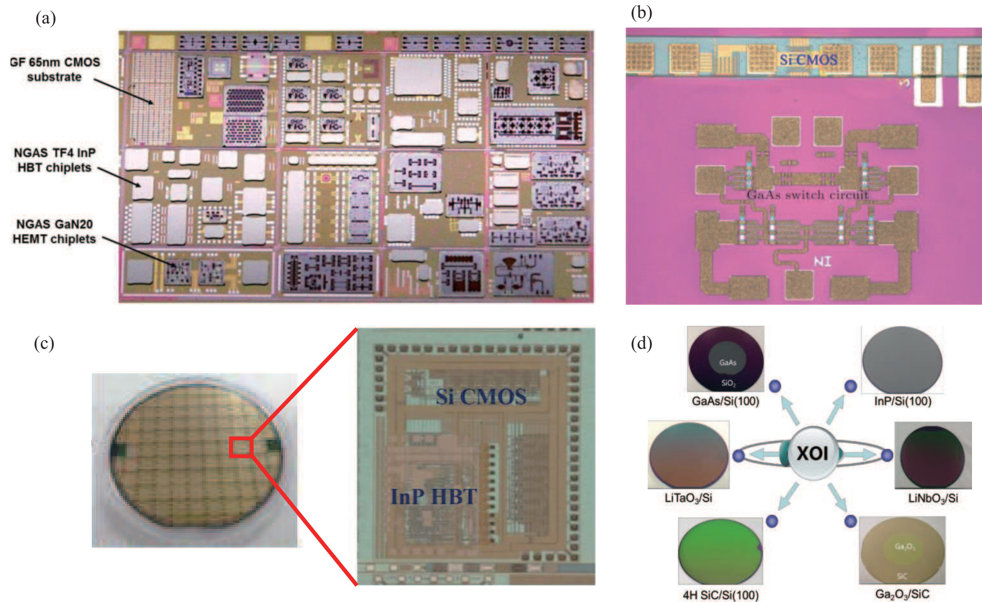

Figure 8 (Color online) Some examples of heterogeneous integration chips. (a) CMOS to III-V chiplet integration [57] @Copyright 2017 IOP Publishing. (b) GaAs pHEMT epi-layer lift-off and transferred on the silicon CMOS circuit [58] @Copyright 2016 IOP Publishing. (c) InP HBT/Si CMOS-based heterogeneous integrated circuit [59] @Copyright 2009 IOP Publishing. (d) Wafer-scale XOI heterogeneous integration materials fabricated by ion-cutting technique.

2.5 Cross-dimensional heterogeneous integration

The cross-dimensional heterogeneous integration technology is a crucial method that does not completely rely on linewidth reduction. It aims to integrate different materials (InP, GaN, GaAs, SiC, Si, organic molecules, etc.), different structures, and different functions (optoelectronic, RF, sensing, biological, etc.) through interdisciplinary and multi-disciplinary integration, architecture design, fusion algorithm and micro/nano 3D integrated manufacturing process [55, 56]. The chip integrated by cross-dimensional heterogeneous integration solution can take full advantages of different materials, devices, and functions, which has a smaller volume, lighter weight, higher performance, and more functions. The key point of cross-dimensional heterogeneous integration is to break through the three integration levels in materials, devices and system functions. Combining the elements of microelectronics, optoelectronics, microelectromechanical systems (MEMS), architectures, and algorithms via cross-dimensional heterogeneous integration with new design ideas, design methods, manufacturing methods, at the micro/nanoscale, a microsystem can be achieved with multi-functional integration of signal perception, signal processing, signaling execution and enabling. Cross-dimensional heterogeneous integration will open up a new way for the development of microelectronic technology in the post Moore's era.

The heterogeneous integration technology has been developed rapidly since the 1990s when Defense Advanced Research Projects Agency (DARPA), the US Department of Defense, first proposed the concept of heterogeneous integration. As shown in Figure 8(a), with chiplet heterogeneous integration technology, GaN, InP, and GaAs chips and MEMS high-Q components were heterogeneously integrated on silicon CMOS wafers, and the distance between heterogeneous interconnects was less than 5 μm [57]. It takes the advantage of the high RF performance of compound semiconductors and the high integration density of CMOS circuits. Based on the epitaxial lift-off and transfer technology, the researchers realized the monolithic integration of the digital control switch circuit of GaAs PHEMT and Si CMOS as illustrated in Figure 8(b) [58]. It has been reported that InP heterojunction bipolar transistors (HBTs) can be integrated on Si materials platforms within silicon fabrication facilities [59], as presented in Figure 8(c). Based on the ion-cutting technology [60–63], the researchers realized different kinds of wafer scale heterogeneous integration materials, e.g., GaAs/Si, InP/Si, $LiTaO_3$/Si, $LiNbO_3$/Si, SiC/Si, and Ga_2O_3/SiC, as shown in Figure 8(d).

The main technical challenges of cross-dimension heterogeneous integration are as follows. (1) Cross-material-dimension heterogeneous integration technology may introduce the material matching, integrated manufacturing, thermal design and electromagnetic compatibility characteristics. (2) Hetero-

geneous integrated transmission technology across device dimensions induces transmission problems in systems of different materials, structures and functions. (3) It is still open to reveal the mechanism of the electromagnetic signal on cross-dimensional heterogeneous integration, especially the mode radiation mechanism and feed coupling mechanism of the electromagnetic signal in different materials, structures, and modules. (4) The integration technology of cross-dimension heterogeneous integrated manufacturing and its compatibility with the existing mature chip level and board level process interconnection integration technology need to be solved. (5) It is highly desirable to develop novel cross-dimensional heterogeneous integration technology design tools.

Cross-dimensional heterogeneous integration technology is one of the best solutions to realize the system with the comprehensive balance of function, performance, cycle and cost. It is a technical bridge from chip to system integration. The characteristic size of silicon-based devices mainly depends on the machining accuracy of the lithography machine, and the lithography machine with the machining accuracy beyond 14 nm has been a long-standing problem in the development of silicon-based ICs in China. Cross-dimensional heterogeneous integration makes full use of various process advantages to achieve circuit performance improvement. It is expected to realize the giant integration of quantum chips, brain-like chips, 3D memory chips, multi-core distributed memory chips, optoelectronic chips, and general-purpose computing chips, and eventually solve the power consumption bottleneck, computing power bottleneck, and function expansion of general-purpose and special-purpose chip technologies.

2.6 Advanced electronic packaging technology

The electronic packaging technology plays an important role in chip performance. The basic functions of electronic packaging include signal distribution, power distribution, heat dissipation, and protection in mechanical, electromagnetic, and chemical aspects.

The typical packaging in the early stage of ICs was through-hole package, such as dual in-line package, single in-line package, and pin grid array package. Owing to the progress of PCB technology and the increase of system complexity, surface mount technology package developed rapidly [1]. In the past two decades, there have been various advanced packaging technologies [64–66], such as flip-chip (FC), multi-chip module (MCM), chip-scale package (CSP), 2.5D/3D integrated packaging based on through-silicon via (TSV), fan-out wafer-level package (FOWLP), fan-out panel level package (FOPLP). At present, wafer-level packaging (WLP) is used in massive manufacturing. In 3D IC, the chip technology and packaging technology are merged closer and closer together. TSV becomes the core of 3D integrated packaging. The SiP technology can be used to integrate different types of discrete devices, such as microprocessors, memory, and sensors into one package [67]. Besides, the chiplet-based technology, which merges multiple heterogeneous dies of diverse functional circuit blocks into a single package, has gained widespread attention. The chiplets prepared with distinct process nodes can be integrated through the SiP technology [68]. Therefore, SiP can provide enhanced performance and a smaller size compared to combining these components at the circuit board level. For instance, the increasing demand for mobile phones and tablets has driven the industry to 3D SiP technology, which can integrate many different functions (digital logic, memory, analog/mixed-signal) with a small form factor. The TSV technology provides an approach for implementing multifunctional integration with a higher packing density for a system in a package.

The development trend for packaging technology is driven by the continuous increase in demands for the smaller form factor, faster, high-density interconnection at cheaper cost.

2.7 AI chip

The wide applications of AI and deep learning call for energy-efficient neural network training. To accelerate the calculation of the deep neural networks while minimizing the energy consumption, AI accelerators have been developed rapidly in recent years.

Artificial neural networks (ANNs) include artificial neurons that receive, process, and transmit signals, and artificial synapses that connect neurons and change the strength of the connection during the learning process. In Figure 9(a), the basic structure of ANN is introduced. The artificial neurons in the input layer receive the signal and propagate it to the neurons in the hidden layer. The signal is first weighted and summed through artificial synapses, then processed by the nonlinear activation function in the hidden layer, and finally propagated to the output layer. The weight of the artificial synapse (W_{ij}) is adjustable during training. The core computation unit of ANN is vector-matrix multiplication (VMM). Figure 9(b)

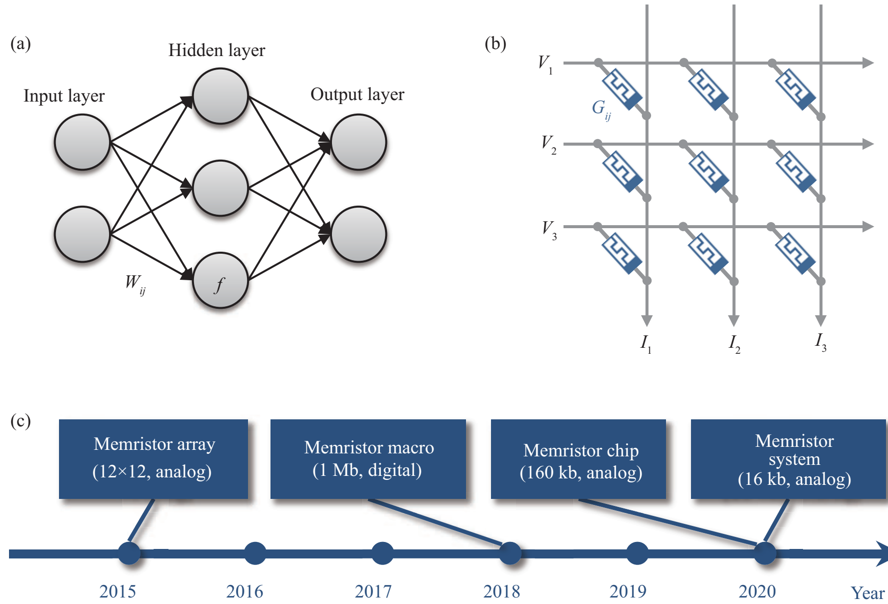

Figure 9 (Color online) AI theory and chips. (a) The diagram of the basic structure of ANN; (b) the memristor array used for VMM; (c) the history of the memristor chips [69–72].

shows a hardware VMM unit based on a memristor array, where the calculation and storage are performed at the same location of the array. The input signal (voltage vector) is passed to each device with different conductivity, and a current is generated according to Ohm's law ($I = VG$, where G is the conductivity) as the calculation result. The current of each column follows Kirchhoff's law, which makes the total current of each column be the sum of currents from all the devices in the same column ($I_j = \sum_i V_i G_{i,j}$). In the ANNs based on a memristor array, each memristor device is an artificial synapse. The conductance of the device can be adjusted continuously and the value can be stored for a long time. Besides, since the weights of artificial synapses may be negative and the conductance of the device is always positive, a common technique is to express the weight as the conductance difference between the two devices.

Several represented studies on memristor chips are presented in Figure 9(c). A passive 12×12 memristor array was proposed in 2015 [69]. A simple 1 layer ANN was realized by such an array and the network can be trained in situ to perform the classification of 3×3-pixel black/white images into three classes. Although the passive array (no transistors in the memristor crossbar) is attractive for its potential in low power consumption, the sneak path current and half-select issues usually limit its large-scale integration. In 2018, a memristor macro with 1 Mb capacity was proposed by using 1T1R active array [70]. Although the macro can perform a convolutional neural network (CNN) with very low latency, the output of the memristor cell is only 1 bit which largely limits its performance. In 2020, a 160 kb memristor chip with a 2T2R active array was proposed [71], where 3-bit signed weight was used in one memristor cell. The chip is fully integrated for a complete multi-layer ANN and achieves 78.4 TOPS/W peak energy efficiency and 94.4% test accuracy in the MNIST dataset. The first memristor system with 8 processing elements (PEs) was proposed in 2020 [72]. Each PE chip consists of a 128×16 memristor array and on-chip decoder circuits. The 5 layer memristor-based CNN was realized in such a system and 96.9% test accuracy in the MNIST dataset was achieved with 110 times better energy efficiency and 30 times better performance density compared with Tesla V100 GPU.

Besides, traditional volatile-memories such as SRAM and DRAM can also be used to realize in-memory computing. The core idea of SRAM based in-memory computing is equivalent to multiply-accumulate (MAC) operation by using XNOR accumulation [73, 74]. Recently, TSMC researchers have proposed an SRAM array with a high power efficiency of 89 TOPS/W which can be further improved by 19 times under a 5 nm design [75]. DRAM-based in-memory computing mainly uses the charge sharing mechanism between DRAM cells [76, 77].

In recent years, the on-chip integration of memristors has greatly developed, and some simple ANNs have been partially or fully implemented by memristor arrays. In the future, larger-scale device integration and more complex ANN implementation will be the focus of research. The high-density 3D integration will be the main technical route and one of the main challenges is to implement high-performance selectors

to eliminate sneak current paths. At the same time, the architecture and data path of the 3D large-scale memristor chips need to be carefully designed to accommodate the high bandwidth and fine-grained interconnection across multiple functional layers. In addition, the future memristor chips need to tolerate the non-ideal characteristics (such as noise and nonlinearity) of the device.

2.8 CNT electronics

CNT has long been regarded as a strong replacement channel material for next-generation high performance and low power FET and IC technology beyond silicon [78], since the discovery of multi-walled CNT in 1991 [79], due to its excellent properties for application as the channel in transistors. They offer extremely high carrier mobility and saturation velocity, which in principle should provide high performance and high-speed device operation [80]. The intrinsically small body (or atomic thickness) of CNT guarantees perfect electrostatic control of carrier movement in the nanotube and they should be free of SCE. Furthermore, CNTs also offer low surface state density and have symmetric band structure (with an identical effective mass of both electrons and holes), which allows a very high gate efficiency, a low supply voltage and the potential to build symmetric CMOS.

The development timeline of CNT FET and CMOS ICs with some major milestones has been summarized in Figure 10. The first FET based on CNT was reported in 1998 [81,82]. CNT FETs showed limited on-state current performance, due to the poor metal/CNT contact. The major significant break to solve the ohmic contact to CNTs was made in 2003 [83], when high work function metal palladium was used as the source and drain contact, thus leading to perfect p-type ballistic CNT FETs. In 2006, the first complex IC (5-stage ring oscillator) was successfully demonstrated and showed 52 MHz operating speed [84]. Although the p-FETs based on CNTs were approaching the performance limit, the development of n-type FET proved to be a slow and frustrating process. In 2007, another important breakthrough in the fabrication of n-type CNT FETs was made. Perfect n-type ohmic contact was observed to form between the scandium contacts and semiconducting CNTs with diameters larger than 1 nm. As a result, performance approaching that of ballistic limits was demonstrated in scandium contacted CNT FETs [85]. Based on the ohmic contact for both n-FETs and p-FETs, the polarity and performance of the CNT FETs can be easily determined by the contact of the CNT to the source/drain, instead of by the CNT channel in between the source/drain [86]. Therefore, the so-called doping-free process to construct CMOS CNT electronics was proposed in 2009 and yielded almost perfectly symmetric CMOS characteristics.

The performance limits of CNT-based CMOS have been investigated using individual CNTs. In 2017, an extremely scaled CNT CMOS FET was demonstrated. To increase the gate control on the channel, a new device structure was developed where both the source and drain were thinned down to atomic thickness (by using graphene) to reduce their electrostatic coupling to the channel. The subthreshold characteristics of the device were thus improved significantly, yielding a remarkably small subthreshold swing of 73 mV/dec at a gate length of 5 nm [87]. To further reduce the subthreshold swing to sub-60 mV/dec level, a new Dirac or cold electron source FET was developed in 2018, which offered sub-60 mV/dec subthreshold swing at room temperature and significantly reduced off-state leakage while delivering comparable on-state current at 0.5 V to that of Intel's 14-nm CMOS FETs at 0.7 V [88].

Although the prototype CNT FETs based on individual CNTs are approaching the physical limit, the requirement of scalability and integration density of devices call for the wafer-scale CNT arrays [89]. The ideal material system is well established to be aligned CNT arrays with a well-defined and consistent 5- to 10-nm pitch, a semiconducting purity estimated to be $> 99.9999\%$ [78, 89]. One way to produce aligned CNT material is based on chemical vapor deposition (CVD) process followed by post-treatment. The major drawbacks of CVD processed aligned CNTs are either low CNT density resulting in poor performance or low semiconducting purity. In 2013, based on the combination of CVD-based process and post-treatment CNT arrays, researchers demonstrated a prototype "computer" that could run an operating system capable of multitasking. However, the performance of this integrated system is very low, the operating speed is only kHz, which is orders of magnitude lower than commercial silicon chips [90]. Another way of producing aligned CNTs is the solution-processed method. The major advantage of this method is its high CNT semiconducting purity, and it could provide wafer-scale assembly capability. In 2020, a method called multiple dispersion and dimension-limited self-alignment was investigated to produce extremely high semiconducting purity and well-aligned CNT arrays (within the alignment of 9 degrees) with a tunable density of 100 to 200 CNTs/μm on a 4-inch silicon wafer. FETs based on the CNT array showed better performance than that of commercial silicon MOSFETs with similar gate

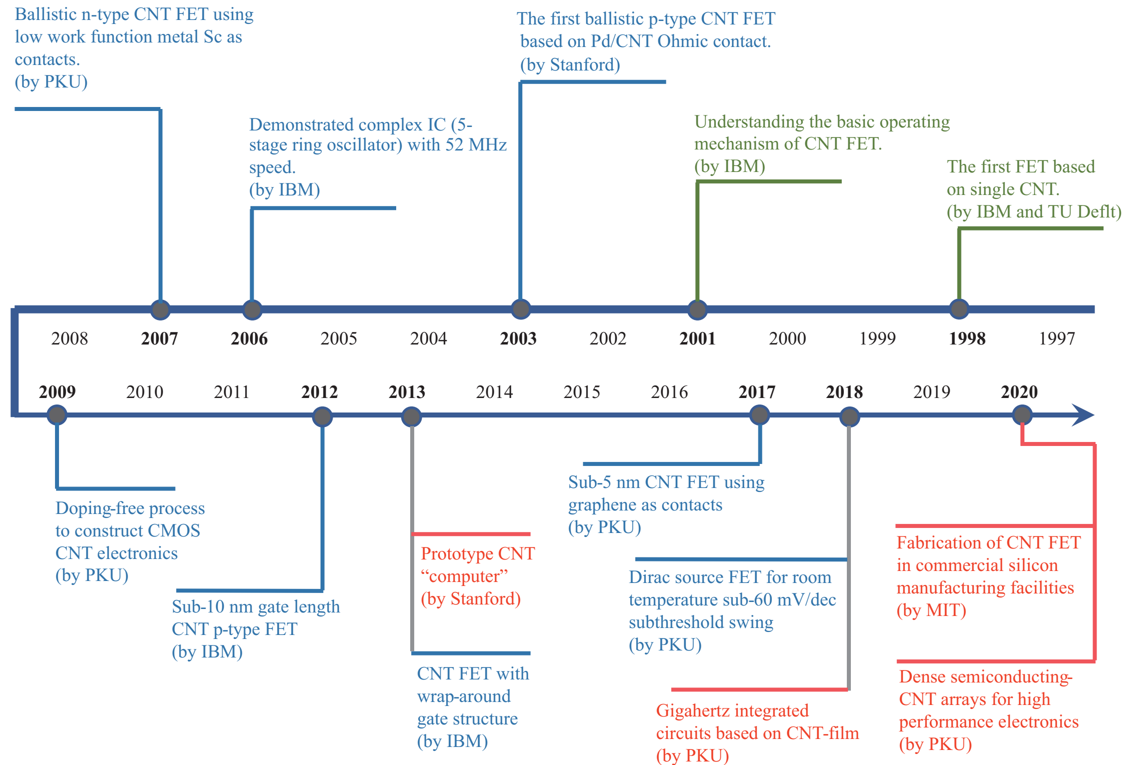

Figure 10 (Color online) The development timeline of CNT FET and CMOS ICs.

length, and batch-fabricated top-gate 5-stage ring oscillators exhibited a highest maximum oscillating frequency of > 8 GHz [91].

CNT-based high-performance CMOS electronics have been scaled down to sub-10 nm planar nodes, and the demonstration of CNT-based ICs has shown both scalability and performance better than silicon-based 0.18 μm CMOS. Further development of this CNT-electronics platform will require optimization of both the material preparation and corresponding device fabrication. Improvement of the uniformity of the tube-to-tube pitch, direction, and diameter of CNTs on a large scale is necessary for ultra-large-scale integration of CNT ICs, particularly for sub-10 nm technology nodes. Moreover, the CNTs in the array need to be further cleaned. At the device level, fabricating CNT CMOS FETs with standard industry processes, which are compatible with conventional CMOS processes, is an important requirement for the development of CNT-based ICs [92]. Finally, the adoption of the multilayer interconnect technology widely used in Si technology and the optimization of the device structure would also be expected to further enhance the actual potential of CNT-based technology.

2.9 Wide and ultra-wide bandgap materials and devices

WBG devices provide obvious advantages for military and commercial systems due to their high output power density, high operating voltage and high input impedance. Note that, to fabricate high-power and high-frequency transistors, semiconductor materials with a large breakdown voltage (BV) and high electron velocity are desired. Table 1 lists the main parameters of GaN and other materials. Johnson's figure of merit (JM) can be calculated to compare the power-frequency limits of different materials [93]. It can be seen from Table 1 that GaN and SiC with higher JM are more suitable for making high-power and high-frequency transistors.

Thanks to the high lattice matching, high thermal conductivity, and good heat dissipation of GaN and SiC, GaN-based HEMTs (high electron mobility transistors) fabricated on SiC substrate can fully meet the requirements including high power, high efficiency and high frequency. However, the cost of SiC is always high, which slows down its commercial process. Commonly used substrate materials for GaN epitaxial growth include sapphire and silicon. Although sapphire can be produced stably at a low cost and large size, its high lattice mismatch with GaN (reaching 14) and poor thermal conductivity can affect the improvement of GaN-based devices in power density. Compared to sapphire, Si material has

Table 1 Main parameters of GaN and other materials [4, 5]

	Si	GaAs	H-SiC	GaN	Diamond
E_g (eV)	1.1	1.42	3.26	3.39	5.45
n_i (cm^{-3})	1.5×10^{10}	1.5×10^{6}	8.2×10^{-9}	1.9×10^{-10}	1.6×10^{-27}
ε_r	11.8	13.1	10	9.0	5.5
μ_n (cm^2/Vs)	1350	8500	700	1200 (bulk) 2000 (2DEG)	1900
v_{sat} (10^7cm/s)	1.0	1.0	2.0	2.5	2.7
E_{br} (MV/cm)	0.3	0.4	3.0	3.3	5.6
Θ (W/cm·K)	1.5	0.43	3.3–4.5	1.3	20
$\mathrm{JM} = \frac{E_{\mathrm{br}} v_{\mathrm{sat}}}{2\pi}$	1.5	0.43	3.3–4.5	1.3	20

larger size, lower cost, and better thermal conductivity. Therefore, it has a larger commercial prospect.

Based on the material properties, GaN-based HEMTs have a lot of advantages over existing semiconductor technologies [94]. Owing to the higher power density, the size of GaN-based HEMTs can be greatly reduced under the same output power so that the matching of the amplifier is easier and the power loss is lower. Moreover, the ability of GaN-based HEMTs to work at high BVs can not only reduce the need for voltage conversion, but also provide the potential for high efficiency, which is a key parameter of the amplifier. Besides, the WBG also enables it to work at high temperatures. Therefore, GaN-based HEMTs are promising in high-power and high-frequency areas.

For the better application of GaN-based devices in high-power and high-frequency areas, tremendous exploratory studies have been completed ranging from material growth, device processing and device structure [95–99].

In 2004, GaN-based HEMTs on SiC with field-plate (FP) length of 1.1 μm were reported, which achieved a continuous wave output power (P_{out}) density of 32.3 W/mm and power-added efficiency (PAE) of 54.8% at 4 GHz [100]. This FP structure can effectively modulate the electric field distribution between the source and drain of the device and reduce the peak value of the electric field, thus increasing the BV and improving the gain and power of devices.

Except for FP structure, SiN_x passivation is also a key to improve characteristics of GaN-based devices [101]. Although SiN_x passivation can be used to reduce the dispersion, the reproducibility of BV, gate leakage, and effectiveness of dispersion elimination are strongly process-related. In 2004, a deep gate recess was used in the fabrication of GaN-based devices to achieve the desired transconductance and a thin $\mathrm{SiO_2}$ layer was deposited on the drain side of the gate recess to reduce gate leakage current and improve BV. A BV of 90 V was achieved without a surface passivation layer. A record Pout density of 12 W/mm with an associated PAE of 40.5% were measured at 10 GHz [102].

In 2010, a hybrid MOSFET structure with a lower on-resistance (R_{on}) and a high BV was reported [103]. By introducing a RESURF, the structure can realize very low R_{on} characteristics. Unlike the Schottky barrier gate used in conventional HEMT devices, MOSFET means that the device can apply a higher gate voltage, which can increase the 2D electron gas concentration and the peak current. Besides, the insulator and AlGaN epitaxial layer can ensure extremely low leakage and allow a large negative to positive gate voltage swing.

In addition to investigations in high-power areas, the RF performances of GaN-based devices have also been studied in recent years due to their potential in MMW field. In 2014, a novel transistor structure, super-lattice castellated FET (SLCFET) based on a GaN super-lattice channel with a 3D gate on SiC substrate was reported [104]. These transistors showed an RF switch figure-of-Merit (FOM) of 2.1 THz, which is 3–10 fold greater than that of the conventional transistors. The f_T and $f_{\max}$ of SLCFET structure with $L_g = 0.25$ μm are 52 and 53 GHz, respectively. In 2015, GaN HEMTs with ultrathin AlN barrier were presented. By using sub-10 nm AlN barrier, devices with 120 nm gate achieved an f_T of 75 GHz and an $f_{\max}$ of 202 GHz [105]. In addition to using AlGaN and AlN barriers, InAlGaN is also a very attractive material. A novel InAlGaN/GaN HEMT with $L_g = 80$ nm was demonstrated in 2015, which also improved the RF performances of devices [106]. In 2016, the devices using N-polar GaN material with 45 nm gate achieved peak f_T of 160 GHz and $f_{\max}$ of 269 GHz [107].

Except for d-mode GaN based HEMTs, e-mode GaN based NMOS transistors with excellent RF characteristics have also been investigated recently. In 2019, high-K dielectric e-mode GaN based transistors fabricated on an HR Si (111) substrate with $L_g = 50$ nm were reported. These devices showed excellent RF performance of $f_T = 190$ GHz and $f_{\max} = 300$ GHz [108]. By introducing robust gate dielectric and

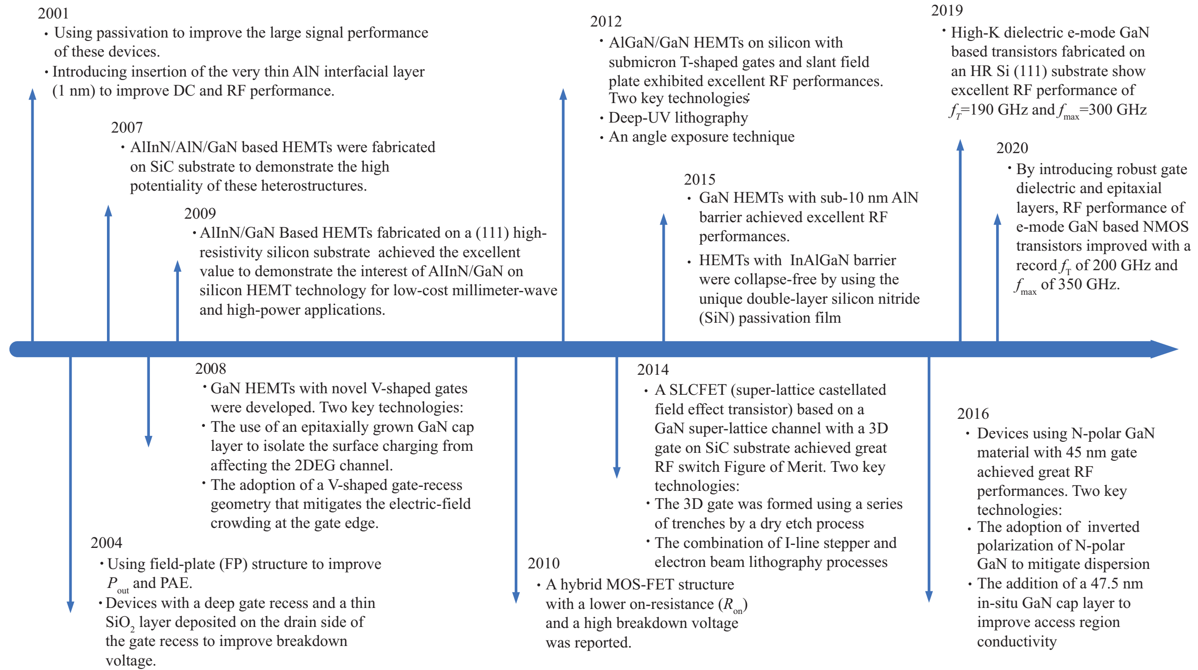

Figure 11 (Color online) Development of GaN based devices in high-power and high-frequency areas.

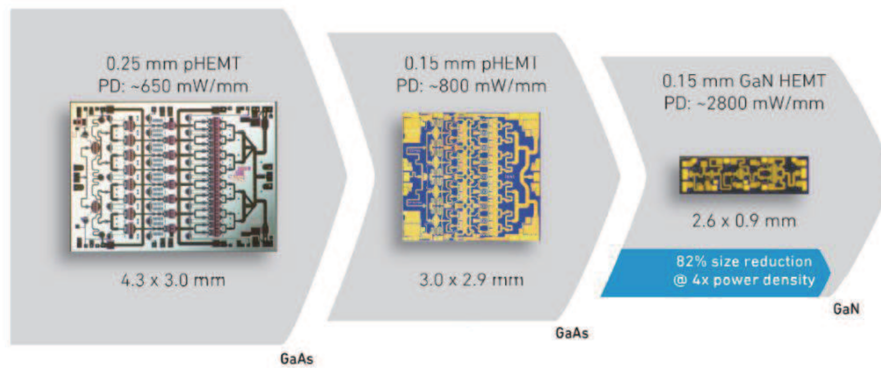

Figure 12 (Color online) Comparison of GaN and GaAs MMIC in size and power density.

epitaxial layers, the RF performance of e-mode GaN based NMOS transistors mentioned above was improved with a record f_T of 200 GHz and $f_{\max}$ of 350 GHz in 2020 [109]. Figure 11 shows the development of GaN based devices in high-power and high-frequency areas.

Figure 12 shows the size and power comparison of the monolithic microwave integrated circuit (MMIC). Based on the powerful performance of the GaN-based devices, the related circuit design has a smaller size and stronger performance than GaAs MMIC.

In addition to three-terminal devices such as HEMT, WBG semiconductors are also suitable for the two-terminal Schottky diodes (SBDs) [110, 111]. The SBD is suitable for the microwave rectification due to low turn-on voltage ($V_{\rm on}$), while the WBG nature endows the SBD to possess a high BV. For instance, in order to improve the ideality of SBD, a GaN SBD with a hybrid anode and a groove-type low work-function metal structure was proposed [110], which showed several advantages such as ultra-low $V_{\rm on}$, high reliability, low leakage, and high consistency. In addition, as shown in Figures 13(a) and (b), a lateral GaN-on-SiC SBD was proposed, which achieved a very low $V_{\rm on}$ of 0.38 V, an ideality factor of 1.03, a subthreshold swing of 62 mV/dec, and a BV of more than 200 V @IR = 3 μA/mm with an electrode spacing of 4 μm [110]. A 5.8 GHz high-power microwave rectification circuit was realized based on this SBD. The single-tube microwave rectification power reached 6 W, and the rectification peak efficiency was higher than 70%, as shown in Figure 13(c). The single-tube circuit power is 50–100 times that of Si and GaAs SBDs, which represents the highest international level of microwave high-power SBD and

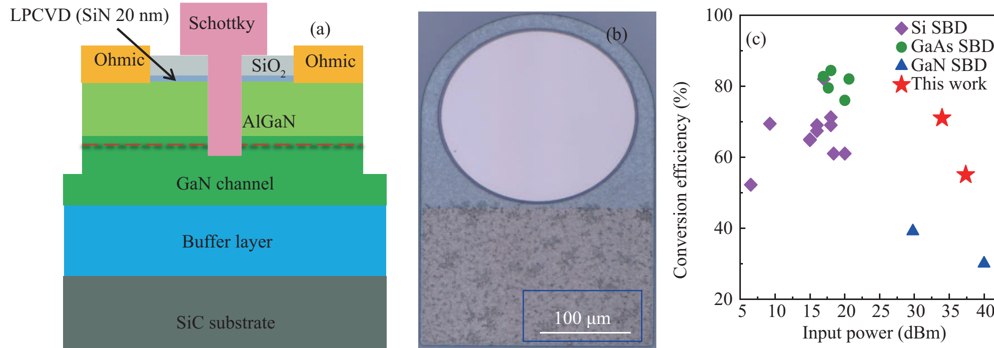

Figure 13 (Color online) (a) Cross-sectional schematic view of the lateral GaN-on-SiC SBD; (b) microscopy image of the fabricated lateral GaN SBD; (c) conversion efficiency versus input power of some state-of-the-art rectifier circuit with Si, GaAs, and vertical GaN SBDs. The groove-type lateral GaN SBD presents the best combination of ηRF/DC and Pin [111] @Copyright 2020 IEEE.

rectifier circuit.

GaN-based SBDs also show a great potential for the next generation power electronics [112–114]. By using lower work-function metal molybdenum (Mo), lateral GaN SBDs demonstrated an ultra-low $V_{\rm on}$ of 0.31 V, a high BV of 2.46 kV with a high-power FOM of 2.65 GW/cm^2 [113]. The development of anode engineering could further improve device characteristics to BV $>$ 3 kV, $V_{\rm on}$ of 0.38 V, and a high-power FOM of more than 3 GW/cm^2 [114].

$\rm Al_xGa_{1-x}N$ as ultra-wide bandgap semiconductor is more suitable for high temperature and high voltage applications due to its larger Eg (3.4 to 6.2 eV) [115–118]. Thanks to the improvements of the material growth quality, $\rm Al_xGa_{1-x}N$-based electronic devices have also made a lot of progresses in recent years. The 5% Al composition AlGaN channel SBD shows better temperature degradation immunity when compared with GaN channel SBD in both on and off states. The leakage current of the AlGaN channel SBD only increased by 3.5 times from 300 to 425 K, while that of GaN channel SBD increased by 27.8 times [115]. As for high-Al application [116], it is suggested to use 85% AlGaN as an alternative channel material by considering the shallower activation energy of donor in GaN [117], without compromising the breakdown characteristics of AlN. Lateral $\rm Al_{0.85}Ga_{0.15}N$ SBDs demonstrate a surface root-mean-square roughness of 0.33 nm, a high BV of 2 kV and an ideality factor of 2.3 [118].

Although GaN-based devices have been extensively studied, there are still lots of problems to be solved. In the future, the heat dissipation problem caused by ultra-high power requires special attention. The research on the reliability of the device is still not in-depth enough. Other problems, related to commercialization, such as gold-free ohmic contact research, need more attention as well. Only when such problems are solved can the commercialization of GaN be further promoted to integrate with existing Si-based CMOS processes.

2.10 Quantum chip

The silicon-based quantum computing processes quantum information based on quantum chips fabricated by modern silicon semiconductor technologies. On the one hand, it continues the Moore's Law of classical semiconductor chips. When the characteristic size of classical processors decreases to 5 nm or less, the quantum tunneling effect appears. Using the quantum tunneling effect to build more powerful processors has become one of the important topics of information processing. On the other hand, it has many excellent characteristics, such as compatibility with modern semiconductor technology, the characteristic scale of nanometer scale, easy integration with classical measurement and control circuits, and is expected to break through the limitation of the extremely low-temperature working environment. It is considered as one of the most promising systems to realize universal quantum computing.

Silicon-based spin qubits, in a narrow sense, refers to quantum dots defined by constrained voltages applied to gates constructed on silicon metal oxide or Si/SiGe heterojunction. In the condition of the filling of a single electron, by applying a certain magnetic field, electron spin state splitting into spin up and spin down. Using this two-level system coding quantum bit 0 and 1 states, we can build a single spin qubit and further realize silicon quantum computing. In a broad sense, the substrate materials used for the preparation of quantum dots can also be Si nanowires, Ge heterojunction, and other semiconductor materials. The formation of single electron can be obtained utilizing donor doping or single donor injection

Year	2004–2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
UNSW	Multi-layer SET	Few electron QD		Charge sensing of QD				Single and two qubit gate		• ST readout • EDSR • Optimized pulse • RB of two qubit gate • Gate sensor			
		Single shot read of donor			Donor qubit		High fidelity		Two qubit gate				
TuDelft	• Single shot read • Two spin states • Single qubit control • EDSR	• Nuclear • Echo		Two qubit gate		Si/SiGe single qubit			Benchmarking		Two qubit gate		RB of two qubit
				Triple quantum dot				Spin shuttle		Simulation			
Princeton		• EDSR of InAs nanowire QD • cQED of InAs nanowire QD					Maser			Strong coupling	Spin photon coupling		
		• Si/SiGe DQD						New Si DQD		QD array	Single and two qubit gate		shuttle
Sandia			MOS DQD		Donor implantation			Cryo-HEMT			• ST qubit • Latching PSB • Dot-donor		New MOS QD
HRL				ST qubit		New DQD			EO qubit		Benchmarking of EO qubit		
CEA					Single QD		Rf sensing		Hole qubit		Electron qubit		rf

Figure 14 (Color online) Research trends of major research groups in semiconductor quantum computing. SET: single electron transistor. EDSR: electric-dipole spin resonance; QD: quantum dot; cQED: circuit quantum electrodynamics; DQD: double quantum dot; ST: singlet-triplet; HEMT: high electron mobility transistor; RB: randomized benchmarking; PSB: Pauli spin blockade; EO: exchange only.

in Si substrate, besides the use of gate trapped quantum dots. The coding of qubits includes not only the single spin state, but also the two-electron spin singleton-triplet state, the spin-charge hybrid state.

Figure 14 shows the research trends of major international research groups in semiconductor quantum computing (including earlier research on III-V semiconductors). In recent years, a series of important research advances have been made in the fields of spin qubit readout, logic gate control, remote qubit coupling, high-temperature qubit. Since 2014, silicon-spin-based single qubit, two qubits manipulation, three qubits entanglement, single photon-spin strong coupling, fast high-fidelity single spin RF-readout, the millimeter-level resonance of two distant spins and high-temperature spin qubits (> 1 K) have been realized successfully [119,120]. Among these experiments, the fidelity of single-qubit manipulation reached 99.9% [121], the fidelity of two qubits manipulation achieved 98% [122], the fidelity of three qubits entanglement reached 88% [123], the fidelity of spin qubit readout reached 99% [124], and the readout time decreased to 1 millisecond. Especially, the 1 K temperature control of silicon-based qubits not only enables the qubit number to be highly integrated, but also enables the control circuits of the qubits to be integrated with the qubits on the same chip, which provides an extended scheme for the practical quantum processor.

Apart from lab studies, some foundry based qubit structures are under development. In 2018, Intel Corporation reported a novel dual nested gate integration process for creating Si quantum dots (QDs) based on their FinFET technology. In 2019, they demonstrated Coulomb blockade in the QD and tunable tunnel coupling between QDs [125], suggesting this gate architecture could be used to form spin qubits. In 2016, several groups also reported single-electron control in a two dimensional array of Si QDs that fabricated using a technology adapted from a commercial fully-depleted silicon-on-insulator transistor (FD-SOI) technology [126]. In 2020, a CMOS quantum IC with a 2.8 GHz excitation and nA current sensing of an on-chip double QD using the same technology was announced [127]. These gate architectures for Si QDs, with the power from industry, may shed light on the realization of large-scale silicon quantum processors.

Although silicon-based quantum computing has developed rapidly in recent years, some fundamental problems restrict its further development. Research difficulties and challenges are summarized as follows. First, high scalability, high fidelity, high-speed qubit readout. The existing qubit readout is realized based on the applied charge detector. Its small detection range and low detection bandwidth limit the expansion scheme of the silicon-based spin qubit. Building detection circuit based on the qubit gate has become the future choice. This technology is effective but difficult, and has not been applied to multi-bit control and bit fidelity measurement. Second, fabrication of substrates with low nuclear spin noise, low charge noise and high homogeneity. High fidelity qubit manipulation has been achieved due to the use of purified Si

substrates, which have reduced the nuclear spin noise in the substrate to a very low level. However, a recent study found that the charge noise, as a low-frequency noise, prevented the further improvement of the fidelity. The uniformity of the substrate is also an important problem, which has a crucial impact on the yield of semiconductor quantum chips. The preparation of high-quality substrates is likely to be one of the potential bottlenecks in the future. Third, design and manufacture of low-temperature readout and control circuits. Current international mainstream qubit measurement is based on room temperature analog and digital control circuits, through the circuits in the refrigerator connected to the chip unit. As the number of qubits expands, likely, existing refrigerators will not be able to support the large wiring. Designing the readout and control circuitry on the same substrate as the qubits will be the solution in the future.

2.11 FEC: research status and development trend

Flexible electronics refer to the fabrication of organic/inorganic materials-based electronic devices on flexible/ductile plastic or thin-metal substrates. Flexible electronics can break through the limitation of classical silicon-based electronics and subvert the original information carrier form, which effectively improves the awareness, storage, and display capabilities of information. FECs are chips made by flexible electronic technology, and FECs not only integrate electronic circuits, electronic components, materials, nanotechnology, and other technologies in the field, but also involve the semiconductor, testing, materials, chemicals, PCBs, display panels, and other industries [128]. Owing to the flexible, ultra-light and ultra-thin characteristics of FECs, they show great potential in flexible wearable devices, flexible bionic devices, flexible brain-machine interfaces, flexible portable devices, and flexible communication devices [129], which are expected to be applied in many future scenes, such as the IoT, mobile health, and smart city, as shown in Figure 15.

To fabricate FECs, flexible electronics are utilized, where special wafer thinning process or directly depositing IC on the flexible substrate is adopted followed by applying mechanics and packaging design methods. The wafer thinning process is the technology that reduces the thickness of the silicon wafer below 50 μm, and such thin thickness improves the flexibility and stability of the wafer. If the thickness of the silicon substrate can be further reduced to 10 μm, the silicon substrate can be almost transparent, and such ultra-thin silicon is an ideal substrate for FECs manufacture. However, the process of cutting silicon substrates down to the micron level is extremely difficult. Another way is to directly deposit ICs on a flexible substrate, which is produced by organic semiconductor materials. Because organic semiconductor materials have the advantages of low cost, high flexibility, and low-temperature processing on any substrate, they could be used for large-scale production with large-area, low cost, and high flux. Most importantly, some specific organic semiconductors with good biocompatibility or biodegradability could enable the seamless interconnection of biological systems with organic flexible electronic arrays, which are suitable for many human-friendly applications, such as electronic skin, smart prosthetics, and wearable human activity/health monitoring devices.

Currently, the flexible electronics and their application in FECs have been rapidly developed, and many excellent achievements and top research teams have emerged. Bao et al. [130] developed organic electronic materials inspired by human skin, enabling the flexible materials to achieve unprecedented application prospects in medical devices, energy storage and environmental applications. Innovative research on soft materials, such as polymers, liquid crystals and biological tissues, can control and induce new electronic and photon responses in these materials [131]. Someya et al. [132] successfully developed the world's lightest and thinnest FECs and applied them in wearable electronic products. Huang et al. [133–137] constructed the theoretical framework of organic optoelectronics and realized the high performance and multi-functional organic semiconductors.

Although showing great potential, FECs still encounter significant challenges. Compared with traditional electronic devices, the performance of various electronic devices composed of FECs is still inadequate. To promote the further development of FECs, some key technical problems need to be overcome. For example, the design, preparation, and processing technologies of materials and components for FECs are desired to be developed to fabricate FECs with superior performance. In addition, design, integration technology, and system architecture scheme for multifunctional FECs are required to expand the application scenarios of FECs. Most importantly, the packaging technology for FECs is highly expected to be explored to ensure the long-term service of FECs.

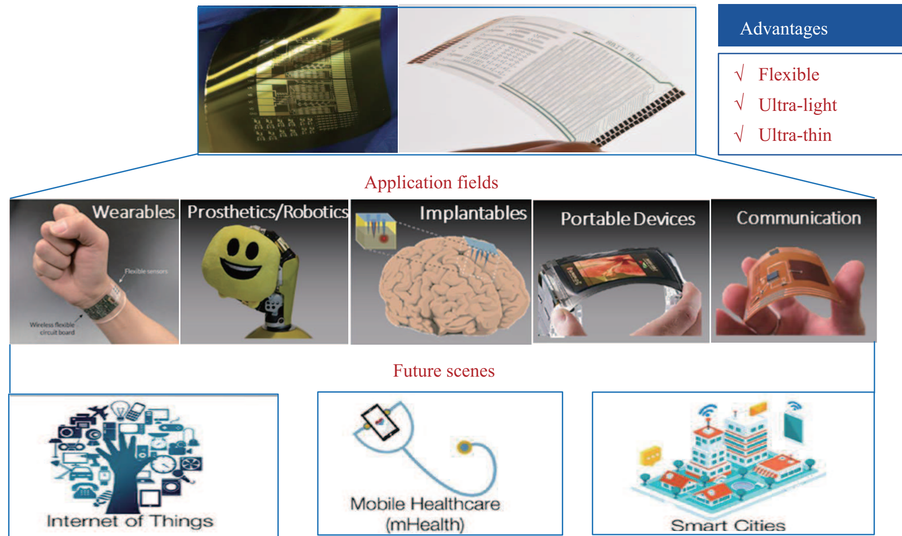

Figure 15 (Color online) The technical advantages, application fields, and future scenes of FECs [128] @Copyright 2018 Springer Nature.

Table 2 Comparison between the different photonic integration technologies[a)]

Optical function	InP	SOI	Si_3N_4/SiO_2	LNOI
Passive waveguide	⋆	⋆ ⋆	⋆ ⋆ ⋆	⋆ ⋆ ⋆
Laser/amplifier	⋆ ⋆ ⋆	–	–	–
Modulator	⋆ ⋆	⋆ ⋆	⋆	⋆ ⋆ ⋆
Switch	⋆ ⋆	⋆ ⋆	⋆	⋆ ⋆ ⋆
Detector	⋆ ⋆ ⋆	⋆ ⋆	–	–
Fiber coupling	⋆	⋆	⋆ ⋆	⋆
Integration scale	⋆ ⋆	⋆ ⋆ ⋆	⋆ ⋆ ⋆	⋆ ⋆ ⋆

a) ⋆⋆⋆ represents very good; – represents challenging/no.

3 Optoelectronic chips: research status and development trend

It is well-known that electronics has dominated information processing while photonics has revolutionized information communication. Thus, photonic circuits and electronic circuits are complementary. PICs have been developed rapidly recently. PICs manipulate lights instead of electrical signals, using on-chip optical waveguides, beam couplers, lasers, electro-optic modulators, and photodetectors [138].

Nowadays photonic integration mainly relies on the following three integration platforms, i.e., InP-based integration, silicon photonic integration (SOI) and silicon nitride/silicon dioxide-based integration (Si_3N_4/SiO_2) [139–142]. Each of these three integration material platforms has unique advantages, but also has some disadvantages. InP-based integration is good at lasing, amplification, detection, but it is very difficult to make active and passive integration circuits with large scale and low traveling loss. SOI is compatible with CMOS integration technologies, but it has difficulty to realize lasers and optical amplifiers. Si_3N_4/SiO_2 has excellent performance to make complex passive integration circuits with low loss, but it is difficult to realize high-speed active devices. Table 2 presents the performance comparison between the different photonic integration technologies, where the emerging lithium niobate on insulator (LNOI) technology shows the excellences of ultra-high-speed electro-optic modulation and low-loss optical waveguides.

3.1 Hybrid photonic integration

The rapid development of information society increases the request for photonic integrated devices in terms of rate, power consumption, size, integration. It is difficult to meet these demands if only depending on a single existing integration platform. It is necessary to employ hybrid photonic integration technology with multi-material systems to make full use of the advantages of different integrated materials, thus high-performance photonic integrated devices can be achieved.

The main hot topics and trends in this area are presented as follows. First, for the existing photonic integrated material systems, they mainly focus on the highly standardized (generic) process of chip fabrication, improving the device performance and increasing the integration scale. Second, for the emerging integrated material system, the main focus is to take the advantages of new materials and overcome the deficiencies of the existing photonic integrated materials in terms of modulation speed, detection, operating band. For example, to increase the electro-optical modulation bandwidth up to 100 GHz and beyond, the thin-film based LNOI photonic integration with ultra-high-speed modulation and low-loss optical waveguide has become a research hotspot [143–145]. To extend the operating band of photonic integration to the mid-infrared region, Chalcogenide photonic integration has attracted much attention [146, 147]. The 2D atomic material, such as graphene, offers high mobility and wide wavelength areas, which also draws much attention in the photonic integration [148]. Third, hybrid photonic integration technology becomes attractive. For example, InP-based and silicon-based hybrid integration has drawn attention to overcome the problem of the luminescence in silicon-based integration [149, 150]. Hybrid integration of photonics and microelectronics within a wafer or between wafers has been emerged to combine the photonic circuits and microelectronics circuits into one chip.

In recent years, some research breakthroughs have been made. In 2019, hybrid silicon and LN Mach-Zehnder modulators for 100 Gbit/s and beyond were demonstrated [151]. The researchers demonstrated the low-noise and low-power avalanche intermediate infrared photodetector based on 2D atomic materials [152]. Hybrid photonic integration is facing the following challenges and opportunities. First, how to employ the successful experiences of CMOS to increase the efficiency of the chip fabrication. Second, how to overcome the disadvantages of the existing integrated materials and improve the chip performance in terms of speed, power, operating band, size. Third, how to realize hybrid integrated chips that include photonic circuits and microelectronic circuits. In a long term, photonic integration must closely collaborate with material science. It is best to develop a new integrated material platform, just like the case in microelectronics, one integration material platform can support the whole device system through CMOS technologies.

3.2 Silicon photonics/optoelectronics

Silicon photonics is a convergence platform of photonics and microelectronics, which can enable complex optical functionalities on a compact chip by using commercial CMOS chip fabrication facilities [153–156]. The silicon integrated optical chips that can generate, modulate, process and detect light signals have been recognized as a promising solution of cost-effectively meeting the ever-increasing demands on data speed and bandwidth.

Over the past two decades, basic key devices such as modulators, photodetectors, multiplexers/demultiplexers, and heterogeneous lasers on a silicon substrate have made considerable progress, while some of them even surpass the performance of traditional devices based on III-V and planar lightwave circuit (PLC) platform. On this basis, silicon-based products used in optical interconnection, high-speed optical switching, and optical transmission have been successfully commercialized [157, 158][1)]. For datacenter optical modules, the 100G products (PSM4 and CWDM4) have explosively grown and the 400G market (DR4 and FR4) has gradually matured. In 2018, SiFotonics launched the 400G DR4 silicon-optical fully integrated chip with the integration of more than 30 active/passive silicon optical devices. For optical switching, silicon photonic engines co-packaged with switching application specific integrated circuit (ASICs) may provide an alternative to pluggable optical transceivers in large data centers, accompanied by a substantial reduction in power consumption and cost. For optical transmission, the silicon photonic technology can realize the integration of coherent modulators and multiplexers/demultiplexers, which is expected to reduce the cost and thus makes the long-distance coherent optical modules sink to the core and convergence layer.

In addition to data centers and long-distance coherent optical modules, silicon-based optoelectronic technologies and devices can also be used in 5G fronthaul networks, quantum communications, sensing, AI, and other fields [159–162]. Intel's 100G silicon optical transceivers can meet the bandwidth requirements of 5G fronthaul applications and meet industrial-grade temperature requirements. In the field of the quantum technology, researchers have developed the most complex optical quantum chip by integrating hundreds of optical quantum devices on one silicon chip, realizing high-dimensional, high-precision,

1) Intel. Product brief: silicon photonics 100G optical transceiver. 2019. https://www.intel.cn/content/www/cn/zh/architecture-and-technology/silicon-photonics/optical-transceiver-100g-cwdm4-qsfp28-extended-temperature-brief.html.

high-stability and programmable quantum entanglement, quantum manipulation, quantum transport and quantum measurement [159]. In the field of AI, the high-throughput and large-scale matrix operations required by AI processor chips can be completed by silicon optical neural network computing units [160]. It has been shown that optical neural network chips have two orders of magnitude faster than traditional electronic computers. The power consumption is reduced by three orders of magnitude. In the field of sensing, MIT, Voyant Photonics, and other teams have launched all-solid-state light detection and ranging (LIDAR) based on silicon-based optical phased array chips, which have the advantages of high integration, fast scanning speed, small size, and low cost [161]. In terms of new microprocessor technologies, DARPA and Intel are working to achieve the integration of silicon optical chips and high-performance microelectronic chips, and have verified a new generation of FPGAs with integrated silicon optical I/O chips [162]. CPU and ASIC chips are expected to increase the throughput rate of processing by 100 times, while reducing energy consumption to 1/10.

3.3 Integrated microwave photonics

Microwave photonics is a frontier interdisciplinary bridging RF engineering and photonics science, which focuses on the generation, processing, manipulation and measurements of RF signals using advanced optical devices and technologies. It has been considered as one of the key technologies to overcome the speed and bandwidth bottlenecks of modern information systems, with the unique advantages of ultra-wide instantaneous bandwidth, low frequency-dependent loss and immunity to electromagnetic interference.

However, the majority of current microwave photonic systems are constructed by using discrete and bulky optoelectronic devices, which are at a disadvantage compared with their electronic counterparts in terms of size, power consumption, stability and cost. Recently, the integrated microwave photonics paves a new avenue to catch up with and surpass their electronic counterparts. The integrated optoelectronic devices and chips offer advantages such as small volume, light weight, low power computation and broad bandwidth [163–165]. Integrated microwave photonics chips and systems have great potential in many fields such as future air-space-ground integrated networks, next-generation wireless broadband access networks, radar and electronic warfare systems.

In a typical microwave photonic system, the RF signal is first up-converted to the optical domain with the help of a laser and an optical modulator. Then the output of the optical modulator is processed by an optical signal processor. The processed optical signal is finally converted back to the RF domain by a photodetector. The high-performance laser sources, optical modulators and photodetectors are the fundamental devices or modules to implement a microwave photonic system. The performance of these key devices should be optimized to boost the performance of an integrated microwave photonic system. For the integrated lasers, high output power, narrow linewidth, and low noise are desired. For the integrated optical modulators, high-efficiency, large bandwidth, and high linearity are wanted. As for the integrated photodetectors, the required performance includes broad bandwidth, high power handling capability and high responsibility. In recent years, the performance of these key integrated devices is gradually augmented, which lays a strong foundation for integrated microwave photonic chips.

Currently, there are mainly three key material platforms for monolithic integration of microwave photonic systems: InP, SOI, and Si_3N_4. In addition to these three key material platforms, there are other technology platforms that can be used to build integrated microwave photonic systems, such as LN, chalcogenide glass, and 2D materials. Nevertheless, none of these above-mentioned material platforms can fulfill all the required performance for integrated microwave photonic systems by themselves. As a result, hybrid integration among different material platforms, as well as co-integration of photonic and microelectronic circuits are the bright future for integrated microwave photonics. To develop fully integrated microwave photonic chips, key technologies such as compatible manufacturing and packaging of different material platforms, as well as driver design and testing of large-scale integrated optoelectronic chips should be developed.

In recent years, integrated microwave photonic chips with versatile functionalities, multiple channels, and reconfigurability have been developed. Various functional units are desired to be implemented in integrated microwave photonic chips, such as low-noise integrated laser sources and frequency combs, advanced optical modulation, optoelectronic oscillators (OEOs), microwave photonic filters, programmable signal processors, arbitrary waveform generators, microwave photonic RF front ends, microwave photonic beamformers, and photonic ADCs [165]. Shen et al. [166] experimentally demonstrated a CMOS-compatible soliton microcombs co-integrated with a pump laser, which represented a milestone towards

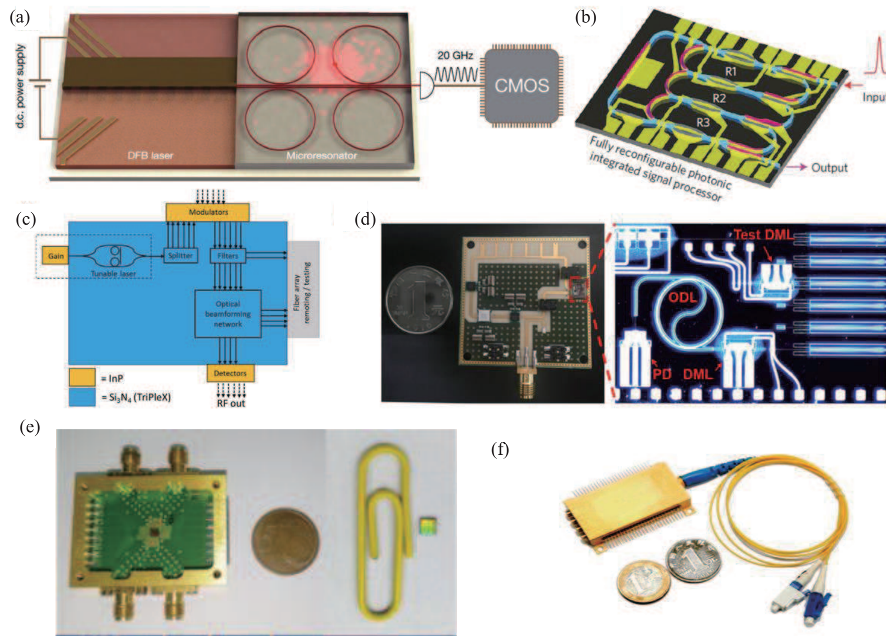

Figure 16 (Color online) Selected functional units implemented in integrated microwave photonics. (a) Low-noise integrated optical frequency combs [166] @Copyright 2020 Nature Publishing Group. (b) Integrated programmable signal processor [167] @Copyright 2016 Nature Publishing Group. (c) Integrated microwave photonic beamformer [168] @Copyright 2019 IEEE. (d) Integrated OEO [169] @Copyright 2018 IEEE. (e) Chip-based microwave-photonic radar for high-resolution imaging [170] @Copyright 2020 John Wiley & Sons. (f) Multifunctional photonic integrated circuit [172] @Copyright 2019 John Wiley & Sons.

mass production of optical frequency combs. Liu et al. [167] experimentally demonstrated a fully reconfigurable photonic integrated signal processor based on an InP-InGaAsP material system, which was capable of performing temporal integration, temporal differentiation and Hilbert transformation. Grootjans et al. [168] demonstrated a broadband continuously tunable delay microwave photonic beamformer based on the hybrid integration of InP and Si_3N_4 chips. Hao et al. [169] demonstrated an integrated OEO featuring monolithic integration of all its photonic parts and RF subsections, where all the photonic parts of the OEO were monolithically integrated on a common InP substrate. In addition, Li et al. [170] experimentally demonstrated a chip-based microwave-photonic radar which enabled high-resolution imaging based on a silicon photonic platform, where high-precision range measurement with a resolution of 2.7 cm and an error of less than 2.75 mm was obtained. Note that, the integrated microwave photonics has been gradually promoted from the proof-of-concept stage to the practical application stage. For instance, Zou et al. [170, 171] demonstrated a multifunctional PIC for diverse applications across the microwave signal generation, transmission, and processing. The PIC was fabricated on an InP platform by monolithically integrating four laser diodes and two modulators. It was further applied to realize remote electromagnetic environment surveillance along in-service high-speed railways. Figure 16 shows some selected functional units implemented in integrated microwave photonics in recent years [167, 169, 170, 172].

At the same time, with the rapid growth of data volume and processing functions in practical applications, multi-channel and multi-band integrated microwave photonic chips are required. Reconfigurability is another important requirement of integrated microwave photonic chips. A network of various tunable active or passive devices should be integrated on such a chip, realizing multi-functional microwave photonic signal processing functions. Moreover, AI can also be used to empower integrated microwave photonics, where neural networks would be adopted to optimize the performance of the integrated chip, thereby promoting the further development of integrated microwave photonics.

3.4 Photonic neuromorphic computing chips: from devices to integrated circuits

As Moore's law becomes slower and the computing consumption of von Neumann's bottleneck can no longer be afforded, neuromorphic computing is one of the promising candidates for a new computation paradigm in the post Moore's era. In the past decades, significant breakthroughs have been made in neuromorphic microelectronics. There are various prototypical neuromorphic processers, including SpiNNaker, Neurogrid, TrueNorth, Loihi, and Tianjic [173]. As an alternative, photonics neuromorphic,

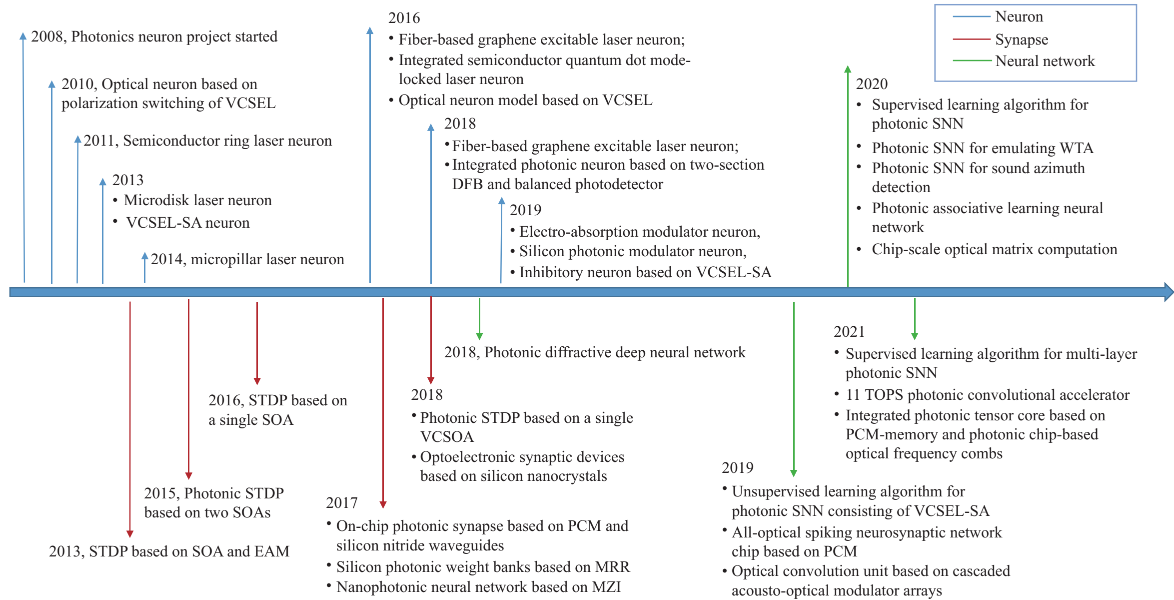

Figure 17 (Color online) Timeline of advances in photonics neuromorphic [174, 178]. VCSEL: vertical-cavity surface-emitting lasers; VCSEL-SA: vertical-cavity surface-emitting lasers with embedded saturable; SOA: semiconductor optical amplifier; EAM: electro-absorption modulator; MRR: microring resonator; MZI: Mach-Zehnder interferometer; VCSOA: vertical-cavity semiconductor optical amplifier; DFB: distributed feedback laser; WTA: winner-take-all.

which is an emerging field that combines the photonics and neuroscience, has attracted lots of attention due to the fascinating advantages such as high speed, wide bandwidth, and massive parallelism [174–178].

In the past decade, tremendous impressive exploratory studies have been completed, ranging from physical mechanisms, devices, architecture, algorithms, and systems to advance the field of photonic neuromorphic computing, as shown in Figure 17. From the device perspective, it has been proved that various optical devices can emulate the neuron-like [179–181] and synapse-like functions [182, 183]. There have been numerous attempts to implement photonics spiking neurons like graphene excitable lasers, distributed feedback lasers, vertical-cavity surface-emitting lasers, or micropillars [174, 178]. The spike temporal encoding, excitatory and inhibitory response, refractory period, and spike-timing dependent plasticity (STDP) have been demonstrated numerically and experimentally. From the architectures and algorithm perspective, different architectures, such as optical matrix computation [184, 185], photonic in-memory computing [186], photonic convolutional neural network [187, 188], photonic reservoir computing [189, 190], and photonic SNN [191, 192], have been explored in recent years.

Very recently, some photonic neural network chips have been successfully demonstrated [160, 193–199]. In 2017, Shen et al. [160] proposed and fabricated a fully optical neural network consisting of 56 Mach-Zehnder interferometers (MZIs) on a silicon PIC, and demonstrated the vowel recognition task. Tait et al. [193] reported a broadcast-and-weight system that was isomorphic to a continuous-time recurrent neural network by using microring weight banks. In 2018, Lin et al. [194] demonstrated a 3D-printed all-optical diffractive deep neural network architecture that implemented classification of handwritten digits and fashion products, as well as the function of an imaging lens at a THz spectrum. In 2019, Feldmann et al. [195] proved that the integrated phase-change photonic SNN, consisting of 4 neurons with 15 synapses each, was capable of simple pattern recognition tasks and can adapt to specific patterns. The spiking neurosynaptic network worked in the optical domain benefits from the high bandwidth and fast data transfer rates intrinsic to light. In 2021, Feldmann et al. [197] demonstrated a computationally specific integrated photonic hardware accelerator (tensor core) for convolutional processing that was capable of operating at speeds of trillions of multiply-accumulate operations per second (10^{12} MAC operations per second or tera-MACs per second). The tensor core achieved parallelized photonic in-memory computing using phase-change-material memory arrays and photonic chip-based optical frequency combs. Such a highly parallelized framework could potentially process an entire image in a single step and at high speed. In addition, Xu et al. [198] demonstrated an optical convolutional accelerator operating at 11.322 TOPS for vector processing, and used a matrix-based approach to perform convolutions of large-scale images with 250000 pixels at a matrix processing speed of 3.8 TOPS.

Photonics neuromorphic is still in its early developmental stage compared with the electronic counterparts. It encounters significant challenges, such as the online training algorithm for the photonic SNN, the large-scale integrated technology for photonic neural network chips, and the real-world applications of photonic neural networks. To obtain significant breakthroughs in the neuromorphic photonics, continuing research and development efforts are required in this interdisciplinary approach involve the neuroscience, device physics, hardware architectures, learning algorithms, photonic integration chips, and neuromorphic applications. Note, thanks to the hybrid optical-electronic integration technology, a hybrid integrated hardware framework that takes full advantage of both photonic and electronic processors could revolutionize the photonic neural network hardware soon, and have important applications in fields of communication, data-center operation, and cloud computing.

3.5 GaN-based optoelectronic integration

The current development of GaN-based devices such as light sources (LEDs and LDs), power transistors (HEMTs and MOSFETs) pave the ways for the GaN-based optoelectronic integration [200–203]. Besides, GaN-based photodetectors with high specific detectivity and fast response speed have been demonstrated [204–206]. The high refractive index of GaN also makes it an ideal material for passive optics, like gratings, distributed Bragg reflectors (DBRs), and resonant cavities [207–209]. InGaN/GaN waveguide and coupler can manipulate light route in monolithic chips with high data transmission speed [210, 211]. All these progresses have shown a strong potential to realize GaN-based optoelectronic integration. In this case, signal transmission, detection, processing, and amplification can be achieved within a single GaN-based chip.

Present effort for the monolithic integration of III-nitride has been put mainly on the partial integration of GaN-based LEDs with other GaN-based electronics or optoelectronic devices. GaN-based transistors, such as HEMTs and MOS, are thought to be good candidates to integrate with GaN-based LEDs for low-power voltage supplies. In this way, residual materials can be removed, parasitic resistance and capacity can be dramatically reduced, and their response times can be reduced [212–215]. Recently, Lu et al. [214] proposed an integrated device between vertical MOSFET and LED (Figure 18(a)). By controlling the gate bias of the VMOSFET, the light intensity emitted from the integrated VMOSFET-LED device could be well modulated. The integration between GaN-based LEDs with foreign power transistors, like Si-based MOSFET, has also been studied. However, these LEDs often have a longer response time than GaN-based transistors due to a lower electron saturation velocity [216].

Monolithic integration among GaN-based LEDs, photodetectors, and waveguides can realize the transformation and processing between the optical signal and electric signal. Photonic integration of an InGaN/GaN multiple-quantum well LED (MQW-LED), waveguide, ring resonator, and photodiode on a Si wafer has been proposed (Figure 18(b)) [217]. The manipulation of in-plane light coupling and propagation by the waveguide and the ring resonator has been illustrated and experimental results show that in-plane full-duplex light communication at a transmission rate of 30 Mbps is achieved. The LEDs and photodetectors can be interconnected by the linear or bent suspended waveguide bridges, which exhibit rapid response on nanosecond time scales and enable transmission of data signals at rates of 250 Mb/s (Figures 18(c) and (d)) [211, 218]. Some studies have shown that optical wireless multiple-input multiple-output systems and advanced modulation and encoding schemes can increase the data rates up to 10 Gb/s [219, 220]. Till now, with improving the crystal quality and structure, GaN-based LEDs with modulation bandwidth more than GHz have already been realized [221–224], these LEDs with integrated power transistors can promote the evolution of future integrated smart-lighting systems.

Considering the relatively low coupling efficiency between LEDs and other GaN devices, GaN laser-based monolithic integrating schemes have also been proposed [225, 226]. Among different optical laser cavities, whispering-gallery-mode (WGM) cavities [227, 228] are the most suitable structure for integration because the internal reflections are along their curved surfaces and can be directionally coupled to external waveguides. In 2018, Tabataba-Vakili et al. [229] demonstrated the first GaN-based WGM blue microlasers integrated with a photonic circuit. However, making a viable photonic circuit platform on silicon with efficient electrical injection in the GaN microdisks remains a challenge for this type of GaN integration [229, 230]. With the development of porous GaN [208, 231], which has a low refractive index, the WGM lasers with lateral porous GaN DBR have been proposed (Figures 18(e) and (f)) [232]. Compared with the traditional mushroom configuration, WGM lasers with porous GaN DBR not only have better device performance, but are more reliable, which lay a foundation for future photonic integration.

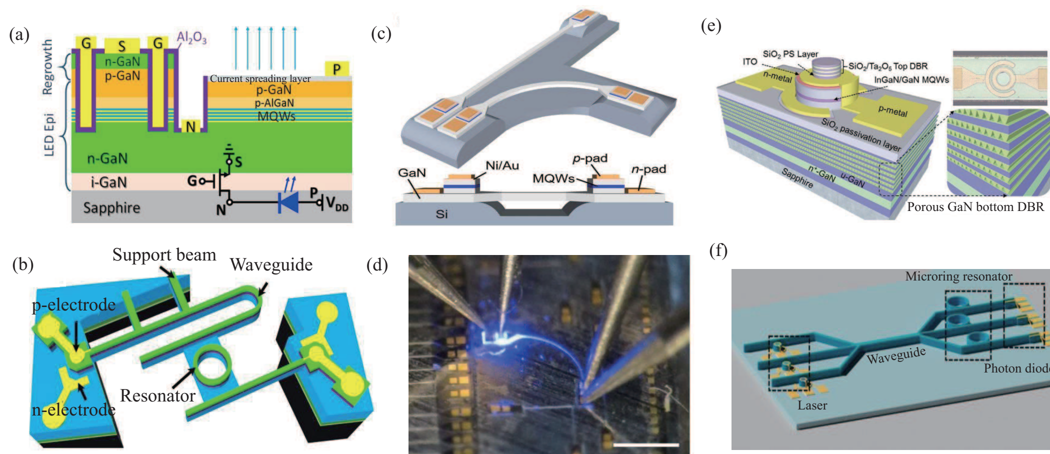

Figure 18 (Color online) Current typical schematic of GaN-based structure for integration. (a) Schematic of monolithically integrated GaN-based MOSFET-LED device and equivalent circuit diagram [214] @Copyright 2016 American Institute of Physics. (b) Schematic of monolithic GaN-based integration of light source, waveguide, ring resonator [217] @Copyright 2017 IOP Publishing. (c) Schematic diagrams of the integration of LEDs, photodetectors, and waveguides [211]. (d) Microphotographs of the integration of LEDs, photodetectors, and waveguides [211]. (e) Schematic and optical image of GaN-based photodiode with a lateral porous GaN DBR [232] @Copyright 2020 John Wiley and Sons. (f) Schematic of optoelectronic integration using WGM GaN-based microdisk lasers.

Present studies have shown a promising prospect for the GaN monolithic integration technique to make a great difference to visible light communications, low-consumption illuminations, and other multi-functionalities. However, despite these efforts, the monolithic integration of III-nitride on a single chip is still only an emerging technology and far from mature. High-quality crystal requirement, growth lattice mismatch, high price and other drawbacks still seriously limit GaN-based integration efficiency.

4 Conclusion and perspective

In the post Moore's era, the rapid development of AI, IoT, edge computing, mobile computing, and 5G makes ICs more vigorous. The traditional von Neumann architecture has been unable to meet the needs of high-efficiency computing. Significant advancement has been achieved from the material, device, IC architecture, integrated technology, design methodology, and packaging technology, which has a revolutionary impact on the current IC technology. In the future, microelectronics, optoelectronics, and intelligent microsystem technologies will develop along the direction of miniaturization, integration and intelligence. Further innovations and breakthroughs from the entire IC industry chain are still desired to ensure the security of the semiconductor industry in China.

Acknowledgements This work was supported by National Outstanding Youth Science Fund Project of National Natural Science Foundation of China (Grant No. 62022062), National Natural Science Foundation of China (Grant Nos. 61974177, 61674119), Fundamental Research Funds for the Central Universities (Grant No. JB210114). The authors would like to thank the experts and researchers who provided the materials for this review.

References

1 Hao Y, Jia X Z, Dong G, et al. Introduction to Microelectronics (in Chinese). 2nd ed. Beijing: Publishing House of Electronics Industry, 2011
2 Hao Y, Zhang J F, Shen B, et al. Progress in group III nitride semiconductor electronic devices. J Semicond, 2012, 33: 081001
3 Tsao J Y, Chowdhury S, Hollis M A, et al. Ultrawide-bandgap semiconductors: research opportunities and challenges. Adv Electron Mater, 2018, 4: 1600501
4 Zhou H, Zhang J C, Zhang C F, et al. A review of the most recent progresses of state-of-art gallium oxide power devices. J Semicond, 2019, 40: 011803
5 Zhang H P, Yuan L, Tang X Y, et al. Progress of ultra-wide bandgap Ga_2O_3 semiconductor materials in power MOSFETs. IEEE Trans Power Electron, 2020, 35: 5157–5179
6 Moore G E. Cramming more components onto integrated circuits. Electronics, 1965, 38: 114–117
7 Moore G E. Progress in digital integrated electronics. In: Proceedings of IEEE Int'l Electron Devices Meeting Technical Digest, 1975. 11–13
8 Dennard R H, Gaensslen F H, Yu H N, et al. Design of ion-implanted MOSFET's with very small physical dimensions. IEEE J Solid-State Circ, 1974, 9: 256–268
9 Salahuddin S, Ni K, Datta S. The era of hyper-scaling in electronics. Nat Electron, 2018, 1: 442–450
10 Shalf J M, Leland R. Computing beyond Moore's law. Computer, 2015, 48: 14–23

11 Arden W, Brillouët M, Cogez P, et al. "More-than-Moore" White Paper. IRTS, 2010
12 Khan H N, Hounshell D A, Fuchs E R H. Science and research policy at the end of Moore's law. Nat Electron, 2018, 1: 14–21
13 Borkar S. Design challenges of technology scaling. IEEE Micro, 1999, 19: 23–29
14 Collaert N. Device architectures for the 5 nm technology node and beyond. 2016. https://bjpcjp.github.io/pdfs/chips/SEMICON_Taiwan_2016_collaert.pdf
15 Jacob A P, Xie R L, Sung M G, et al. Scaling challenges for advanced CMOS Devices. Int J High Speeed Electron Syst, 2017, 26: 1740001
16 Barraud S, Previtali B, Vizioz C, et al. 7-levels-stacked nanosheet GAA transistors for high performance computing. In: Proceedings of IEEE Symposium on VLSI Technology, 2020
17 Veloso A, Eneman G, Huynh-Bao T. Vertical nanowire and nanosheet FETs: device features, novel schemes for improved process control and enhanced mobility, potential for faster & more energy efficient circuits. In: Proceedings of IEEE International Electron Devices Meeting, 2019. 230–233
18 Perrine B. 3D sequential integration. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2019
19 Han G Q, Hao Y. Design technology co-optimization towards sub-3 nm technology nodes. J Semicond, 2021, 42: 020301
20 Borghetti J, Snider G S, Kuekes P J, et al. 'Memristive' switches enable 'stateful' logic operations via material implication. Nature, 2010, 464: 873–876
21 Chi P, Li S C, Xu C, et al. PRIME: a novel processing-in-memory architecture for neural network computation in ReRAM-based main memory. SIGARCH Comput Archit News, 2016, 44: 27–39
22 Manipatruni S, Nikonov D E, Lin C C, et al. Scalable energy-efficient magnetoelectric spin-orbit logic. Nature, 2019, 565: 35–42
23 Zhang W Q, Gao B, Tang J S, et al. Neuro-inspired computing chips. Nat Electron, 2020, 3: 371–382
24 Hickmott T W. Low-frequency negative resistance in thin anodic oxide films. J Appl Phys, 1962, 33: 2669–2682
25 Beck A, Bednorz J G, Gerber C, et al. Reproducible switching effect in thin oxide films for memory applications. Appl Phys Lett, 2000, 77: 139–141
26 Ovshinsky S R. Reversible electrical switching phenomena in disordered structures. Phys Rev Lett, 1968, 21: 1450–1453
27 Wong H S P, Raoux S, Kim S B, et al. Phase change memory. Proc IEEE, 2010, 98: 2201–2227
28 Chappert C, Fert A, van Dau F N. The emergence of spin electronics in data storage. Nat Mater, 2007, 6: 813–823
29 Wang X B, Chen Y R, Xi H W, et al. Spintronic memristor through spin-torque-induced magnetization motion. IEEE Electron Device Lett, 2009, 30: 294–297
30 Rizzo N D, Houssameddine D, Janesky J, et al. A fully functional 64 Mb DDR3 ST-MRAM built on 90 nm CMOS technology. IEEE Trans Magn, 2013, 49: 4441–4446
31 Lee K, Kim W J, Lee J H, et al. 1 Gbit high density embedded STT-MRAM in 28 nm FDSOI technology. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2019
32 Jerry M, Chen P Y, Zhang J C, et al. Ferroelectric FET analog synapse for acceleration of deep neural network training. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2017
33 Tang J, Bishop D, Kim S, et al. ECRAM as scalable synaptic cell for high-speed, low-power neuromorphic computing. In: Proceedings of IEEE International Electron Devices Meeting, 2018
34 Ni K, Yin X Z, Laguna A F, et al. Ferroelectric ternary content-addressable memory for one-shot learning. Nat Electron, 2019, 2: 521–529
35 Li B Z, Gu J J, Jiang W Z. Artificial intelligence (AI) chip technology review. In: Proceedings of International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), 2019. 114–117
36 Akopyan F, Sawada J, Cassidy A, et al. TrueNorth: design and tool flow of a 65 mW 1 million neuron programmable neurosynaptic chip. IEEE Trans Comput-Aided Des Integr Circ Syst, 2015, 34: 1537–1557
37 Davies M, Srinivasa N, Lin T H, et al. Loihi: a neuromorphic manycore processor with on-chip learning. IEEE Micro, 2018, 38: 82–99
38 Naffziger S, Lepak K, Paraschou M, et al. 2.2 AMD chiplet architecture for high-performance server and desktop products. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2020. 44–45
39 You X H, Zhang C, Tan X S, et al. AI for 5G: research directions and paradigms. Sci China Inf Sci, 2019, 62: 021301
40 Ali A, Dinc H, Bhoraskar P, et al. A 12b 18 GS/s RF sampling ADC with an integrated wideband track-and-hold amplifier and background calibration. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2020. 250–252
41 Shibata H, Taylor G, Schell B, et al. An 800 MHz-BW VCO-based continuous-time pipelined ADC with inherent anti-aliasing and on-chip digital reconstruction filter. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2020. 260–262
42 Holt W M. Moore's law: a path going forward. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2016. 8–13
43 Zhu S, Xu B W, Wu B, et al. A skew-free 10 GS/s 6 bit CMOS ADC with compact time-domain signal folding and inherent DEM. IEEE J Solid-State Circ, 2016, 51: 1785–1796
44 Seok E, Cao C H, Shim D, et al. A 410 GHz CMOS push-push oscillator with an on-chip patch antenna. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2008. 472–473
45 Sengupta K, Hajimiri A. A 0.28 THz 4×4 power-generation and beam-steering array. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2012. 256–258
46 Han R, Afshari E. A 260 GHz broadband source with 1.1 mW continuous-wave radiated power and EIRP of 15.7 dBm in 65 nm CMOS. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2013. 138–139
47 Tousi Y, Afshari E. A scalable THz 2D phased array with +17 dBm of EIRP at 338 GHz in 65 nm bulk CMOS. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2014. 258–259
48 Meng X Y, Chi B Y, Wang Z H. CMOS cross-coupled oscillator operating close to the transistor's $f_{\max}$. IEEE Microw Wirel Compon Lett, 2017, 27: 1131–1133
49 Park J D, Kang S, Thyagarajan S V, et al. A 260 GHz fully integrated CMOS transceiver for wireless chip-to-chip communication. In: Proceedings of Symposium on VLSI Circuits (VLSIC), 2012. 48–49
50 Wang Z, Chiang P Y, Nazari P, et al. A 210 GHz fully integrated differential transceiver with fundamental-frequency VCO in 32 nm SOI CMOS. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2013. 136–137
51 Deng X D, Li Y, Li J, et al. A 320-GHz 1×4 fully integrated phased array transmitter using 0.13-μm SiGe BiCMOS technology. IEEE Trans THz Sci Technol, 2015, 5: 930–940

52 Brayton R, Cong J. Electronic Design Automation: Past, Present, and Future. NSF Workshop Report, 2009
53 Brayton R, Cong J. NSF workshop on EDA: past, present, and future (part 1). IEEE Des Test Comput, 2010, 27(2): 68–74
54 Brayton R, Cong J. NSF workshop on EDA: past, present, and future (part 2). IEEE Des Test Comput, 2010, 27(3): 62–74
55 Chen W, Bottoms W R. Heterogeneous integration Roadmap. In: Proceedings of International Conference on Electronics Packaging (ICEP), 2017
56 Hancock T M, Demmin J. Heterogeneous and 3D integration at DARPA. In: Proceedings of IEEE International 3D Systems Integration Conference, 2019. 27–29
57 Gutierrez-Aitken A, Scott D, Sato K, et al. Diverse accessible heterogeneous integration (DAHI) foundry at Northrop Grumman Aerospace Systems (NGAS). ECS Trans, 2017, 80: 125–134
58 Wu L S, Zhao Y, Shen H C, et al. Heterogeneous integration of GaAs pHEMT and Si CMOS on the same chip. Chin Phys B, 2016, 25: 067306
59 Fitzgerald E A, Bulsara M T, Bai Y, et al. Monolithic III-V/Si integration. ECS Trans, 2009, 19: 345–350
60 Lin J J, You T G, Wang M, et al. Efficient ion-slicing of InP thin film for Si-based hetero-integration. Nanotechnology, 2018, 29: 504002
61 Shi H N, Huang K, Mu F W, et al. Realization of wafer-scale single-crystalline GaN film on CMOS-compatible Si(100) substrate by ion-cutting technique. Semicond Sci Technol, 2020, 35: 125004
62 Xu W H, Wang Y B, You T G, et al. First demonstration of waferscale heterogeneous integration of Ga_2O_3 MOSFETs on SiC and Si substrates by ion-cutting process. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2019
63 Yan Y Q, Huang K, Zhou H Y, et al. Wafer-scale fabrication of 42 rotated Y-cut $LiTaO_3$-on-insulator (LTOI) substrate for a SAW resonator. ACS Appl Electron Mater, 2019, 1: 1660–1666
64 Al-Sarawi S F, Abbott D, Franzon P D. A review of 3-D packaging technology. IEEE Trans Comp Packag Manufact Technol B, 1998, 21: 2–14
65 Tummala R R. Packaging: past, present and future. In: Proceedings of the 6th International Conference on Electronic Packaging Technology, 2005
66 Ulrich R K. Advanced Electronic Packaging. 2nd ed. Hoboken: John Wiley & Sons, 2006
67 Gambino J P, Adderly S A, Knickerbocker J U. An overview of through-silicon-via technology and manufacturing challenges. MicroElectron Eng, 2015, 135: 73–106
68 Li T, Hou J, Yan J L, et al. Chiplet heterogeneous integration technology-status and challenges. Electronics, 2020, 9: 670
69 Prezioso M, Merrikh-Bayat F, Hoskins B D, et al. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. Nature, 2015, 521: 61–64
70 Chen W H, Li K X, Lin W Y, et al. A 65 nm 1 Mb nonvolatile computing-in-memory ReRAM macro with sub-16 ns multiply-and-accumulate for binary DNN AI edge processors. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2018. 494–496
71 Liu Q, Gao B, Yao P, et al. A fully integrated analog ReRAM based 78.4TOPS/W compute-in-memory chip with fully parallel MAC computing. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2020. 500–502
72 Yao P, Wu H Q, Gao B, et al. Fully hardware-implemented memristor convolutional neural network. Nature, 2020, 577: 641–646
73 Jiang Z W, Yin S H, Seo J S, et al. XNOR-SRAM in-bitcell computing SRAM macro based on resistive computing mechanism. In: Proceedings of the on Great Lakes Symposium on VLSI, 2019. 417–422
74 Valavi H, Ramadge P J, Nestler E, et al. A 64-tile 2.4-Mb in-memory-computing CNN accelerator employing charge-domain compute. IEEE J Solid-State Circ, 2019, 54: 1789–1799
75 Chih Y D, Lee P H, Fujiwara H, et al. An 89TOPS/W and 16.3TOPS/mm^2 all-digital SRAM-based full-precision compute-in memory macro in 22 nm for machine-learning edge applications. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2021. 252–254
76 Seshadri V, Lee D, Mullins T, et al. Ambit: in-memory accelerator for bulk bitwise operations using commodity DRAM technology. In: Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture, 2017. 273–287
77 Li S C, Niu D M, Malladi K, et al. DRISA: a DRAM-based reconfigurable in-situ accelerator. In: Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture, 2017. 288–301
78 Tulevski G S, Franklin A D, Frank D, et al. Toward high-performance digital logic technology with carbon nanotubes. ACS Nano, 2014, 8: 8730–8745
79 Iijima S. Helical microtubules of graphitic carbon. Nature, 1991, 354: 56–58
80 Dürkop T, Getty S A, Cobas E, et al. Extraordinary mobility in semiconducting carbon nanotubes. Nano Lett, 2004, 4: 35–39
81 Tans S J, Verschueren A R M, Dekker C. Room-temperature transistor based on a single carbon nanotube. Nature, 1998, 393: 49–52
82 Martel R, Schmidt T, Shea H R, et al. Single- and multi-wall carbon nanotube field-effect transistors. Appl Phys Lett, 1998, 73: 2447–2449
83 Javey A, Guo J, Wang Q, et al. Ballistic carbon nanotube field-effect transistors. Nature, 2003, 424: 654–657
84 Chen Z H, Appenzeller J, Lin Y M, et al. An integrated logic circuit assembled on a single carbon nanotube. Science, 2006, 311: 1735–1735
85 Zhang Z Y, Liang X L, Wang S H, et al. Doping-free fabrication of carbon nanotube based ballistic CMOS devices and circuits. Nano Lett, 2007, 7: 3603–3607
86 Zhang Z Y, Wang S, Wang Z X, et al. Almost perfectly symmetric SWCNT-based CMOS devices and scaling. ACS Nano, 2009, 3: 3781–3787
87 Qiu C G, Zhang Z Y, Xiao M M, et al. Scaling carbon nanotube complementary transistors to 5-nm gate lengths. Science, 2017, 355: 271–276
88 Qiu C G, Liu F, Xu L, et al. Dirac-source field-effect transistors as energy-efficient, high-performance electronic switches. Science, 2018, 361: 387–392
89 Franklin A D. Electronics: the road to carbon nanotube transistors. Nature, 2013, 498: 443–444
90 Shulaker M M, Hills G, Patil N, et al. Carbon nanotube computer. Nature, 2013, 501: 526–530
91 Liu L J, Han J, Xu L, et al. Aligned, high-density semiconducting carbon nanotube arrays for high-performance electronics. Science, 2020, 368: 850–856
92 Bishop M D, Hills G, Srimani T, et al. Fabrication of carbon nanotube field-effect transistors in commercial silicon manufacturing facilities. Nat Electron, 2020, 3: 492–501

93 Johnson E O. Physical limitation on frequency and power parameters of transistors. In: Proceedings of IRE International Convention Record, 1991. 295–302
94 Mishra U K, Parikh P, Wu Y F. AlGaN/GaN HEMTs-an overview of device operation and applications. Proc IEEE, 2002, 90: 1022–1031
95 Shen L M, Heikman S, Moran B, et al. AlGaN/AlN/GaN high-power microwave HEMT. IEEE Electron Device Lett, 2001, 22: 457–459
96 Sarazin N, Jardel O, Morvan E, et al. X-band power characterisation of AlInN/AlN/GaN HEMT grown on SiC substrate. Electron Lett, 2007, 43: 1317–1318
97 Chu R M, Shen L, Fichtenbaum N, et al. V-gate GaN HEMTs for X-band power applications. IEEE Electron Device Lett, 2008, 29: 974–976
98 Sun H F, Alt A R, Benedickter H, et al. 102-GHz AlInN/GaN HEMTs on silicon with 2.5-W/mm output power at 10 GHz. IEEE Electron Device Lett, 2009, 30: 796–798
99 Chang C H, Hsu H T, Huang L C, et al. Fabrication of AlGaN/GaN high electron mobility transistors (HEMTs) on silicon substrate with slant field plates using deep-UV lithography featuring 5W/mm power density at X-band. In: Proceedings of Asia Pacific Microwave Conference, 2012. 941–943
100 Wu Y F, Saxler A, Moore M, et al. 30-W/mm GaN HEMTs by field plate optimization. IEEE Electron Device Lett, 2004, 25: 117–119
101 Tilak V, Green B, Kaper V, et al. Influence of barrier thickness on the high-power performance of AlGaN/GaN HEMTs. IEEE Electron Device Lett, 2001, 22: 504–506
102 Shen L, Coffie R, Buttari D, et al. High-power polarization-engineered GaN/AlGaN/GaN HEMTs without surface passivation. IEEE Electron Device Lett, 2004, 25: 7–9
103 Ikeda N, Niiyama Y, Kambayashi H, et al. GaN power transistors on Si substrates for switching applications. Proc IEEE, 2010, 98: 1151–1161
104 Robert R S, Stewart E J, Freitag R, et al. The super-lattice castellated field effect transistor (SLCFET): a novel high performance transistor topology ideal for RF switching. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2014
105 Medjdoub F, Herbecq N, Linge A, et al. High frequency high breakdown voltage GaN transistors. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2015
106 Makiyama K, Ozaki S, Ohki T, et al. Collapse-free high power InAlGaN/GaN-HEMT with 3 W/mm at 96 GHz. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2015
107 Romanczyk B, Guidry M, Wienecke S, et al. W-Band N-Polar GaN MISHEMTs with high power and record 27.8 efficiency at 94 GHz. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2016
108 Then H W, Dasgupta S, Radosavljevic M, et al. 3D heterogeneous integration of high performance high-K metal gate GaN NMOS and Si PMOS transistors on 300 mm high-resistivity Si substrate for energy-efficient and compact power delivery, RF (5G and beyond) and SoC applications. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2019
109 Han W T, Radosavljevic M, Jun K, et al. Advances in research on 300 mm Gallium Nitride-on-Si(111) NMOS transistor and silicon CMOS integration. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2020
110 Dang K, Zhang J C, Zhou H, et al. Lateral GaN Schottky barrier diode for wireless high-power transfer application with high RF/DC conversion efficiency: from circuit construction and device technologies to system demonstration. IEEE Trans Ind Electron, 2020, 67: 6597–6606
111 Dang K, Zhang J C, Zhou H, et al. A 5.8-GHz high-power and high-efficiency rectifier circuit with lateral GaN Schottky diode for wireless power transfer. IEEE Trans Power Electron, 2020, 35: 2247–2252
112 Zhang T, Zhang J C, Zhou H, et al. A 1.9-kV/2.61-mΩ cm^2 lateral GaN Schottky barrier diode on silicon substrate with tungsten anode and low turn-on voltage of 0.35 V. IEEE Electron Device Lett, 2018, 39: 1548–1551
113 Zhang T, Zhang J C, Zhou H, et al. High-performance lateral GaN Schottky barrier diode on silicon substrate with low turn-on voltage of 0.31 V, high breakdown voltage of 2.65 kV and high-power figure-of-merit of 2.65 GW cm^{-2}. Appl Phys Express, 2019, 12: 046502
114 Zhang T, Zhang J C, Xu S, et al. A $>$ 3 kV/2.94 mΩ cm^2 and low leakage current with low turn-on voltage lateral GaN Schottky barrier diode on silicon substrate with anode engineering technique. IEEE Electron Device Lett, 2019, 40: 1583–1586
115 Zhang T, Zhang J C, Zhang W H, et al. Investigation of an AlGaN-channel Schottky barrier diode on a silicon substrate with a molybdenum anode. Semicond Sci Technol, 2021, 36: 044003
116 Fu H, Baranowski I, Huang X, et al. Demonstration of AlN Schottky barrier diodes with blocking voltage over 1 kV. IEEE Electron Device Lett, 2017, 38: 1286–1289
117 Borisov B, Kuryatkov V, Kudryavtsev Y, et al. Si-doped $Al_xGa_{1-x}N$ ($0.56 \leqslant x \leqslant 1$) layers grown by molecular beam epitaxy with ammonia. Appl Phys Lett, 2005, 87: 132106
118 Zhang Y N, Zhang J C, Liu Z H, et al. Demonstration of a 2 kV $Al_{0.85}Ga_{0.15}N$ Schottky barrier diode with improved on-current and ideality factor. IEEE Electron Device Lett, 2020, 41: 457–460
119 Yang C H, Leon R C C, Hwang J C C, et al. Operation of a silicon quantum processor unit cell above one kelvin. Nature, 2020, 580: 350–354
120 Petit L, Eenink H G J, Russ M, et al. Universal quantum logic in hot silicon qubits. Nature, 2020, 580: 355–359
121 Yoneda J, Takeda K, Otsuka T, et al. A quantum-dot spin qubit with coherence limited by charge noise and fidelity higher than 99.9%. Nature Nanotech, 2018, 13: 102–106
122 Huang W, Yang C H, Chan K W, et al. Fidelity benchmarks for two-qubit gates in silicon. Nature, 2019, 569: 532–536
123 Takeda K, Noiri A, Nakajima T. Quantum tomography of an entangled three-spin state in silicon. 2020. ArXiv:2010.10316
124 Yoneda J, Takeda K, Noiri A, et al. Quantum non-demolition readout of an electron spin in silicon. Nat Commun, 2020, 11: 1144
125 Pillarisetty R, George H C, Watson T F, et al. High volume electrical characterization of semiconductor qubits. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2019. 7–11
126 Franceschi S D, Hutin L, Maurand R, et al. SOI technology for quantum information processing. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), 2016. 3–7
127 Guevel L L, Billiot G, Jehl X, et al. A 110 mK 295 µW 28 nm FDSOI CMOS quantum integrated circuit with a 2.8 GHz excitation and nA current sensing of an on-chip double quantum dot. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC), 2020. 306–308

128 Gupta S, Navaraj W T, Lorenzelli L, et al. Ultra-thin chips for high-performance flexible electronics. NPJ Flexible Electron, 2018, 2: 1–17
129 Huang S Y, Liu Y, Zhao Y, et al. Flexible electronics: stretchable electrodes and their future. Adv Funct Mater, 2019, 29: 1805924
130 Son D, Kang J, Vardoulis O, et al. An integrated self-healable electronic skin system fabricated via dynamic reconstruction of a nanostructured conducting network. Nat Nanotech, 2018, 13: 1057–1065
131 Song E, Li J, Won S M, et al. Materials for flexible bioelectronic systems as chronic neural interfaces. Nat Mater, 2020, 19: 590–603
132 Matsuhisa N, Inoue D, Zalar P, et al. Printable elastic conductors by in situ formation of silver nanoparticles from silver flakes. Nat Mater, 2017, 16: 834–840
133 Wang N N, Cheng L, Ge R, et al. Perovskite light-emitting diodes based on solution-processed self-organized multiple quantum wells. Nat Photon, 2016, 10: 699–704
134 Cao Y, Wang N N, Tian H, et al. Perovskite light-emitting diodes based on spontaneously formed submicrometre-scale structures. Nature, 2018, 562: 249–253
135 Gu L, Shi H F, Bian L F, et al. Colour-tunable ultra-long organic phosphorescence of a single-component molecular crystal. Nat Photonics, 2019, 13: 406–411
136 Ren H, Yu S D, Chao L F, et al. Efficient and stable Ruddlesden-Popper perovskite solar cell with tailored interlayer molecular interaction. Nat Photonics, 2020, 14: 154–163
137 Liang C, Gu H, Xia Y, et al. Two-dimensional Ruddlesden-Popper layered perovskite solar cells based on phase-pure thin-films. Nat Energy, 2020, 6: 38–45
138 Bogaerts W, Pérez D, Capmany J, et al. Programmable photonic circuits. Nature, 2020, 586: 207–216
139 Smit M, Williams K, van der Tol J. Past, present, and future of InP-based photonic integration. APL Photonics, 2019, 4: 050901
140 Hoefler G E, Zhou Y, Anagnosti M, et al. Foundry development of system-on-chip InP-based photonic integrated circuits. IEEE J Sel Top Quantum Electron, 2019, 25: 1–17
141 Billah M R, Blaicher M, Hoose T, et al. Hybrid integration of silicon photonics circuits and InP lasers by photonic wire bonding. Optica, 2018, 5: 876–883
142 You J, Luo Y K, Yang J, et al. Hybrid/integrated silicon photonics based on 2D materials in optical communication nanosystems. Laser Photonics Rev, 2020, 14: 2000239
143 Wang C, Zhang M, Chen X, et al. Integrated lithium niobate electro-optic modulators operating at CMOS-compatible voltages. Nature, 2018, 562: 101–104
144 Wang C, Zhang M, Yu M J, et al. Monolithic lithium niobate photonic circuits for Kerr frequency COMB generation and modulation. Nat Commun, 2019, 10: 978
145 Li M X, Ling J W, He Y, et al. Lithium niobate photonic-crystal electro-optic modulator. Nat Commun, 2020, 11: 4123
146 Lin H T, Song Y, Huang Y Z, et al. Chalcogenide glass-on-graphene photonics. Nat Photon, 2017, 11: 798–805
147 Shen W H, Zeng P Y, Yang Z L, et al. Chalcogenide glass photonic integration for improved 2 μm optical interconnection. Photon Res, 2020, 8: 1484–1490
148 Romagnoli M, Sorianello V, Midrio M, et al. Graphene-based integrated photonics for next-generation datacom and telecom. Nat Rev Mater, 2018, 3: 392–414
149 Guo X H, He A, Su Y K. Recent advances of heterogeneously integrated III-V laser on Si. J Semicond, 2019, 40: 101304
150 He A, Guo X, Wang H W, et al. Ultra-compact coupling structures for heterogeneously integrated silicon lasers. J Lightw Technol, 2020, 38: 3974–3982
151 He M B, Xu M Y, Ren Y X, et al. High-performance hybrid silicon and lithium niobate Mach-Zehnder modulators for 100 Gbit s^{-1} and beyond. Nat Photonics, 2019, 13: 359–364
152 Gao A Y, Lai J W, Wang Y J, et al. Observation of ballistic avalanche phenomena in nanoscale vertical InSe/BP heterostructures. Nat Nanotechnol, 2019, 14: 217–222
153 Zhou Z P, Tu Z J, Yin B, et al. Development trends in silicon photonics. Chin Opt Lett, 2013, 11: 012501
154 Zhou Z P, Yin B, Michel J. On-chip light sources for silicon photonics. Light Sci Appl, 2015, 4: 358
155 Chen X, Milosevic M M, Stankovic S, et al. The emergence of silicon photonics as a flexible technology platform. Proc IEEE, 2018, 106: 2101–2116
156 Su Y K, Zhang Y, Qiu C, et al. Silicon photonic platform for passive waveguide devices: materials, fabrication, and applications. Adv Mater Technol, 2020, 5: 1901153
157 Driscoll J B, Doussiere P, Islam S, et al. First 400G 8-channel CWDM silicon photonic integrated transmitter. In: Proceedings of the 15th International Conference on Group IV Photonics (GFP), 2018
158 Fathololoumi S, Nguyen K, Mahalingam H, et al. 1.6 Tbps silicon photonics integrated circuit for co-packaged optical-IO switch applications. In: Proceedings of Optical Fiber Communication Conference, 2020
159 Qiang X G, Zhou X Q, Wang J W, et al. Large-scale silicon quantum photonics implementing arbitrary two-qubit processing. Nat Photon, 2018, 12: 534–539
160 Shen Y C, Harris N C, Skirlo S, et al. Deep learning with coherent nanophotonic circuits. Nat Photonics, 2017, 11: 189–190
161 Poulton C V, Yaacobi A, Cole D B, et al. Coherent solid-state LIDAR with silicon photonic optical phased arrays. Opt Lett, 2017, 42: 4091–4094
162 Wade M, Anderson E, Ardalan S, et al. TeraPHY: a chiplet technology for low-power, high-bandwidth in-package optical I/O. IEEE Micro, 2020, 40: 63–71
163 Capmany J, Novak D. Microwave photonics combines two worlds. Nat Photon, 2007, 1: 319–330
164 Yao J. Microwave photonics. J Lightw Technol, 2009, 27: 314–335
165 Marpaung D, Yao J, Capmany J. Integrated microwave photonics. Nat Photon, 2019, 13: 80–90
166 Shen B Q, Chang L, Liu J Q, et al. Integrated turnkey soliton microcombs. Nature, 2020, 582: 365–369
167 Liu W, Li M, Guzzon R S, et al. A fully reconfigurable photonic integrated signal processor. Nat Photon, 2016, 10: 190–195
168 Grootjans R, Roeloffzen C, Taddei C, et al. Broadband continuously tuneable delay microwave photonic beamformer for phased array antennas. In: Proceedings of the 49th European Microwave Conference (EuMC), 2019. 812–815
169 Hao T F, Tang J, Domenech D, et al. Toward monolithic integration of OEOs: from systems to chips. J Lightw Technol, 2018, 36: 4565–4582
170 Li S M, Cui Z Z, Ye X W, et al. Chip-based microwave-photonic radar for high-resolution imaging. Laser Photonics Rev, 2020, 14: 1900239

171 Zou X H, Bai W L, Chen W, et al. Microwave photonics for featured applications in high-speed railways: communications, detection, and sensing. J Lightw Technol, 2018, 36: 4337–4346
172 Zou X H, Zou F, Cao Z Z, et al. A multifunctional photonic integrated circuit for diverse microwave signal generation, transmission, and processing. Laser Photonics Rev, 2019, 13: 1800240
173 Roy K, Jaiswal A, Panda P. Towards spike-based machine intelligence with neuromorphic computing. Nature, 2019, 575: 607–617
174 Prucnal P R, Shastri B J, de Lima T F, et al. Recent progress in semiconductor excitable lasers for photonic spike processing. Adv Opt Photon, 2016, 8: 228–299
175 Shastri B J, Tait A N, Lima T D, et al. Principles of neuromorphic photonics. 2018. ArXiv:1801.00016
176 Peng H T, Nahmias M A, de Lima T F, et al. Neuromorphic photonic integrated circuits. IEEE J Sel Top Quantum Electron, 2018, 24: 1–15
177 Robertson J, Wade E, Kopp Y, et al. Toward neuromorphic photonic networks of ultrafast spiking laser neurons. IEEE J Sel Top Quantum Electron, 2020, 26: 1–15
178 Xiang S Y, Han Y N, Song Z W, et al. A review: photonics devices, architectures, and algorithms for optical neural computing. J Semicond, 2021, 42: 023105
179 Xiang S Y, Wen A J, Pan W. Emulation of spiking response and spiking frequency property in VCSEL-based photonic neuron. IEEE Photonics J, 2016, 8: 1–9
180 Xiang S Y, Zhang Y H, Guo X X, et al. Photonic generation of neuron-like dynamics using VCSELs subject to double polarized optical injection. J Lightw Technol, 2018, 36: 4227–4234
181 Xiang J L, Torchy A, Guo X, et al. All-optical spiking neuron based on passive microresonator. J Lightw Technol, 2020, 38: 4019–4029
182 Ren Q S, Zhang Y L, Wang R, et al. Optical spike-timing-dependent plasticity with weight-dependent learning window and reward modulation. Opt Express, 2015, 23: 25247
183 Xiang S Y, Han Y N, Guo X X, et al. Real-time optical spike-timing dependent plasticity in a single VCSEL with dual-polarized pulsed optical injection. Sci China Inf Sci, 2020, 63: 160405
184 Zhou H L, Zhao Y H, Xu G X, et al. Chip-scale optical matrix computation for pagerank algorithm. IEEE J Sel Top Quantum Electron, 2020, 26: 1–10
185 Zhou H L, Zhao Y H, Wei Y X, et al. All-in-one silicon photonic polarization processor. Nanophotonics, 2019, 8: 2257–2267
186 Ríos C, Youngblood N, Cheng Z, et al. In-memory computing on a photonic platform. Sci Adv, 2019, 5: 5759
187 Xu S F, Wang J, Wang R, et al. High-accuracy optical convolution unit architecture for convolutional neural networks by cascaded acousto-optical modulator arrays. Opt Express, 2019, 27: 19778
188 Xu S F, Wang J, Zou W W. Optical patching scheme for optical convolutional neural networks based on wavelength-division multiplexing and optical delay lines. Opt Lett, 2020, 45: 3689–3692
189 Vandoorne K, Mechet P, van Vaerenbergh T, et al. Experimental demonstration of reservoir computing on a silicon photonics chip. Nat Commun, 2014, 5: 3541
190 Guo X X, Xiang S Y, Zhang Y H, et al. Polarization multiplexing reservoir computing based on a VCSEL with polarized optical feedback. IEEE J Sel Top Quantum Electron, 2020, 26: 1–9
191 Xiang S Y, Zhang Y H, Gong J K, et al. STDP-based unsupervised spike pattern learning in a photonic spiking neural network with VCSELs and VCSOAs. IEEE J Sel Top Quantum Electron, 2019, 25: 1–9
192 Xiang S Y, Ren Z X, Song Z W, et al. Computing primitive of fully VCSEL-based all-optical spiking neural network for supervised learning and pattern classification. IEEE Trans Neural Netw Learn Syst, 2020. doi: 10.1109/TNNLS.2020.3006263
193 Tait A N, de Lima T F, Zhou E, et al. Neuromorphic photonic networks using silicon photonic weight banks. Sci Rep, 2017, 7: 1–10
194 Lin X, Rivenson Y, Yardimci N T, et al. All-optical machine learning using diffractive deep neural networks. Science, 2018, 361: 1004–1008
195 Feldmann J, Youngblood N, Wright C D, et al. All-optical spiking neurosynaptic networks with self-learning capabilities. Nature, 2019, 569: 208–214
196 Wetzstein G, Ozcan A, Gigan S, et al. Inference in artificial intelligence with deep optics and photonics. Nature, 2020, 588: 39–47
197 Feldmann J, Youngblood N, Karpov M, et al. Parallel convolutional processing using an integrated photonic tensor core. Nature, 2021, 589: 52–58
198 Xu X, Tan M X, Corcoran B, et al. 11 TOPS photonic convolutional accelerator for optical neural networks. Nature, 2021, 589: 44–51
199 Wu H Q, Dai Q H. Artificial intelligence accelerated by light. Nature, 2021, 589: 25–26
200 Morkoc H, Mohammad S N. High-luminosity blue and blue-green gallium nitride light-emitting diodes. Science, 1995, 267: 51–55
201 Khan A, Balakrishnan K, Katona T. Ultraviolet light-emitting diodes based on group three nitrides. Nat Photon, 2008, 2: 77–84
202 Jia Y Q, Ning J, Zhang J C, et al. Transferable GaN enabled by selective nucleation of AlN on graphene for high-brightness violet light-emitting diodes. Adv Opt Mater, 2020, 8: 1901632
203 Peng R S, Meng X J, Xu S R, et al. Study on dislocation annihilation mechanism of the high-quality GaN grown on sputtered AlN/PSS and its application in green light-emitting diodes. IEEE Trans Electron Device, 2019, 66: 2243–2248
204 Liu L, Yang C, Patané A, et al. High-detectivity ultraviolet photodetectors based on laterally mesoporous GaN. Nanoscale, 2017, 9: 8142–8148
205 Li J, Xi X, Li X D, et al. Ultra-high and fast ultraviolet response photodetectors based on lateral porous GaN/Ag nanowires composite nanostructure. Adv Opt Mater, 2020, 8: 1902162
206 Li J, Xi X, Lin S, et al. Ultrahigh sensitivity graphene/nanoporous GaN ultraviolet photodetectors. ACS Appl Mater Interfa, 2020, 12: 11965–11971
207 Noda S, Fujita M. Light-emitting diodes: photonic crystal efficiency boost. Nat Photon, 2009, 3: 129–130
208 Zhang C, Park S H, Chen D, et al. Mesoporous GaN for photonic engineering-highly reflective GaN mirrors as an example. ACS Photonics, 2015, 2: 980–986
209 Lee S M, Gong S H, Kang J H, et al. Optically pumped GaN vertical cavity surface emitting laser with high index-contrast nanoporous distributed Bragg reflector. Opt Express, 2015, 23: 11023–11030
210 Gao X M, Shi Z, Jiang Y, et al. Monolithic III-nitride photonic integration toward multifunctional devices. Opt Lett, 2017,

42: 4853–4856
211 Li K H, Fu W Y, Cheung Y F, et al. Monolithically integrated InGaN/GaN light-emitting diodes, photodetectors, and waveguides on Si substrate. Optica, 2018, 5: 564–569
212 Liu Z J, Huang T D, Ma J, et al. Monolithic integration of AlGaN/GaN HEMT on LED by MOCVD. IEEE Electron Device Lett, 2014, 35: 330–332
213 Liu C, Cai Y F, Zou X B, et al. Low-leakage high-breakdown laterally integrated HEMT-LED via n-GaN electrode. IEEE Photon Tech Lett, 2016, 28: 1130–1133
214 Lu X, Liu C, Jiang H X, et al. Monolithic integration of enhancement-mode vertical driving transistorson a standard InGaN/GaN light emitting diode structure. Appl Phys Lett, 2016, 109: 053504
215 Cai Y F, Gong Y P, Bai J, et al. Controllable uniform green light emitters enabled by circular HEMT-LED devices. IEEE Photon J, 2018, 10: 1–7
216 Tsuchiyama K, Yamane K, Utsunomiya S, et al. Monolithic integration of Si-MOSFET and GaN-LED using Si/SiO_2/GaN-LED wafer. Appl Phys Express, 2016, 9: 104101
217 Gao X M, Yuan J L, Yang Y C, et al. A 30 Mbps in-plane full-duplex light communication using a monolithic GaN photonic circuit. Semicond Sci Technol, 2017, 32: 075002
218 Li K H, Cheung Y F, Fu W Y, et al. Monolithic integration of GaN-on-sapphire light-emitting diodes, photodetectors, and waveguides. IEEE J Sel Top Quantum Electron, 2018, 24: 1–6
219 Chun H, Rajbhandari S, Faulkner G, et al. LED based wavelength division multiplexed 10 Gb/s visible light communications. J Lightw Technol, 2016, 34: 3047–3052
220 Rajbhandari S, McKendry J J D, Herrnsdorf J, et al. A review of gallium nitride LEDs for multi-gigabit-per-second visible light data communications. Semicond Sci Technol, 2017, 32: 023001
221 Zhao L X, Zhu S C, Wu C H, et al. GaN-based LEDs for light communication. Sci China-Phys Mech Astron, 2016, 59: 107301
222 Zhu S C, Lin S, Li J, et al. Influence of quantum confined Stark effect and carrier localization effect on modulation bandwidth for GaN-based LEDs. Appl Phys Lett, 2017, 111: 171105
223 Rashidi A, Monavarian M, Aragon A, et al. Nonpolar m-plane InGaN/GaN micro-scale light-emitting diode with 1.5 GHz modulation bandwidth. IEEE Electron Device Lett, 2018, 39: 520–523
224 Cao H C, Lin S, Ma Z H, et al. Color converted white light-emitting diodes with 637.6 MHz modulation bandwidth. IEEE Electron Device Lett, 2019, 40: 267–270
225 Vahala K J. Optical microcavities. Nature, 2003, 424: 839–846
226 Feng M X, Wang J, Zhou R, et al. On-chip integration of GaN-based laser, modulator, and photodetector grown on Si. IEEE J Sel Top Quantum Electron, 2018, 24: 1–5
227 Tamboli A C, Haberer E D, Sharma R, et al. Room-temperature continuous-wave lasing in GaN/InGaN microdisks. Nat Photon, 2007, 1: 61–64
228 Simeonov D, Feltin E, Bühlmann H J, et al. Blue lasing at room temperature in high quality factor GaN/AlInN microdisks with InGaN quantum wells. Appl Phys Lett, 2007, 90: 061106
229 Tabataba-Vakili F, Doyennette L, Brimont C, et al. Blue microlasers integrated on a photonic platform on silicon. ACS Photonics, 2018, 5: 3643–3648
230 Tabataba-Vakili F, Rennesson S, Damilano B, et al. III-nitride on silicon electrically injected microrings for nanophotonic circuits. Opt Express, 2019, 27: 11800–11808
231 Yang C, Liu L, Zhu S C, et al. GaN with laterally aligned nanopores to enhance the water splitting. J Phys Chem C, 2017, 121: 7331–7336
232 Li J, Yang C, Liu L, et al. High responsivity and wavelength selectivity of GaN-based resonant cavity photodiodes. Adv Opt Mater, 2020, 8: 1901276