Recent Advances and Future Prospects for Memristive Materials, Devices and Systems

Min-Kyu Song[1,2,‡], Ji-Hoon Kang[1,2,‡], Xinyuan Zhang[2,3], Wonjae Ji[4], Alon Ascoli[5], Ioannis Messaris[5], Ahmet Samil Demirkol[5], Bowei Dong[6], Samarth Aggarwal[6], Weier Wan[7], Seok-Man Hong[8], Suma George Cardwell[9], Irem Boybat[10], Jae-sun Seo[11], Jang-Sik Lee[4], Mario Lanza[12], Hanwool Yeon[13], Murat Onen[14], Ju Li[3,15], Bilge Yildiz[3,15], Jesús A. del Alamo[14], Seyoung Kim[4], Shinhyun Choi[8], Gianluca Milano[16], Carlo Ricciardi[17], Lambert Alff[18], Yang Chai[19], Zhongrui Wang[20], Harish Bhaskaran[6], Mark C. Hersam[21,22,23], Dmitri Strukov[24], H.-S. Phillip Wong[7], Ilia Valov[25,26], Bin Gao[27], Huaqiang Wu[27], Ronald Tetzlaff[5], Abu Sebastian[10], Wei Lu[28], Leon Chua[29], J. Joshua Yang[30], Jeehwan Kim[1,2,3]*

[1]Department of Mechanical Engineering, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, 02139, United States.

[2]Research Laboratory of Electronics, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts 02139, United States.

[3]Department of Materials Science and Engineering, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, 02139, United States.

[4]Department of Materials Science and Engineering, Pohang University of Science and Technology (POSTECH), Pohang, 37673, Republic of Korea.

[5]Chair of Fundamentals of Electrical Engineering, Institute of Principles of Electrical and Electronic Engineering, Faculty of Electrical and Computer Engineering, School of Engineering Sciences, Technische Universität Dresden, Dresden, 01069, Germany.

[6]Department of Materials, University of Oxford, Oxford, OX1 3PH, United Kingdom.

[7]Department of Electrical Engineering, Stanford University, Stanford, California, 94305, United States.

[8]The School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, 34141, Republic of Korea.

[9]Sandia National Laboratories, Albuquerque, New Mexico, 87123, United States.

[10]IBM Research Europe, 8803 Rüschlikon, Switzerland.

[11]School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, Arizona, 85281, United States.

[12]Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, Saudi Arabia.

[13]School of Materials Science and Engineering, Gwangju Institute of Science and Technology (GIST), Gwangju, 61005, Republic of Korea.

[14]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, 02139, United States.

[15]Department of Nuclear Science and Engineering, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, 02139, United States.

[16]Advanced Materials Metrology and Life Sciences Division, Istituto Nazionale di Ricerca Metrologica (INRiM), Strada delle Cacce, Torino, 10135, Italy.

[17]Department of Applied Science and Technology, Politecnico di Torino, c.so Duca degli Abruzzi, Torino, I-10129, Italy.

[18]Advanced Thin Film Technology Division, Institute of Materials Science, Technische Universität Darmstadt, Darmstadt, 64287, Germany.

[19]Department of Applied Physics, Hong Kong Polytechnic University, Hong Kong, 999077, China.

[20]Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, China.

[21]Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois, 60208, United States.

[22]Department of Chemistry, Northwestern University, Evanston, Illinois, 60208, United States.

[23]Department of Electrical and Computer Engineering, Northwestern University, Evanston, Illinois, 60208, United States.

[24]Department of Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, California, 93106, United States.

[25]Research Centre Juelich, PGI-7, Wilhelm-Johnen-Str., Juelich 52425, Germany.

[26]Institute of Electrochemistry and Energy Systems “Acad. E. Budewski”, Bulgarain Academy of Sciences, “Acad. G. Bochev 10” str., 1113 Sofia, Bulgaria.

[27]School of Integrated Circuits, Tsinghua University, Tsinghua University, Beijing, 100084, China.

[28]Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, Michigan, 48109, United States.

[29]Department of Electrical Engineering and Computer Sciences, University of California Berkeley, Berkeley, California, 94720, United States.

[30]Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, California, 90089, United States.

‡These authors contributed equally.

Abstract

Memristive technology has been rapidly emerging as a potential alternative to traditional CMOS technology, which is facing fundamental limitations in its development. Since oxide-based resistive switches were demonstrated as memristors in 2008, memristive devices have garnered significant attention due to their bio-mimetic memory properties, which promise to significantly improve power consumption in computing applications. Here, we provide a comprehensive overview of recent advances in memristive technology, including memristive devices, theory, algorithms, architectures, and systems. In addition, we discuss research directions for various applications of memristive technology, including hardware accelerators for artificial intelligence, in-sensor computing, and probabilistic computing. Finally, we provide a forward-looking perspective on the future of memristive technology, outlining the challenges and opportunities for further research and innovation in this field. By providing an up-to-date overview of the state-of-the-art in memristive technology, this review aims to inform and inspire further research in this field.

KEYWORD: Memristor, Compute-in-memory, Resistive switching memory, Ferroelectric memory, Phase change memory, Ion-intercalation resistors, Memtransistors, Neuromorphic computing, In-sensor computing

1. Introduction

Over the past 70 years, Complementary Metal-Oxide Semiconductor (CMOS) technology has advanced rapidly, making it one of the most influential technologies in human history. Exponential scaling of the switching channel in CMOS devices, followed by Moore's law, is one of the key enablers for the advancement.[1] However, CMOS scaling is approaching fundamental physical limits, which has necessitated the need to look for alternative switching devices to continue improving computational performance.[2-4] An ionic-based resistive switching device (*i.e.*, a memristor) is a leading candidate in this context.[5-8] Due to its programmable analog memory effect, which resembles the function of biological synapses in the human brain, memristors have been regarded as a building block of neuromorphic (i.e., brain-like) electronics. This technology has the potential to not only provide biomimetic devices that support human-like data processing but also for influencing even the most basic forms of electronics (Figure 1).

A memristor can be defined as a passive component that remembers the amount of charge that has passed through it.[9, 10] In addition to redox-based resistive switching memory (also known as resistive random-access memory – ReRAM), which was regarded as the memristor, various types of devices, including two-terminal devices (i.e., phase-change memory and magnetic tunnel junctions) and three-terminal devices (i.e., ferroelectric transistors, ion-intercalation resistors, and memtransistors), have been proposed to show a 'memristive' property.[6] Thus, in this review, a memristive device refers to any device whose resistance is programmable under voltage or current inputs and is retained without a power supply. One of the key features of memristive devices is that they can be used as computing units in the form of crossbar arrays. By Ohm's law and Kirchhoff's law, a crossbar array can naturally perform matrix-vector multiplication (MVM) (Figure 2),[11, 12] which is also called compute-in-memory (CIM). While the conventional von

Neumann computing architecture suffers from the high energy costs of data transmission between processing units and memory units, CIM can significantly improve computing power efficiency through massively parallel processing. Since data processing based on artificial neural networks for artificial intelligence (AI) requires numerous MVM, CIM has received great attention for AI accelerator development. Furthermore, memristive devices can be used for implementing alternative computing paradigms, such as spiking neural networks, in-sensor computing, and probabilistic computing.

The purpose of this review is to comprehensively highlight recent advances in memristive technology and to discuss opportunities and challenges for future research. We first discuss recent advances in various types of memristive devices based on material properties before then focusing on studies of memristor theory. We also present the development of memristive algorithms, architectures, and systems in addition to emerging memristive applications and AI accelerators. Finally, forward-looking perspectives on the future prospects of memristive technology are presented.

2. Memristive behaviors of various materials and devices

2.1) Material design of redox-based resistive switching memory

One of the most prominent examples for the important synergy between nanoionics and nanoelectronics are memristive devices. Three different research areas, namely resistive memories,[13, 14] (nano)ionic-based devices[15] and the memristor theory[10, 16] were established in early 1970's and 1980's, that developed in parallel and finally merged in 2008, building the fundament for memristive technologies. In 2008, Strukov *et al*.[9] have proposed resistive switching memories as the missing 4th circuit element – the memristor – as theoretically proposed by L. Chua in 1970, thus initiating the fusion of these three research fields. The modern developments in memristive technologies cover a variety of applications beyond memories such as sensors, nanoactuators, selector devices, memristive transistors, artificial neurons and synapsis for brain inspired computing. A simple classification, based on the physical processes responsible for the resistive switching is presented in Figure 3.

Among the most important categories of resistive switches are the electrochemical metallization cells (ECM, also known as conductive bridge RAM – CBRAM or programmable metallization cells),[17] valence change mechanism memories (VCM, also known as metal oxide resistive memory – OxRAM)[18-20] and thermochemical mechanism (TCM)[21] devices. Recent review papers focus on details on their basic operation principle, memory and computing related applications.[5] All types of resistive switches share a simple two-terminal MIM-structure where the resistance of the insulator can be tuned between at least two different resistance levels (a high-resistive state, HRS, and a low resistive state, LRS). Many devices allow intermediate resistance levels (multilevel-switching) or analogue-switching, resulting in storing multiple bits in a single memory cell. A schematic presentation of filamentary switching ECM and VCM devices is shown in Figure 4.

Redox reactions as nanoionic signature of the resistance transition have been widely accepted.[22, 23]

Adapting memristive devices for different applications and controlling the functionalities is a challenging but essential task of the current research in this field. In many cases the function can be induced by using different pulse scheme or different amplitude and/or magnitude of the external stimuli. However, a more important but in same time more challenging task is using an approach based on materials design. This approach requires a deep understanding of the relation between materials properties, the physicochemical interactions and processes within the device, and the resulting charge dynamics and functionalities. Despite the apparent simplicity in structure and materials, memristive cells are complex nanoscale systems where mechanical, chemical, and electrochemical interactions are present, as shown in Figure 5.

The main electrochemical processes are redox reactions at the interfaces and ionic motion (diffusion and/or migration) within the switching film. However, chemical interactions at the interfaces, and with the molecules from local environment often play important role in determining factors such as variability, state stability endurance and retention. For example, it was recently found that the capping layer is significantly influencing the device performance, despite it is not directly participating into the switching process.[24] Moreover, not only the assemble of materials is important but as well their thicknesses should be adapted and coordinated. In macroscopic systems electrolytes are considered as infinite source of ions and adding or removing some (e.g., due to reactions) is not changing the properties of the system. In nanoscale electrochemical cells this is typically not the case. Adding or removing charges (ions and/or electrons) can have significant impact on the physicochemical properties and thus, on the switching behaviour and functionalities. Thus, thicknesses of the different layers in one device should be carefully selected. In addition,

extracting or adding ions, can change the kinetics from field accelerated to classical diffusion driven transport.[25]

Speaking on materials design an essential factor is the purity of the used materials and often impurities may unexpectedly play the role of doping. In many cases, the level of impurities is not considered where in the same time levels of ppm can significantly change the transport, switching kinetics and functionalities.[25] Such impurities can be immobile, being part of the initial material, but they can be also introduced on a later stage due to incorporation of protons (for example introduced during preparation or incorporated from local environment) or ions from the electrochemically active electrodes.[25, 26] Impurities and doping are important not only considering the switching film, but as well electrodes. It has been also demonstrated that introducing other components in the metallic films (alloying) can be essential for reaching optimal performance.[27]

Specifically, the complex defect structure of oxide functional layers that lead to the formation of the conducting filament has been recently studied. While the important role of oxygen vacancies has been discussed extensively in the literature including quantization effects, [28-31] the interplay of (oxygen) point defects with two dimensional defects like grain and phase boundaries has been mostly neglected. In molecular beam epitaxy (MBE) grown HfO_2 dielectric layers deposited onto TiN bottom electrodes, it is possible to enforce the growth of threading grain boundaries that all have an equivalent crystallographic orientation.[32] Using such model structures, the influence of grain boundaries could be directly studied. Surprisingly, different sets of grain boundaries lead to markedly different forming voltages, V_f. Low-symmetry grain boundaries are connected to large V_f with a broad probability distribution, while high-symmetry grain boundaries show forming free behavior with a sharp distribution of values. (Figure 6) It turns out, that the symmetric grain boundary attracts to larger extent oxygen vacancies which, in turn, lead to a higher concentration

of electronic states in the band gap of HfO_2 close to the Fermi level. Therefore, this particular set of grain boundaries is an ideal precursor for the formation of a conductive filament.[33] The understanding and control of complex defects and their interaction is a key to manipulate conductive filament towards multiple resistive states. While HfO_x is one of the mostly used materials in VCM, other materials like Y_2O_3 with a high amount of intrinsic oxygen vacancies might be even more suited to control the transition between a large number of resistive states.[34, 35]

Two-dimensional (2D) layered materials have also been introduced as switching medium in memristors.[36-39] The best resistive switching performance for memory and neuromorphic applications has been obtained in multilayer hexagonal boron nitride (h-BN) produced using chemical vapor deposition (CVD).[40-42] The reasons are: i) h-BN is an insulator with a band gap ~6 eV, meaning that it can block current in HRS and reduce energy consumption[40]; and ii) CVD provides the right amounts of native defects that enable stable RS.[43, 44] A recent study reported crossbar arrays of small (<0.053 μm^2) h-BN memristors with high endurance (>5×10^6 cycles), and low energy consumption per state transition (~1.41 pJ).[42] It has been studied that the switching occurs by metal penetration from the electrodes[40, 41] as in ECM. Native defects in CVD h-BN, as in VCM, also facilitate the formation of the filament (i.e., lower the energy-to-breakdown) and enable resistive switching.[41] The difference is that in CVD h-BN the native defects are surrounded by extremely stable 2D layered material.[41] This limits the number of atoms that are involved in the switching and avoids lateral propagation of the filament, which should have positive effects to reduce variability and increase endurance.[44, 45] The fact that the resistive switching in CVD h-BN is oxygen-free,[41] should also prevent filament self-disruption, which should enhance the retention time. The problem of most studies in the field of 2D material-based memristors is that they report the performance of few (<5) large devices (>1 μm^2), while the community should be focusing on

presenting data collected in multiple (>100) small devices (<0.1μm^2), if possible arranged in a crossbar structure.[46, 47]

Despite the impressive progress made in the field, materials design of memristive devices has not reached its full potential and fundamental studies on the relation between materials, processes and functionalities are essential for the further progress in tuning and controlling memristive functionalities.

2.2) Optical memory and photonic tensor core based on phase change materials

Phase change materials (PCM) have long been employed as storage medium for optical data storage such as compact disks since the 1980s due to their pronounced change in dielectric function upon phase switching. Some families of PCM include chalcogenide glasses like GeSbTe and certain metal oxides like VO_2. The amorphous phase and crystalline phases of PCM exhibit long-range disorder and long-range order respectively.[48] To switch from the crystalline state to amorphous state, a heat pulse is applied to PCM to elevate the temperature above melting temperature, followed by a rapid quenching (cooling rate higher than 10^9 K/s) to cool down the material. To crystallize PCM, a heat pulse with intermediate power is applied to heat up the material above the glass transition temperature, but not necessarily higher than the melting temperature. The prominent properties of PCM data storage include long data retention at room temperature (more than hundreds of years),[48-50] low switching energy (hundreds of pico-Joule to few nano-Joule), fast switching dynamics at nanosecond scale, and many distinguishable stable intermediate states (tens of states).

Microscale and nanoscale integrated optical memory is a growing field in the past decade. As an emerging field, it is not surprising that this field creates several different names for different devices depending on their properties and functionalities. Here, these emerging devices are

categorized into three groups: optical memristor, optomemristor, and electro-optic memristor (Figure 7a). Optical memristors are optical devices that can store multi-level data in the non-volatile physical states of materials similar to electrical memristors. But optical memristors read out data from the optical domain, requiring different physical states to have different optical properties. Optomemristors are essentially electrical memristors whose switching dynamics and electrical response can be controlled optically. Electro-optic memristors are simultaneously optical and electrical data storage devices. This property is called 'dual electrical-optical functionality', meaning electro-optic memristors can be programmed optically and readout electrically and vice versa.

For PCM optical memristor, data is stored either in the absorption coefficient or the refractive index of PCM. In the first type, the amplitude of incoming light is modulated upon phase switching in conventional PCMs like $Ge_2Sb_2Te_5$ (GST) and AgInSbTe (AIST). In the second type, the phase of incoming light is modulated in low-loss PCMs like Sb_2S_3[51] and Sb_2Se_3.[52] The change in absorption coefficient and refractive index of PCM upon phase switching can exceed unity, which is several orders higher than other mechanisms for non-volatile optical memories. A simple PCM optical memristor can be constructed with thin film deposition of PCM on waveguide for interaction with the evanescent field of waveguide mode.[53, 54] In the case of devices based on GST, the amorphous phase of GST has a lower extinction coefficient, resulting in weak attenuation of light. For the crystalline phase, a higher extinction coefficient leads to strong attenuation of light. Therefore, by switching between amorphous and crystalline states amplitude modulation can be achieved. Further, by varying the fraction of amorphous to crystalline phase, multiple intermediate transmission states or multiple memory levels can be achieved.[53, 55] All-optical programming is adopted in PCM optical memristors – this effectively is a direct photonic analog of electrical

memristors, modulating transmission of light via absorption. The optical programming pulse is absorbed by the PCM thin film resulting in heating of the thin film. By varying the power and shape of such programming pulses, both amorphization and recrystallization can be achieved. Further by varying the power of programming pulses, multiple memory levels can be achieved. To date, a maximum of 5 bits in PCM optical memristor has been reported.[55] Using the all-optical programming approach to reconfigure integrated photonic circuits with PCM memory devices, in-memory matrix computation using photonic synapses and all-optical synaptic neural network have been demonstrated.[56-59] Besides the all-optical programming approach, there has been an increasing interest in electrically controlled programming approach due to the rise in popularity of low-loss PCM for phase modulation. This has until recently been dominated by microheater-based systems. Unlike all optical approaches, microheater-based approaches are power-hungry due to large-area heating and efforts have been made to reduce the programming energy by optimizing microheater designs.[60-62]

Recently, the photonic tensor core for parallelized in-memory MVM is a significant advance in PCM photonics.[63, 64] As shown in Figure 7b, by designing an all-optical integrated photonic circuit with spatially distributed PCM photonic memories and layout architecture analogous to electronic crossbar arrays, the photonic tensor core is able to perform parallel MVM using wavelength division multiplexing. The huge bandwidth provided by wavelength division multiplexing and fast photonic processing renders the photonic tensor core with ultra-high computation speed (more than one trillion operations per second) and density (more than one trillion operations per mm^2) for convolutional processing, which forms the backbone of modern artificial intelligence algorithms. An electro-optic version of photonic tensor core is also developed to reduce the complexity of programming PCM memories.[63] Other major advances in PCM photonics include

optomemristor and electro-optic memristors, which links and synergizes the optical and electrical properties of PCM. PCM optomemristor shows memristor functionalities that cannot be realized solely in the electrical domain, but requires control from the optical domain. A PCM optomemristor formed by Pt/$GeSe_3$/Ag stack[65] was recently reported with optoelectronic properties (Figure 7c). This device is essentially a redox type memristor where a conductive channel is formed in the $GeSe_3$ between the two electrodes, but such resistive switching can be optically controlled. Optical illumination causes the shift of resistive switching voltage towards a smaller value. Hence, the device can be optically controlled to switch at a lower or higher voltage relative to its intrinsic switching voltage. Using this property presented in the Pt/$GeSe_3$/Ag optomemristor, i.e., resistive switching controlled by illumination, the emulation of multi-factor neuromorphic computation is demonstrated as an example of reinforcement learning in hardware. Towards electro-optic memristors featuring dual electrical-optical functionality, meaning the device can be programmed optically and readout electrically and vice versa, a plasmonic nanogap enhanced PCM device has been demonstrated.[66, 67] As shown in Figure 7d, a plasmonic nanogap is integrated with a tapered dielectric waveguide to achieve efficient dielectric waveguide mode to plasmonic mode conversion. The PCM sits at the center of the nanogap to bridge the two tapered electrodes and form a conductive path. Since PCMs are both electrically and optically active, this device can be switched and read out flexibly in both the electrical and optical domain, offering another dimension for on-chip addressing and control of PCM memories.

Moving forward to practical applications, the challenges of PCM optical memory devices include limited endurance due to phase segregation, long-term stability in elevated temperature, ultrafast programming, and integration with CMOS electronics. Phase segregation can be eliminated by using single-element PCM such as Te[68] or Sb.[69-71] However, the data retention time

in Te or Sb -based optical memory should be improved to achieve non-volatility. Long-term stability is limited by the typically low crystallization temperature of PCM, which can be enhanced by doping PCMs with high melting point elements such as Ge.[72] Achieving ultrafast programming speeds up to GHz and integration with CMOS electronics remain some of the most important daunting challenges in PCM material discovery.

2.3) Memristive behavior of ferroelectric transistors and arrays based on hafnia

Various types of materials such as perovskites, 2D materials, polymers, and fluorite oxides have been found to have ferroelectricity and studied for next-generation memory devices.[73-77] The memory effect of the ferroelectric materials such as is attributed to the switching of electric dipole alignments between an upward and downward direction, driven by electric field. Ferroelectric memories include ferroelectric random-access memory (FeRAM), ferroelectric tunnel junction (FTJ), and ferroelectric transistors. Similar to DRAM, FeRAM has one transistor-one capacitor (1T-1C) structure and uses a ferroelectric capacitor to store charges, but the memory states determined by polarization switching of ferroelectric layer is non-volatile. However, destructive readout process and large foot-print are critical drawbacks of FeRAMs. In FTJs, ultrathin ferroelectric materials are sandwiched with two electrodes, which act as tunneling barrier. Polarization switching induces resistive switching by changing barrier height for tunneling. In ferroelectric transistors, the ferroelectric layer is used as a gate-insulating layer.[78] Polarization charge in the ferroelectric layer modulates the channel conductivity resulting in changing threshold voltage (V_t) of the transistor.

Recently, memristive devices using ferroelectricity and their array structure have been actively studied to achieve high-performance, high-density neural networks.[79-82] Using partial polarization characteristics of ferroelectric gate insulators, precise control of carriers inside the channel layer

can be achieved, which can lead to conductance modulation in the channel layer.[82-86] Among diverse ferroelectric materials, hafnia-based ferroelectric materials have been widely investigated for neuromorphic applications due to their CMOS compatibility, fast switching speeds, and high scalability.[87-91] At the erased state (i.e., downward polarization), the electrons inside the channel are depleted, which leads to low channel conductance (Figure 8a).[79, 82] When a programming pulse with increasing amplitude is applied to the gate, the direction of polarization gradually changes to an upward direction, and electrons are accumulated at the interface between the ferroelectric layer and the channel layer. This gradual switching of polarization under incremental pulses results in a gradual increase in channel conductance.[82, 86] Thus, voltage pulses with incremental pulse amplitude or width are usually used to achieve linear and symmetric conductance change in ferroelectric transistors.[82, 92] With incremental pulse schemes, classification accuracy as high as 91% is expected for handwritten digits, which is similar to that of ideal synapses.[82] Although the incremental voltage scheme can result in highly linear and symmetric conductance modulation, the conductance of the ferroelectric transistor should be determined prior to the conductance modulation. Thus, incremental pulse schemes lead to longer training time and higher energy consumption than identical pulse schemes.[86] However, when identical pulse schemes are used for ferroelectric transistors, abrupt conductance changes are reported (Figure 8b, c).[86] Also, it was confirmed that the accuracy of the neural network is significantly degraded with identical pulse schemes compared to that with incremental pulse schemes.[86] (Figure 8d) Thus, it is a challenge to develop a ferroelectric transistor compatible with an identical pulse scheme or operation method, which can decrease the training time and energy consumption.

The array structure for ferroelectric synaptic transistors and its operation method, which can decrease the training time, was demonstrated by Kim *et al*.[91] Synaptic transistor array was

fabricated using low temperature (400 °C) process by integrating a ferroelectric thin film transistor with an oxide semiconductor channel, $InZnO_x$, and ferroelectric gate insulator, $HfZrO_x$ (Figure 8e). In this work, the row- and column-wise parallel programming method, in which the ferroelectric transistor in the same row and column can be programmed simultaneously, was proposed to decrease the training time of ferroelectric transistor arrays. Using the optimized parallel programming method, the conductance of the selected cell can be selectively and linearly modulated (Figure 8f). Also, the conductance of unselected cells that share the same row and column with the selected cells has remained in their states without change. Based on the electrical characteristics of the ferroelectric synaptic transistor and its array, the performance of convolutional neural networks with VGG-8 structure was demonstrated using the simulation method.[93] In a simulation based on the electrical characteristics of the ferroelectric synaptic transistor array, convolutional neural networks achieved an image accuracy of 90.3% on the complex image dataset, CIFAR-10 (Figure 8g). This report experimentally demonstrated the neural network based on ferroelectric transistors and their convolution operations. Also, a row- and column-wise parallel programming method, which could decrease the training time, was experimentally proposed. At present, the demonstration of ferroelectric transistors for neuromorphic applications is usually done with small-density arrays or simulation methods without experimental validation of highly-scaled, high-density neural networks based on ferroelectric transistors.[94-96] Further study needs to be done on the demonstration of a highly scaled transistor array for neuromorphic applications with considerations for array operations and their characteristics.

2.4) Ion-intercalation programmable resistors

Recently, a device family of non-volatile three-terminal programmable resistors for deep learning applications has emerged.[97] This device class, often referred to as ECRAM (Electrochemical Random-Access Memory), ENODe (Electrochemical Neuromorphic Device) or EIS (Electrochemical Ionic Synapse), relies on controlled intercalation of dopant ions in a semiconductor channel.[98-111] (Figure 9a) In essence, ions are shuttled back and forth between an ion reservoir (also performing the role of gate) and a channel where the ions behave as dopants in response to the application of positive or negative voltage pulses to the reservoir with respect to the channel. In this way, the conductivity of the channel can be increased or decreased in a controlled fashion. During device programming, the ions are transferred through an electrolyte that separates the gate from the channel while the corresponding electrons flow through the outside circuit. When the gate/reservoir is electrically left open, electrons cannot flow and the ions remain in place in the channel. Hence, the device exhibits non-volatile characteristics.

Ion-intercalation programmable resistors have been demonstrated using Li^+,[98-101] O^{2-},[102-104] and H^+,[106-111] among other ions. The use of H^+ (protons) is particularly attractive because their small radius and light mass promises high speed and energy-efficient operation with minimum volume expansion and contraction. Also, unlike most of other ions, protons are CMOS compatible, a key consideration as deep learning accelerators will need to be constructed on a Si CMOS platform. This has implications for the nature of the ion, but also the choice of gate, channel and electrolyte materials, as well as the processing temperatures involved in device fabrication. Early proton-based device demonstrations were based on polymeric materials for the channel, reservoir and electrolyte.[105, 106] Organic materials are not CMOS compatible in that they cannot withstand the processing temperatures generally involved in Back-End-Of-the-Line (BEOL) CMOS fabrication. Later demonstrations used Pd as an inorganic reservoir/gate and a metal oxide (WO_3) as an

inorganic channel but still used polymeric Nafion as the electrolyte.[107] Proton-based devices have also been demonstrated using inorganic 2D electrolytes.[108]

Recently, the use of phosphosilicate glass (PSG) as electrolyte in combination with a Pd reservoir/gate and a WO_3 channel has been demonstrated as the key innovation to resolve the critical limitation of protonic devices: the absence of a CMOS-compatible, solid-state material system.[109-111] PSG is a most suitable choice for this application. It is a well-known material in Si technology with straightforward BEOL integration capabilities; it is an excellent electrical insulator, and it also exhibits good proton conductivity at room temperature.[112] In fact, PSG has been used as a proton-exchange membrane in microfuel cells.[113] Micron-scale protonic resistors[109] and nanoscale active area devices[110] (Figure 9b, c) exhibit high operation speed (5 ns/pulse), high energy efficiency (~fJ/pulse), many (~1000) nonvolatile conductance states centered around 0.1 μS and spanning a large dynamic range (10×), nearly linear and symmetric modulation for incremental and decremental conductance changes, good retention and high endurance (Figure 9d, e).

A detailed study of the dynamics of these devices has also been carried out by Onen *et al*.[111] They have revealed a device response that includes a transient field-effect induced channel current that results from proton movement through the PSG during the gate voltage pulse. In effect, this device behaves as a MOSFET with a threshold voltage that shifts negative when the PSG is flooded with protons. In addition, when the magnitude of the gate voltage pulse exceeds a certain minimum, the expected non-volatile proton intercalation induced channel current was observed. High-speed measurements reveal that in response to 5 ns gate voltage pulses, this non-volatile conductance modulation occurs in an impulse-like fashion, without any extended equilibration period. (Figure 9f)

The non-volatile programmable protonic ion intercalation resistors described here display state-of-the-art combined material, processing, and performance properties of potential for analog deep learning applications. The operating gate voltage should be improved by engineering the PSG and the Pd/PSG and PSG/WO_3 interfaces.

2.5) Other emerging memristive devices

The development of memristive materials and devices including Spin-transfer Torque Magneto-resistive RAM (STT-MRAM), memtransistor, and gate-injection device can offer possibilities to improve memristive device performance and neuromorphic systems.[114-116] Magnetic device has been studied for several decades, and various operation mechanisms have been reported.[117-119] After predicting the Spin-transfer Torque (STT) effect without the external magnetic field by Slonczewski[120] and Berger[121] in 1996, STT-MRAM with magnetic tunneling junction (MTJ) has become an important memristive device in neuromorphic applications, exhibiting small device-to-device variation, fast switching speed and excellent endurance.[122-125] However, despite the superior device properties of STT-MRAM, the low resistance of STT-MRAM can lead to large power consumption in a conventional crossbar array for analog multiply-accumulate operations, which are essential operation for neuromorphic systems. To overcome this issue, Jung *et al.* demonstrated a 64 × 64 crossbar array based on STT-MRAM cells with an architecture that includes the resistance of a field-effect transistor (FET), as every MTJ is accompanied by a FET switch.[126] Additionally, a multi-state STT-MRAM design also have been proposed for neuromorphic computing schemes as artificial synapses.[127]

The defining structural characteristic of a 2D material is its nanoscale thickness. For 2D semiconductors, this nanoscale thickness has multiple implications for electronic properties including quantum confinement effects that modulate band structure, reduced dielectric screening

that results in enhanced electrostatic modulation by applied gate electric fields, and charge transport that is strongly affected by structural defects.[128] In the specific case of 2D transition metal dichalcogenides (e.g., MoS_2), nanoscale thickness also implies reduced energy barriers for point defect motion, especially in polycrystalline materials where point defect motion is further enhanced along grain boundaries.[129, 130] These properties of 2D transition metal dichalcogenides present opportunities for memristive materials and devices.[131] For example, the low energy barrier for point defect motion in monolayer polycrystalline MoS_2 implies that the spatial doping profile and thus the charge transport properties can be reconfigured under the application of modest lateral electric fields. In this manner, memristive phenemona have been observed in monolayer polycrystalline MoS_2 devices.[132] By concurrently applying a vertical electric field across a gate dielectric, the memristive response can be electrostatically modulated with the details depending on the grain boundary orientation in the device channel.[132] (Figure 10a-c) Since these devices combine the nonvolatile two-terminal response of a memristor with the volatile gate modulation of a transistor, they have been dubbed memtransistors.[133]

Although early work focused on devices with a small number of grain boundaries and thus were sensitive to stochastic variations in grain boundary orientation from device to device,[132] later work recognized that the incorporation of several grain boundaries per device leads to statistical averaging at the single device level, which substantially improves memtransistor yield and wafer-scale uniformity.[133] Efforts to control polycrystalline grain structure and point defect concentration[134, 135] have also proven successful at improving memtransistor device metrics (e.g., smaller device dimensions, lower voltage operation, and reduced power consumption) in addition to enabling neuromorphic functionality such as gate-tunable potentiation and depression (Figure 10d)[136] and heterosynaptic responses in multi-terminal memtransistors.[133] (Figure 10e) The 2D

memtransistor channel can also be gated from above and below, which enables dual-gate memtransistors (Figure 10f) that have been incorporated into crossbar array architectures with minimal crosstalk, disturbance, or sneak currents.[137] (Figure 10g) Since 2D memtransistors are still in their infancy, many research opportunities remain at the materials level (e.g., exploring different 2D semiconductors, dopants, contacts, and gate dielectrics) that are likely to further diversify the device and circuit possibilities both in conventional neuromorphic computing[138] and emerging bio-realistic systems.[139]

A gate injection-based filed-effect transistor (GIFET) has been demonstrated by employing thermionic emission with superior synaptic characteristics (Figure 11a).[140] To achieve the linearity, the authors suggest an operation method to program and erase the charges in the stored layer by thermionic emission from the gate metal (Figure 11b). The barrier height between the charge store layer (CSL, WO_x) and blocking layer (a-Si:H) should be low to guarantee sufficient current density for a high on/off ratio and to hold electrons. The amount of charge stored in the CSL layer changes the Si channel's depletion region by field-effect, and this is the major mechanism to control synapse weight in GIFET. The LTP-LTD characteristic of the GIFET has high linearity with low asymmetric ratio, and the device also achieves robust endurance ($\geq 2\times10^8$ pulses) with consecutive potentiation and depression pulses and retention properties with a data loss of 5.45 % (13.6 nS) (Figure 11c). GIFET has a similar structure to the currently commercialized CTF (charge trap flash), but the engineering of the gate stack and the changing of the charge injection/extraction mechanism allow GIFET to be a promising memristive device. Based on thermionic emission mechanism, GIFET achieves excellent synaptic properties such as number of conductance state, on/off ratio, spatio-temporal variation, linearity, retention and endurance. Moreover, all the processes and materials for GIFET fabrication were CMOS compatible, which means it could be

used as a near-future neuromorphic device. Further development to reduce programming voltage, integrate into large array, and scale down will provide opportunities for the use of GIFET in neuromorphic hardware.

3. Memristor theory

3.1) Edge of Chaos-Induced Bifurcations in Bio-Mimetic Locally-Active Circuits

Local Activity[141] refers to the capability of a physical system to act as a source of local energy upon suitable polarization. Importantly, Local Activity is a fundamental prerequisite for a system to display complex dynamical behaviors,[142] in certain occasions. Edge of Chaos[141] denotes a particularly-attractive condition, in which a physical system is both locally active and asymptotically stable at some operating point. It is when a system is poised on the Edge of Chaos that emergent phenomena may appear across its physical medium out of some bifurcation.[143] Recurring to the universal Theory of Local Activity and Edge of Chaos is necessary to explain the mechanisms behind the appearance of complex phenomena in any physical system, which has the possibility to exchange energy with the respective environment.[144] The recent availability of memristor nanoscale physical realizations,[145, 146] blessed with the capability to feature a Negative Differential Resistance (NDR),[147] which is a signature for Local Activity, enables the design of circuits, which, once some of their constitutive units are poised on the Edge of Chaos, may reproduce complex bifurcation phenomena, occurring in biological systems, while utilizing a lower number of degrees of freedom than their original counterparts. Two major works provide proof of evidence for this claim.

In particular, Ascoli *et al*. [148] presented the simplest ever-reported bio-inspired physical system, which reproduces the same counterintuitive phenomenon, reported by the American luminary

Stephen Smale in 1974,[149] when, while investigating a model from cellular biology, he witnessed two identical fourth-order reaction cells, "mathematically dead" on their own, pulsing together indefinitely, upon diffusive coupling. The proposed bio-inspired reaction-diffusion network, consisting of two identical and resistively-coupled second-order memristive Pearson-Anson cells, features half the number of states than the corresponding biological system (the mathematical model in ref 150 also reproduces the Smale Paradox from ref 149 through a lower number of degrees of freedom, specifically 6, than the biological counterpart, but it requires two additional state variables as compared to the proposed two-cell array). As shown in Figure 12a, the neural network contains only 9 circuit elements, specifically 2 batteries, 3 linear resistors, 2 linear capacitors, and 2 volatile memristors, manufactured with niobium oxide at the facilities of NaMLab (see ref 146 for details on the physical stack of each of the resistance switching memories, which are referred to as NaMLab memristors, in the remainder of this chapter). Employing powerful tools from the theory of Local Activity, and applying methods from Nonlinear Dynamics to an accurate model of the memristor cellular unit, a comprehensive picture for its local and global behavior was drawn. This study, providing a systematic technique to choose the design parameters of the two-cell array, so as to allow diffusion-driven instabilities to appear therein, further enabled to identify the true origin for Smale's paradoxical observations: the appearance of symmetry-breaking effects, accompanied by the steady-state development of sustained oscillations, i.e. dynamic patterns, across the homogeneous cellular medium (refer to Figure 12b, c) is in fact possible if and only if the two identical reaction cells of the medium are preliminarily poised on a common Edge of Chaos (i.e. locally-active and asymptotically-stable) operating point, before the insertion of a dissipative path, joining the top terminals of their capacitors, let them interact diffusively. More specifically, upon a progressive decrease in the coupling resistance – refer to the

circuit implementation of Figure 12a – the Reaction-Diffusion Memristor Cellular Nonlinear Network (RD M-CNN) first undergoes a supercritical pitchfork bifurcation,[151] which signals the destabilization of the homogeneous solution, with the simultaneous emergence of either of two possible locally-stable static inhomogeneous solutions, depending upon the initial condition, and then a supercritical Hopf bifurcation,[152] which endows the proposed bio-inspired cellular array with the capability to support either of two locally-stable dynamic inhomogeneous solutions, as established by the starting conditions for the cells' states, at the expenses of the static Turing patterns, which concurrently lose stability (consult ref 153 for more details).

In his seminal paper[154] Alan Turing, the Father of Artificial Intelligence, devised the model of a linear dynamical system, consisting of two identical diffusively-coupled second-order reaction cells, which is subject to dissipation-driven instabilities. Due to the lack of a nonlinear saturation mechanism, the destabilization of the homogeneous solution of the array resulted here in the asymptotic divergence of its four state variables. Including a nonlinear term in a second-order Turing-like reaction cell (see ref 155, for example), the diffusion-driven symmetry-breaking phenomenon in the resulting fourth-order dynamical system may lead to the steady-state appearance of static inhomogeneous solutions. Ascoli *et al*. presented the simplest ever-reported two-cell neural network,[156] which, leveraging the NDR of the NaMLab memristor, is capable to support dissipation-induced static pattern formation through half the number of states than Turing-like reaction-diffusion arrays. As shown in Figure 13a, the proposed neural network is composed of just 7 one-ports, specifically 2 DC voltage sources, 3 linear resistors, and 2 niobium oxide-based volatile memristors from NaMLab.[146] In this work, they showed that complexity may emerge in a physical circuit, hosting no other dynamic component besides the 2 locally-active memristors.[145] As reported in ref 156, an in-depth circuit- and system-theoretic analysis allowed Ascoli *et al*. to

gain a thorough understanding of the local and global dynamics of the bio-inspired reaction-diffusion network of Figure 13a, to prove that the preliminary polarization of each of its two identical constitutive units, when isolated one from the other, on some common Edge of Chaos operating point is the *conditio sine qua non* for the destabilization of the homogeneous solution of the array, later on, when a diffusion process establishes their interaction, and to enable the identification of the local supercritical pitchfork bifurcation,[151] which spawns the birth of two locally-stable inhomogeneous static solutions across the respective cellular medium under a sufficiently strong coupling condition (see Figure 13b,c).

All in all, the origin for the diffusion-induced emergent phenomena, appearing in each of the bio-inspired RD-MCNNs, respectively presented in ref 148 and ref 156, is the preliminary polarization of their uncoupled reaction cells on the Edge of Chaos operating regime. Importantly, the NDR of the NaMLab memristor plays a fundamental role for biasing these cells on the Edge of Chaos. Moreover, the two studies revealed the potential of memristors with NDR to enable the reproduction of high-order dynamical phenomena in biology through lower-order electrical circuits.

3.2) High-Frequency AC Response of ReRAM Cells

Depending on the its switching sensitivity to input level and polarity, as well as its memory state(s), a highly nonlinear dynamical process determines the input-induced response of a real-world nonvolatile memristor. The recent research works in ref 158 and ref 159, demonstrated how a system- and circuit-theoretic analysis of a nonlinear system, such as the memristor, enables to draw a comprehensive picture for its response to input/initial condition combinations of interest to the application at hand. In particular, the mechanisms underlying the resistance switching phenomenon that occur in a TaO_x-based nanodevice, manufactured and modeled in Hewlett

Packard (HP) Labs,[160] when stimulated with high-frequency AC periodic inputs, were elucidated on the basis of deep circuit-theoretic study. But what does a high frequency periodic input mean in terms of the memristor response? Quoting the definition provided in ref 158: *"A high-frequency periodic input induces a solution waveform of the device state, which, after some finite time, becomes periodic with such a small peak-to-peak amplitude that the resulting device current-voltage locus exhibits no apparent hysteresis."*

Figure 14a illustrates a seemingly non-hysteretic steady-state current-voltage locus of the TaO_x memristor, induced by a high-frequency zero-mean sinusoidal voltage input (refer to the figure caption for further details). This device is a typical example of a first-order voltage controlled extended memristor,[161] whose high-level differential algebraic equation (DAE) set model reads as

$$i = G(x, v)v, \tag{1}$$

$$\dot{x} = g(x, v), \tag{2}$$

where i denotes the current flowing through the memristor, v is the voltage across the memristor terminals, and x stands for the device only state variable, which physically represents the fraction of the oxide film, which is most conductive, due to a higher distribution of oxygen vacancies within its medium. As a result, the state x of the TaO_x memristor, discussed in this section, may only assume values from the closed set [0,1]. The functions $G(\cdot)$ and $g(\cdot)$ are the memristor memductance and state evolution function, respectively. Further details regarding the physical processes underlying the operation of the TaO_x memristor under study, as well as the specific functional forms of (1) and (2), assumed by its simulation model, can be found in ref 159. Notably, the vast majority of the real-world resistive switches reported to date can be modeled by following the mathematical formulation of extended memristors, as reported in equations (1)-(2). Figure 14b plots the time-waveform of the TaO_x memristor state x, for the same simulation settings as the

ones employed in Figure 14a. The inset in Figure 14b zooms-in to the last three periods of the illustrated x vs t response. Interestingly, at steady-state, x oscillates periodically with a miniscule amplitude about a mean value $\bar{x}_{\mathrm{s}}$, i.e. at steady state, $x(t) \approx \bar{x}_{\mathrm{s}}$ (see the dotted horizontal line in the inset in Figure 14b). Figure 14c visualizes, in detail, the time-response of x over the three first cycles of the periodic input in Figure 14b, that is for $n \in \{1,2,3\}$, where n represents the input cycle number. Δx_n, $\Delta x_n^{(+)}$ and $\Delta x_n^{(-)}$, correspond to the net change, increase and decrease of x over the n^{th} input cycle, respectively. The dotted horizontal lines mark the running mean value $\bar{x}_n$ of the state variable x over each of the three first input cycles. It can be observed, that the memristor state experiences an insignificant net change Δx_n over the course of each cycle n of a high-frequency excitation signal during the transient phase of its oscillation. According to the mathematical analyses introduced in ref 158 and ref 159, which were supported by the afore-described behavioral observations, the value of $\bar{x}$ over each input cycle of a purely-AC periodic high-frequency voltage stimulus $v(t)$, characterized by a very small period T, can be related to the value of Δx over the same cycle through the following analytical equation:

$$\Delta x = \underbrace{\int_0^{T/2} g^{(+)}\big(\bar{x}, v(t)\big)dt}_{\Delta x^{(+)}} + \underbrace{\int_{T/2}^{T} g^{(-)}\big(\bar{x}, v(t)\big)dt}_{\Delta x^{(-)}}. \tag{4}$$

The functions $g^{(+)}(\cdot)$ and $g^{(-)}(\cdot)$ ($g^{(+)}(\cdot) \neq g^{(-)}(\cdot)$) are sign-invariant functions with opposite signs, modeling the device switching kinetics in the positive and negative stimulation cases, respectively. The moment at which transients decay to zero, the time-waveform of x becomes periodic oscillating about $\bar{x}_{\mathrm{s}}$ with a negligible peak-to-peak amplitude (see the inset in Figure 14b). Thus, $\Delta x = 0$ and equation (4) is simplified to the following analytical expression:

$$\underbrace{\int_0^{T/2} g^{(+)}\big(\bar{x}_{\mathrm{s}}, v(t)\big)dt}_{\Delta x^{(+)}} + \underbrace{\int_{T/2}^{T} g^{(-)}\big(\bar{x}_{\mathrm{s}}, v(t)\big)dt}_{\Delta x^{(-)}} = 0. \tag{5}$$

The above equation allows to calculate the mean value $\bar{x}_{\mathrm{s}}$ of the steady-state periodic oscillation in the memristor state x, for a predefined high-frequency periodic voltage input $v(t)$. Importantly, (5) describes the condition for the emergence of the fading memory phenomenon in memristors,[162-164] based on which the steady-state time-response of the state variable $x(t)$, depends on the characteristics of the voltage input $v(t)$, but not on the initial memristor state x_0 (see Figure 14d, e). Through (5), the theoretical investigation performed in ref 158 and ref 159 provided a mathematical proof for the fading memory phenomenon in resistive switches that exhibit asymmetric switching kinetics with respect to the input polarity. The predictive power of equations (4) and (5) were validated through a series of simulation experiments, utilizing periodic inputs with square-wave, sinusoidal, and triangle shapes. As a matter of fact, focusing in the case where the input corresponds to a high-frequency bipolar square-wave AC input, equation (5) may be simplified to a closed-form analytical formula that reads as

$$g^{(+)}(\bar{x}_{\mathrm{s}}, V + V_{\mathrm{o}}) + g^{(-)}(\bar{x}_{\mathrm{s}}, -V + V_{\mathrm{o}}) = 0, \tag{6}$$

where $\hat{v} = V$ is the amplitude of the square-wave voltage input, and V_{o} is the DC offset component (see Figure 14f). The study in ref 158 showed, that, with the use of (6), the TaO_x memristor model may be programmed to any target state from a given set by properly tuning the DC offset level V_{o} of a fixed V-amplitude high-frequency square-wave AC periodic voltage input, as demonstrated in Figure 14g, highlighting a potential application of the discussed theoretical research.

Finally, it should be noted, that the methodologies, described in this manuscript, are not restricted to the specific TaO_x memristor mathematical description, which represents the object of investigation here, but can be applied to any first-order sign-invariant non-volatile memristor DAE set model, irrespective of the expression for its state evolution function, as explained thoroughly

in refs 165-167. Concluding, even though the results discussed in ref 158 and ref 159 are based solely on theoretical studies, it is anticipated that they shall inspire the development of programming algorithms and/or methods for real-world non-volatile memristors in the years to come.

4. Algorithm and architecture

4.1) Algorithms and architectures for high-performance DNN training with analog resistive crosspoint arrays

Background and principle of accelerated DNN training using analog crosspoint arrays. DNNs have become a powerful backbone technology for many artificial intelligence (AI) applications, such as computer vision, speech recognition, and natural language processing. Thus, there is a growing demand for hardware that can perform DNN training faster and more energy-efficiently, as this would enable more advanced AI applications. Training DNNs for large-scale models and datasets is a computationally intensive and time-consuming, even with data center-scale computing resources. Special-purpose digital hardware, including graphics processing units (GPUs), field-programmable gate arrays (FPGAs), and application-specific integrated circuits (ASICs), has been developed and used to address the high computational loads associated with DNN training. As a promising alternative to digital approaches, analog resistive crosspoint array-based architectures have been proposed and studied with expectations of significant performance boost.[168-170]

In principle, analog resistive crosspoint arrays are capable of performing all three core matrix operations (forward pass, backward pass, and weight update) of the standard DNN training algorithm, Stochastic Gradient Descent (SGD), in a fully parallel manner, which can lead to

significant speed-up in DNN training.[169] By storing the weight matrix in the array and applying voltages to the crossbar wires, parallel matrix-vector multiplications corresponding to the forward and backward pass can be performed in parallel using Ohm's law and Kirchhoff's current law.[171] Additionally, with the recent development of array-level parallel update techniques,[169, 172] the entire array can be updated at O(1) complexity, completing fully parallel DNN training on analog resistive crosspoint arrays. However, it has been revealed that the ultimate goal, energy-efficient acceleration of DNN training, can be achieved only when the crosspoint elements meet the stringent requirements.

Device requirements for analog crosspoint array-based DNN training hardware. During the early stages of research, it was anticipated that the error tolerance and noise immunity inherent in neural networks would compensate for imperfections in practical resistive crosspoint devices, thereby ensuring successful DNN training. While it is true that neural networks exhibit some degree of intrinsic error tolerance and noise immunity, it has now become apparent that certain device characteristics can significantly impede the classification accuracy of a DNN if they are not met. Based on simulation studies of neural network training on analog crossbar array architectures,[169, 173-179] four critical device characteristics have been identified as necessary for achieving high classification accuracy: (1) symmetric and bi-directional conductance update, (2) 256 or more conductance states, (3) several seconds of weight retention time constant, and (4) low cycle-to-cycle variation.

Emerging memory device technologies, including ReRAM,[180-183] conductive bridge random-access memory (CBRAM),[184, 185] phase-change memory (PCM),[186, 187] electrochemical random-access memory (ECRAM),[98, 102, 110] ferroelectric field-effect transistor (FeFET),[188] and capacitor-based synapses,[173, 189, 190] have been explored as potential solutions for implementing ideal analog

crosspoint arrays for DNN training. While significant progress has been made in improving the device characteristics of these emerging memory technologies, further improvement in device characteristics is still required to achieve the large acceleration and software-level classification accuracy desired for DNN training.

Algorithms and Architectures to Overcome Device Non-idealities. Recent algorithm studies[191] have demonstrated that it is possible to compensate for some of the critical non-idealities in memory devices used for DNN acceleration. One example of an algorithmic remedy is the Tiki-Taka algorithm (TT),[192] which addresses the challenges of non-ideal memory elements in analog resistive crosspoint arrays. Developed by the IBM group, TT utilizes the interaction between the core array and an auxiliary array to compensate for the negative impact of update asymmetry during training. The auxiliary array accumulates weight change values, which are then buffered and transferred to the core array using a specific threshold. A quantitative study has analyzed the update symmetry requirement for core and auxiliary arrays separately, showing that it could be further relaxed.[193] TTv2, the second version of the Tiki-Taka algorithm, uses an additional digital array to significantly reduce the number of conductance states required.[194] This enables the training of DNNs on extremely noisy devices and reduces the number of states required for the auxiliary device to barely 10 levels. Experimental demonstrations of the algorithms on ECRAM[195] and ReRAM arrays[196] have been reported, which verify the successful training with practical devices. These studies suggest that algorithmic remedies can overcome non-idealities in memory devices and improve software-level classification accuracy. (Figure 15)

On the other hand, architectures have been developed to overcome non-idealities in memory devices. For example, Burr et al. proposed an architecture that utilizes a pair of PCM and capacitor-based synapses (3T1C) to overcome the non-idealities of each cell.[197] While 3T1C cells are capable

of fine-grained, bi-directional, and symmetric conductance updates, they suffer from volatility and significant asymmetry deviation due to the spatial variability of the MOSFET's current level. PCMs, on the other hand, have reliable memory retention but require periodic reset cycles to address uni-directional updates. The architecture that combines PCM and 3T1C achieves highly precise and nonvolatile on-chip learning. By using alternating polarity inversion to correct asymmetry deviation in 3T1C, it is possible to train DNNs on the chip with the same level of precision as the floating-point equivalent reference level. This architectural design has been studied further and is considered a potential candidate for future realization.

The algorithmic and architectural studies demonstrate the potential of analog resistive crosspoint arrays for DNN training, even in the presence of non-ideal memory elements. As research in this field continues, further advances will be made to improve the efficiency and accuracy of DNN training on analog crosspoint arrays.

4.2) Modeling and analysis of device thermal variation for reliable DNN inference

With the employed DNN model, the intense computation inside the ReRAM array increases the power density within the small area, which subsequently heat the device. The increased temperature will weaken the resistive memory cell to retain the programmed values, the conductance states will drift away from the initial value, leading to significant accuracy degradation in CIM.[198, 199] On the algorithm side, structured pruning algorithm remove the redundant weights in a group-wise fashion, creating the column-level or array-level computation skipping during CIM.[200-202] The prevalent usage of pruning algorithms changes the weight distribution due to the reduction of redundancy. In the meantime, the programmed weight values are also affected by the conductance change. Therefore, analyzing the impact of the thermal variations on the sparse model is critical but remains under-explored.

The group Lasso[200] algorithm penalizes the group of weights via regularization during DNN training. On the hardware side, low precision weights can be represented as the shifted positive weights subtracted by the middle-value offset. In NeuroSim,[203] the synaptic ReRAM array will be programmed as the positive weight, and the adjacent dummy column stores the global mid-value offset. With 4-bit weights, shifting and programming all the level "0" (0000) to level "7" (0111) increases the percentage of the LRS inside the synaptic array, regardless of structured or unstructured pruning.

However, since the large number of sparse groups can be skipped during CIM, the structured pruning aims to "concentrate-and-conquer" the sensitive LRS level via group Lasso. On the contrary, element-wise pruning generates a high percentage of non-skippable level "7" in the ReRAM crossbar array, amplifying LRS distortion and degrading the model robustness and accuracy. With the measured conductance from our ReRAM prototype chip,[204] Figure 16a shows the "accuracy-robustness-sparsity" tradeoff with structured and unstructured sparsity.[205] Under the context of the dummy column-based deployment scheme, quantization-aided structured pruning enhances the model's robustness to the thermal variations.

In practice, quantization digitizes each floating-point weight element to the nearest quantization level. In the meantime, group Lasso penalizes the weights in a structured manner. Although the unimportant groups will be collectively quantized to zero, the granularity mismatch between pruning and quantization introduces the residual non-structured zeros inside the weight distribution. The residual element-wise sparsity will further degrade the model's robustness during CIM. Motivated by this, a robustness-aware structured pruning algorithm, minimizing the residual element-wise sparsity with global penalty threshold, leading to the comprehensive pruning candidate selection and accurate group-wise pruning, was proposed. Figure 16b demonstrates the

overall flow of the proposed algorithm. Together with the progressive knowledge distillation (PKD) and batch normalization adaptation (BNA) scheme,[206] the proposed algorithm maintains >92 % and >90 % inference accuracy for ResNet-18 and VGG-8 models, respectively, as shown in Figure 16c-f. Along with the robustness enhancement, the proposed scheme also compresses the dense ResNet-18 and VGG-8 models by 24.5× and 24.9×, respectively.

4.3) Design exploration with analog noises for a computing system

The objective of the hardware-software codesign for analog resistive memory is to deal with the impact of non-idealities of resistive memory on computing performance. Resistive memory mainly suffers from non-idealities. The first one is programming variation when their resistance is changed, which is also accompanied with large programing energy and duration compared to transistor switching. The second one is the read noise, or the conductance fluctuation when the bias voltage is smaller than the programming threshold. These non-idealities may outweigh the energy-area efficiency advantages thanks to in-memory computing in some application scenarios.[50] As a result, the application and the hardware-software need to be designed in a manner that can either leverage or mitigate these non-idealities. The following Table 1 summarizes some of the initiatives along this direction, which are discussed in details in the follows.

Hardware-software codesign that leverages the read noise. Combinatorial optimization can leverage the intrinsic read variation of resistive memory to prevent the system from being trapped in local minimum during the evolution to the global minimum. For example, a frequently employed technique in simulated annealing is to inject noise of decaying amplitude, equivalent to the lowering down of the annealing temperature. To implement that, hardware-wise, heuristic threshold neurons[207] and weight scaling[208] can achieve run-time temperature scaling. In addition, the effective temperature can also be modulated by the strength of self-feedback in a dynamic

system, such as the diagonal resistive memory elements in a hardware recurrent network.[209] Software-wise, simulated annealing algorithms using Hopfield neural network[207, 209] or Boltzmann machines[208] have been devised, with application to representative combinatorial optimization problems such as graph partition.

Hardware-software codesign that leverages the program noise. Neural networks of random features, such as the echo state networks and extreme learning machines, turn the programming noise into an advantage. Such networks are favored in edge learning due to largely reduced training complexity, which complements memristive systems practicing local learning rules.[210, 211] To implement these networks, hardware-wise, the programming stochasticity of the resistive memory will be used to produce random and fixed conductance matrices, serving the weights of random projections in those networks. (see Figure 17a) Software-wise, the echo state graph neural network[212] has been introduced for graph-structured data learning at significantly reduced learning cost and improved efficiency, with applications from social networks to drug discovery. Notably, such an echo state graph neural network can serve as the graph feature extractor when combined with a trainable projection layer and associative memory, forming a memory-augmented graph neural network for few-shot graph learning.[213] In addition, pairing echo state network-based temporal feature extractors with random convolutional-pooling architecture-based spatial feature extractors can naturally learn spatial-temporal signals at low cost.[214]

Bayesian inference may also exploit the programming variation of resistive memory. Hardware-wise, the 1-transistor-1-resistive memory (1T1R) array allows physically implementation of the proposal distribution using the cycle-to-cycle and device-to-device programming variation of resistive memory.[215] Software-wise, the Metropolis-Hasting Markov chain Monte Claro (MCMC)

sampling has been used for both supervised classification and reinforcement learning classical control problems.[215]

Cybersecurity may use the programing stochasticity of resistive memory. Hardware-wise, the electrical responses of a resistive memory crossbar with intrinsic programming variation and nonlinearity to external bias cannot be duplicated, thanks to the underlying ionic nature of resistive switching which serves as the source of entropy.[216] (see Figure 17b) Software-wise, memory-based physical unclonable function enables downstream applications such as authentication.[216]

Hardware-software codesigns that withstand noises. Hyperdimensional computing is an emerging neutrally inspired method for representing and manipulating data in a high-dimensional space. To implement that, hardware-wise, resistive memory array with in-memory search capabilities has been demonstrated.[217, 218] Since high dimensional vectors are mostly orthogonal using the random basis, it naturally mitigates the impact of weight imprecision of the resistive memory. Software-wise, the hyperdimensional encoder and associative memory have been used in supervised classification and few-shot learning of images.[217, 218]

Linear system solver is the basis of scientific computing. To accelerate linear system solver, hardware-wise, hybrid analog-digital system was developed to combine the high efficiency of analog computing and the high precision of digital computing. This allows operations of different precision demanding to be deployed on different platforms.[219] (see Figure 17c) In addition, multiple resistive memory cells can be used to represent a matrix element at higher precision.[220] Software-wise, iterative refinement algorithm using an inner conjugated gradient solver have been implemented on the digital and analog platforms, respectively, for applications such as partial correlation.[219] Moreover, Newton's method has been used for solving partial differential equations.[220]

Future Perspectives. The pioneer exploration of hardware-software codesign so far has well illustrated the advantage of analog in-memory computing in multiple important branches of computing. To move further, the controllability, such as the tunable range of the read and programming noise, as well as the programming energy and duration are the parameters to be optimized to practically parallel or even excel the digital counterparts.

5. Computing-in-memory system for AI accelerators

5.1) Development of computing-in-memory based AI accelerators

There is a significant interest in developing custom accelerators for deep learning inference. Most of the research and commercial effort is centered on digital accelerators that optimize the computation and dataflow specifically for inference applications. Digital accelerators typically contain arrays of processor elements (PEs) that can perform several MVM operations in pipeline-parallel fashion. However, despite all the advances, empirical evidence based on commercial AI accelerators indicate plateauing of energy efficiency at 10 TOPS/W or 100 fJ/Operation.[221] This is attributed to the energy costs associated with data movement as indicated in Figure 18a. MVM operations cost tens of fJ whereas accessing on-chip SRAM costs about 1pJ/byte. Even with significant weight reuse, each MVM operation comes with significant memory access overhead.[222] A promising compute paradigm that aims to address this is analog CIM.[223-235] The essential idea is to fabricate a weight stationary array of synaptic unit cells based on ultra-dense memory devices (see Figure 18b). Each unit cell holds one synaptic weight that can remain stationary for many inferences. The MVM operations themselves are performed by exploiting physical attributes such as Kirchhoff's circuit laws or capacitive charge sharing. CIM based AI accelerators improve AI inference energy-efficiency in three ways. First, it performs computation directly inside statically

stored weights of AI models, thus eliminating energy-consuming data movement of weights from memory to separate compute units. Second, in conventional memory, data is read one row at a time. Each time, column pre-charge energy and sense-amplifier analog-to-digital conversion energy are incurred. In CIM, however, both column pre-charge and analog-to-digital conversion are performed only once. Therefore, energy spent inside the array is reduced. Finally, depending on specific implementations of CIM, analog multiply-accumulate operations performed by memory cells and metal wires can be more energy-efficient than digital counterparts performed by logical gates.

A prominent candidate for CIM is an approach based on SRAM cells and switched capacitors which is almost ready for commercialization. However, the primary drawback is that SRAM-based unit cell arrays are not dense enough to hold all the weights of a large DNN due to chip area constraints. This would necessitate frequent DRAM accesses. Compared to this approach, memristive devices could offer much higher weight density. Moreover, the non-volatile storage capability facilitates power cycling of the accelerator. All the demonstrations to-date that are silicon verified are based on resistive memory technology such as PCM,[224, 236] and metal-oxide based ReRAM.[225] Flash memory-based CIM which is closer to commercialization could also afford higher weight density.[226] However, compared to Flash, the backend integration capability of memristive devices ensures that one could exploit more advanced CMOS technology nodes.

The development of CIM system experienced three periods, according to the completeness and diversity of the system. As shown in Figure 18c, the diversity means how many kinds of computing tasks the system can support, and the completeness means whether the system contains all the necessary functional modules. The first period focused on the array level demonstration with a small-scale memristor array.[237-239] For example, Patrick *et al*. demonstrated a sparse coding task

with a 32×32 passive crossbar,[237] and Yao *et al*. demonstrated face recognition with a 128×8 1T1R array.[238] The second period transferred to macro circuit design and multi-array system construction. During this period, the integration density of memristor devices increased fast, and more functional modules were included. The modules mainly concentrated on analog circuits for supporting the computing process, like ADC, special function units, etc. Liu *et al*. reported a monolithic integrated ASIC chip, which integrated two memristor arrays up to 160 Kb density and supported complete MLP inference on chip.[240] Yao *et al*. demonstrated multilayer CNN on a fully-hardware implemented system with eight 2 Kb memristor arrays and all the necessary analog circuits, while the digital parts were realized with an integrated FPGA.[241] During the third period, researches started to explore fully-parallel on-chip neural network inference.[224] Wan *et al*. developed a multi-core CIM chip to perform diversity tasks with high accuracy.[225] Recently, IBM reports a multi-core CIM chip with complete on-chip routing solution,[228] making important progress on the completeness of the CIM system.

To construct a practical CIM system, the correlated software ecology should be developed in addition to the efforts on the hardware part. At present, studies on the software level mainly include two aspects: hardware simulation and algorithm model deployment, as shown in Figure 19. Simulation software aims to model and analyze the hardware behaviors from device level to circuit level and to system level. With the simulation software, design guidelines for each level can be provided. System level simulation tools can model the timing, data scheduling, resource occupancy, etc. Jiang *et al*. developed a high-level architectural simulator to model the timing and resource of each module, like bus, buffer, and computing array, under the real working scenes.[242] Lee *et al*. designed a clock-accurate simulator to model the effects of various NoC structures on the performance of a multi-core system.[243] Circuit level simulation tools aim to estimate the power

consumption, speed, area, and computing accuracy of circuit modules or the whole chip. Zhang *et al*. developed a complete algorithm to device simulator to benchmark the influence of array size, input voltage, ADC precision and non-ideal effects.[244] Peng *et al*. proposed an end-to-end simulator for both training and inference.[93] This simulator considers all the non-ideal effects of the memristor devices and arrays. Device level simulation tools focus on the physical mechanisms of I-V characteristics, retention, endurance, etc., and the influence of these behaviors on the computing accuracy. Liu *et al*. developed a compact model for the reliability characteristics of memristors, and embedded the model into circuit level simulation tools.[245] Gao *et al*. constructed a kinetic Monte Carlo tool to model the distribution of oxygen vacancies during resistive switching process.[246] On the other hand, deployment software aims to link algorithm applications to the CIM hardware, and map the model parameters to the real circuits. Deployment software decides how much the performance and resource of CIM hardware can be utilized. Ambrosi *et al*. developed an Open Neural Network Exchange (ONNX) software stack for algorithm compiling, as well as an instruction set for decoupling the algorithm development and hardware design.[247] Zhang *et al*. proposed a convolutional kernel duplication method to enhance the throughput of different algorithm models.[248] Peng *et al*. proposed a weight placement method to increase the reuse rate of input data,[249] which can also enhance the throughput.

5.2) Passively integrated memristor crossbar array without select-transistors

Lateral dimensions of some emerging memory devices could be scaled to below ten nanometers without sacrificing their analog properties and retention.[250] The density of stand-alone memory devices can be sustained at the circuit level by passively integrating them into crossbar arrays.[251-267] The monolithic (passive) vertical integration allows further increasing effective density.[252-256, 268] Such 3D circuits would enable, for example, much higher performance neuromorphic

computing hardware because large neural models can be entirely stored on a chip thus avoiding energy-taxing and slow off-chip communications.[251, 269]

In the previous decade, there has been a significant effort in developing passively integrated memory crossbar circuits,[27, 40, 115, 237, 251, 255, 257, 259-267] with industry primarily focusing on simpler memory applications.[252-254, 258, 270] For example, the published results supported with proper crossbar array statistics include metal-oxide filamentary[253, 255, 256, 262, 263] and interfacial,[237, 257] solid-state-electrolyte,[258, 271] and 2D-material memory devices.[40] However, the focus now has mainly shifted to sparser, actively-integrated ("1T1R") memristive crossbar circuits in which each memory device is coupled with a "select" transistor.[19] The primary reason is the much more demanding requirements for the device technology in passively integrated circuits.

Indeed, in active crossbar memory circuits, the select transistor and additional control lines simplify access to each cell in the array (Figure 20a). A voltage can be applied to and the current measured from a "selected" memory cell, while ensuring zero biases and currents via all other cells. Furthermore, the select transistor could be used as a current compliance for precise conductance tuning. In passive ("0T1R") crossbar circuits (Figure 20b), a selected cell is directly connected to a set of "half-selected" cells sharing the same electrodes, and indirectly connected (via half-selected cells) to all remaining cells in the array. In the most typical "half-biasing" scheme used for writing the selected cell, half of the write voltages of appropriate polarity are applied to the selected cell's electrodes while all other electrodes are biased to zero (Figure 20b). The first challenge in the such scheme is "sneak-path" currents via half-selected cells. Sneak-path currents increase with scaling up crossbar dimensions and could result in non-negligible IR drops on the electrodes and hence inaccurate operation. This problem is especially bad for forming and write operations because of much larger applied voltages. The second, more serious challenge is due to

variations in memristors' effective switching threshold voltages. Half-selected cells receive half of the write voltage of the selected cell. With large device-to-device variations, such voltage may be large enough to disturb the state of the half-selected memristors (Figure 20c) making the accurate conductance tuning in the whole array challenging. The disturbance is especially bad for the half-biased memristors with lower switching thresholds, when a larger write voltage must be applied to the selected memristor with a higher switching threshold.

Increasing crossbar array dimensions is critical for reducing the area overhead of peripheral circuits. Practical dimensions vary for different applications and, e.g., are around a hundred for some neuromorphic inference applications.[251, 269] However, increasing crossbar array dimensions exacerbate the already discussed half-select disturbance and IR drops. A general solution for the latter problem is to decrease currents via memristors, i.e., making the resistance of the electrode lines negligible to those of the crossbar array's crosspoint devices. Another solution is to utilize "1D1R" circuits[260] based on memristors with rectifying *I-V* characteristics. In a related "1S1R" approach, a memory cell is connected with a highly nonlinear two-terminal selector device, such as the ovonic threshold switch in Intel's 3DXpoint technology.[270] A common strategy for improving device uniformity in filamentary devices is to constrict the electrical and ionic currents during forming step, *e.g.*, by specifically engineering electrodes with protrusions or wedge shapes[272] or creating a special opening for ion transport.[273] A more advanced approach of engineered threading dislocations resulted in the smallest variations reported for the integrated crossbar circuit based on filamentary devices.[115] Interfacial memristors typically feature better device uniformity over filamentary ones due to the non-localized switching mechanism.[261] However, improvements in uniformity may be achieved at the cost of sacrificing analog properties and retention. For example, a lower energy barrier for memory mechanism (and hence lower

retention) is typical for interfacial and solid-state electrolyte memristors and devices with sub-microamp-level write currents. Moreover, the devices with sharp *I-V* nonlinearity (such as in 1S1R devices[270]) are not suitable for an implementation of an accurate product between analog input voltage and device conductance.

This is why several recent studies were focused on filamentary metal-oxide memristors, e.g., based on bilayer Al_2O_3 / TiO_{2-x} structures,[251, 264-266] that have excellent analog properties and retention and can be scaled to below 10 nm.[250] For example, larger crossbar dimensions in ref 251 (Figure 21a) over prior work[262, 263] were achieved by optimizing memristive devices and substantially decreasing electrode resistances with improved fabrication process. Specifically, Al_2O_3 tunnel barrier thickness was optimized to achieve linear device *I-V*s at smaller non-disturbing voltages, which is crucial for accurate in-memory computing, and mild nonlinearly at higher voltages to reduce sneakpath currents during forming and write operations (Figure 20c). The electrode resistance was lowered by using higher-conductance metals, and thicker and more planar geometry by adopting industrial CMOS patterning processes. The improved electrode conductance allowed increasing pre-forming device conductances, which was achieved by increasing oxygen vacancy concerntrations, without running into IR drop problems. That, in turn, led to lowering of forming voltages and hence less damaging, more controllable filament formation process and ultimately more uniform device-to-device *I-V* characteristics. The improved device uniformity allowed programing ~99 % crossbar array memory devices with an effective 4-bit precision (Figure 21b).

In summary, passive crossbar memory circuits enable much higher integration densities that would be crucial for memory and in-memory computing applications. However, such technology is much more challenging as compared to actively integrated crossbar circuits. The further

advances in passive crossbar memory circuits technology would likely largely depend on the progress in its active counterparts, specifically from improvements in device uniformity and lowering of device currents that is desired for reducing select transistor overhead.

5.3) Fully integrated multi-core CIM chips

For the memristor-based AI accelerators to be broadly adopted, CIM systems need to support a wide range of AI tasks while achieving software-comparable inference accuracy. However, the same factors that makes CIM energy-efficient also makes it challenging to simultaneously be reconfigurable and accurate. For instance, due to the high programming cost of today's resistive RAM devices, the statically stored weights cannot be reconfigured for different neural network layers during inference; similarly, the robustness of analog computation is prone to device and circuit non-idealities, which leads to accuracy loss. To address this trade-off between efficiency, reconfigurability, and accuracy, Wan *et al*.[225] demonstrated the NeuRRAM chip (Figure 22a). NeuRRAM integrates 48 CIM cores with a total of 3 million RRAM devices. It simultaneously realized high energy-efficiency, reconfigurability to support diverse AI model architectures including convolutional neural networks (CNN)[275] long short-term memory (LSTM)[276] and Restricted Boltzmann Machine (RBM),[277] and software-comparable accuracy across various AI benchmarks. Importantly, while most AI benchmark results reported in previous studies were obtained using a hybrid of software simulation and hardware measurement, all results reported in this work were measured directly from hardware. Previous CIM chips typically only support MVM in a single direction by hard-wiring input DACs to the rows and output ADCs to the columns, or vice versa. Such design cannot perform MVM in the reverse direction (i.e., with transpose of the weight matrix), which is needed for models such as RBM or computing backward gradient during the back-propagation training. NeuRRAM implements a transposable neurosynaptic array

architecture (TNSA) that allows dynamic reconfigurability of MVM directions with minimal energy or area overheads. Unlike conventional CIM designs where ADCs are at the periphery of memory array, TNSA physically interleaves RRAM weights and ADCs and connect them through bi-directional pass gates (Figure 22b, c). It allows DAC inputs and ADC outputs to be sent and received in any arbitrary directions inside the array. Despite being fabricated at an older technology node, NeuRRAM can achieve 2 times better energy-delay-product, a commonly used metric for energy-efficiency, than previous RRAM-CIM chips. The benefit comes from a voltage-sensing scheme for in-memory MVM.[278] The conventional current-sensing design is to use voltage as input and measure the current as the results based on Ohm's law. Such design has limited parallelism and energy-efficiency due to large array current. The voltage-sensing scheme of NeuRRAM keeps output lines floating during MVM and measures settled open-circuit voltage on the output lines. The design eliminates power- and area-consuming peripheral circuits such as voltage clamps and current mirrors. It also allows simultaneous activation of all rows and columns and thus realizing the maximum parallelism.

To mitigate the impact of hardware non-idealities on inference accuracy, Wan *et al*. proposed a series of algorithm-hardware co-optimization techniques.[225] They include (1) model-driven hardware calibration that calibrates hardware using real data from training datasets, (2) non-idealities-aware training that statistically models hardware non-idealities during model training, and (3) chip-in-the-loop progressive finetuning that uses the chip to perform the forward-pass of fine-tuning such that the model adapts to individual chips' characteristics. Together, these techniques enable NeuRRAM to achieve comparable inference accuracy to software models (with 4-bit weights) across various AI benchmarks, including CIFAR-10[279] and MNIST[280] image

classification, Google speech common recognition,[281] and a Bayesian image recovery task. (Figure 22d)

Figure 23 shows a compute core that is based on PCM embedded in an advanced 14-nm CMOS technology node.[227, 228] Compact current controlled-oscillator-based ADCs allow the MVM operation to be executed at constant time complexity without requiring any time-multiplexing since the pitches of the ADCs and the unit cells match. Despite the integration of the PCM element at a relatively high position in the metal stack, a throughput density of 1.59 TOPS/mm^2 is achieved. It is conceivable that by integrating the PCM devices closer to the transistor-level at a denser pitch, substantially higher compute density can be achieved.

In fact, compute density is one of the most desirable attributes for memristive CIM-based DNN accelerators. It would be a critical breakthrough if we could conceive memory devices that could provide the ability to hold large DNNs in a reasonable die size without ever having to fetch the synaptic weights from DRAM. There is significant commercial opportunity in developing discrete accelerator chips that could run very large neural network models at a power budget unthinkable with state-of-the-art solutions. Yet, maintaining high performance and energy efficiency at the application level requires a careful optimization of the system architecture. Recently, it was shown that a highly heterogeneous and programmable CIM accelerator architecture featuring spatially distributed CIM tiles, a dense and flexible 2D mesh-based interconnect with massively parallel data bus, and special-function digital cores can support inference of a variety of large models with high energy efficiency.[229] Moreover, the architecture could be enhanced with custom digital functionality to achieve a constant and high throughput regardless of the batch size, potentially attractive for throughput-critical applications.[230]

Another commercial opportunity for CIM is the space of embedded systems where CIM tiles can be used alongside programmable processors to enhance the operational flexibility. In such architectures, the CIM tiles could hold the synaptic weights and perform MAC operations whereas the processor could handle the residual digital operations, data management and control.[231] Moreover, a tighter integration between the processor and CIM tiles could be envisaged where the processor uses an instruction set extension to access CIM tiles.[232]

One key challenge for CIM-based DNN accelerators is achieving no or minimal accuracy loss in comparison to high precision digital accelerators. Existing compute precision when using analog encoding of synaptic weights on a single device is equivalent to 3-4-bit fixed point arithmetic. It is expected that with improved devices,[233] programming algorithms[234] and synaptic architectures,[235] this could be improved significantly.

6. Other approaches of memristive technology

6.1) In sensor computing and 3D heterointegration

According to the projection by Semiconductor Research Corporation and Semiconductor Industry Association, the number of sensor node exponentially increases with the development of Internet of Things. By 2032, the number of sensors is expected to be ~45 trillion, which will generate >1 million zettabytes (1027 bytes) of data per year.[282] The massive data from sensor nodes obscures valuable information that we need it most. Abundant data movement between sensor and processing unit greatly increases power consumption and time latency, which poises grand challenges for the power-constraint and widely distributed sensor nodes that are required in Internet of Things. Therefore, it urgently requires computation paradigm that can efficiently process information near or inside sensors, eliminate the redundant data, and reduce frequent data

transfer. The manufacturing technology of sensors and processing units are quite disparate, which unavoidably leads to physical separation between the two units. The data transfer between them inevitably causes time delay and power consumption. Researchers have adopted advanced packaging technique or monolithic integration to reduce their physical distance, which is the so-called near-sensor computing technology.[283] To fuse the functionalities of sensor and computation units, scientists have demonstrated the multiple devices with both sensing and computing functionalities for in-sensor computing that avoids the data transfer outside the sensors.[283-292] Based on the response characteristics of the sensors, different computation functionalities can be realized,[293-296] as shown in Figure 24.

Vision sensor is a good example to illustrate the in-sensor computing. Vision is the dominant source for human to collect information from external environment.[297] Optoelectronic conversion is an indispensable process in the image sensing, which also constitutes abundant information reconstruction and reorganization. The use of optoelectronic devices for image sensing and information processing provides a way for in-sensor computation of vision information. In response to the intensity of light stimulation, the sensors can show linear and nonlinear response characteristics, which allows us to realize feature enhancement or image recognition within in a sensor array.[294] For the sensors with linear response characteristics, they can be adopted for the computation with an artificial neural network. The implementation of machine learning algorithms mainly consists of multiplication and summation. Optoelectronic conversion can be used for the hardware implementation of the multiplication function $I_{ph} = R \times P$, where I_{ph} is the photocurrent, R is the photoresponsivity and P is the power intensity of light stimulation. The photo-responsivity of the photo-sensors can be continuously modulated by gate voltage, emulating the synaptic weight in an artificial neural network. By connecting a number of photo-sensors in

series, the summation can be executed according to Kirchhoff's current law. Mennel *et al*. reported this proof-of-concept demonstration with two-dimensional semiconductors, exhibiting the ultrafast image recognition at the level of nanosecond.[288] Jang *et al*. extended this in-sensor computing strategy with mature Si technology at a large-scale.[295]

The nonlinear response characteristics of photo-sensors can reconstruct the light stimuli, which provides alternative ways for processing visual information. For the sensor with superlinear response characteristics, strong light stimuli can produce high photo-conductance and, while the weak light stimuli generate only negligible photo-conductance. Thus, the sensor array can enhance the feature of the image and reduce the background noise. The photo-responsivity of photo-sensor can be also modulated over a large range to adapt different light stimuli, exhibiting sublinear response characteristics.[290] This design can be used for the in-sensor vision adaptation under different lighting conditions (Figure 25a). For the sensors with threshold response characteristics, we can use it for the nociceptor that can respond to the stimuli beyond a certain threshold.[298, 299]

To achieve low latency and high energy efficiency in real-time AI applications, it is essential to eliminate unnecessary data movement and promote parallel data processing.[256, 300, 301] Three-dimensional heterogeneous integration (3DHI) utilizes high-density, reduced surface area to seamlessly integrate different functional layers such as sensors, processors, and memory.[302-309] Since 3DHI can process large amounts of data obtained from sensor networks with high bandwidth and low latency, efficient sensory data processing can be achieved by combining 3DHI with in-sensor/near-sensor computing architectures.[283] However, challenges related to modality, materials, data processing, and computing architecture limit the development of 3DHI. Regarding materials, the number of layers that can be stacked is limited because intimately connected layers affect each other.[310-314] Normally, this requires costly processes such as wafer bonding or transfer-printing.[314-

[316] Regarding modality, only one sensory modality can be accommodated per one fixed sensor connected to the 3D chip.[317] Regarding data processing, additional processors cannot be added to a fixed 3D chip for various computing environments such as a highly noisy or mixed-precision processing environment.[219] Regarding computing architecture, most of heterogeneous chips developed to date are based on conventional von Neumann architecture, which does not fully utilize high bandwidth area interfaces for parallel data processing. This issue could be solved by using neuromorphic computing architecture.[285, 288, 318, 319] The other issues, however, cannot be solved without having the ability to rearrange and reconfigure each layer in the 3D chip stack.

One potential solution to the aforementioned issue is to utilize 3DHI with embedded neuromorphic computing crossbar arrays, where each heterogeneous chip layer can be stacked and reconfigurable by chip-to-chip light communication. An array of optoelectronics enables chip-to-chip light communication, which makes it possible to replace, insert, stack, and restack chip modules with various functionalities.[320] A stackable hetero-integrated system with embedded neuromorphic AI chips and optoelectronics, including photodiodes and light-emitting diodes (LED) is depicted in Figure 25b-d. As shown in Figure 25b, the hardwire-connected conventional 3DHIs still have significant drawbacks, however stackable hetero-integrated chips have several advantages over them. Like a Lego block, every AI- and optoelectronic-embedded layer allow us to i) replace either the processor or the sensor, depending on the need for either sensing or computing, ii) stack different layers to improve or reconfigure neural networks, and iii) add or remove layers to optimize the processing function. (Figure 25c) Examples of the aforementioned cases are as follows: i) a variety of sensors could be readily replaced, or different pre-trained computing layers could be replaced to process different sensory inputs; ii) pre-trained computing layers could be continuously stacked for highly parallel kernel operations to detect varying and

various signals from sensor layers; and iii) computing layers can be added to manage more complex computing requirements or removed to perform lighter computing tasks. As a result, stackable hetero-integrated chips can be customized depending on the purpose of the chips and the sensing modality. In other words, functional layers can be reconfigured easily due to chip-to-chip light communication. The stackable hetero-integrated chip enables various neuromorphic sensors and processors to be seamlessly integrated. In addition to that, the stackable hetero-integrated chips provide a high degree of freedom for tailoring near-sensor computing design and high versatility in edge computing. Despite the advantages of stackable hetero-integrated chips, there is a very important issue to solve, alignment of each chip. Precise alignment between chips is critical for heterogeneous chip module integration because misalignment has a significant impact on the speed, energy efficiency, and crosstalk of light communication. As a result, for practical application, high-precision alignment techniques that do not require micromanipulators should be developed. For example, additional macro patterning of V-grooves via anisotropic etching of the silicon substrate can be used as a self-alignment scheme.[321, 322] Such structure will enable easy and accurate alignment by hands. Other passive alignment strategies including optical plugs, ferrules, and fiber-arrays can also be used for precise self-alignment of the chips.[323-325] These approaches will enable stackable hetero-integrated chips to be readily reconfigurable.

6.2) Analog circuit design based on memristor arrays

Analog computing, which operates on continuous signals in real-time, has been an important computational paradigm since the advent of electronics. However, the advent of digital computing, with its high accuracy, reliability, and ease of use, has caused analog computing to fall out of favor over the past few decades. The world is analog, where the ubiquitous sensors have led to an analog data deluge. Due to the notoriously energy and time-consuming conversion between analog and

digital domains, recent developments in big data and the Internet of Things (IoT) have created challenges for digital computing to process video/audio information, particularly in terms of energy efficiency and processing throughput. As a result, analog computing is now experiencing a resurgence in interest as a potential solution to these challenges.[12, 50, 239, 326] Despite this growing interest, the development of analog circuits has been impeded by a lack of reconfigurable and scalable platforms for fast analog circuit prototyping and verification, similar to the widely-used field-programmable gate arrays (FPGAs) for digital circuits. This lack of infrastructure has slowed progress in the field, particularly in comparison to the rapid advancements in digital circuit technology. Recently memristors, non-volatile analog memory devices, have been utilized to design analog circuits including analog content-addressable memories (CAM)[327] and field-programmable analog arrays (FPAAs),[328] because of their fast-switching speed and multiple data storage up to thousands of levels.[5, 329-332] Such memristor-based analog circuits can reduce the footprint required for analog circuitry while providing faster programming speeds and more accurate performance with tunability.

The platform for a memristive field-programmable analog array (memFPAA) has been recently demonstrated by Li *et al*.[328] A FPAA usually comprises a monolithic collection of configurable analog blocks (CABs). However, in this study, memristive devices serve as core analog elements while CMOS-based peripheral circuits are used. To investigate the capabilities of the memFPAA, several analog computing tasks have been implemented, including band pass filters, audio equalizers, and an acoustic mixed frequency classifier. As illustrated in Figure 26, the memFPAA is able to emulate the biological hearing system. As the basilar membrane in the human hearing system, which selectively resonating to audio inputs with various frequencies, the band pass filters together with VMM synaptic array and peak detectors can selectively differentiate specific

frequency bands of the sinusoidal inputs. In this case, the post-synaptic neurons receive the weighted sum of signals from the band pass filters, and the neurons connected to the synaptic memristors with higher weights that match the input signal frequencies will produce a larger output voltage. This process closely resembles the biological process where hair cells detect basilar membrane vibrations and send signals to the cerebral cortex. Furthermore, memristor-based FPAAs can be integrated into large-scale circuits to operate as either a general-purpose reconfigurable analog circuit or as an FPGA-like platform that rapidly prototypes analog circuit designs. The results of this study demonstrate the potential of memristive devices for use in analog circuits.

6.3) Quantum effects in memristive devices

The development of device concepts and architectures that rely on quantum effects can enable technologies for information processing impacting the entire sector of electronics. Among emerging technologies, memristive devices represent suitable platforms for exploring quantum effects at room temperature in air.[31, 333, 334] By coupling ionics with electronics, these devices can exhibit under appropriate operational conditions typical features of quantum conductance, where conduction does not follow the Ohm's law. In this conduction regime, the electron flow is regulated by ballistic transport with a discrete number of modes (or channels) that contributes to the overall conduction. This occurs when the filament size of memristive devices is shrunk down to act as a constriction where the lateral size becomes comparable with the electron wavelength, resulting in the formation of a quantum point contact (QPC). In this framework, memristive cells can be schematized as two electrodes *A* and *B* that behave as ideal electron reservoirs in thermal equilibrium and separated by a narrow constriction of length *L* and width *W* (Figure 27a), where the density of states can be calculated by considering the case of a particle (electron) in a box with

infinite walls (Figure 27b).[31] In this case, conductance is quantized in multiples of the fundamental quantum of conductance $G_0 = 2e^2/h$, as theoretically described in the framework of the Landauer theory.[335] While in semiconductors the QPC is a mesoscopic object due to the large Fermi wavelength, QPC in metallic nanofilaments has necessarily atomic dimensions since the Fermi wavelength is of the same order of magnitude as the atomic separation. For this reason, the observation of quantum effects in memristive devices is intrinsically related to the capability of manipulating the filament morphology down to the nearly atomic scale.

Pioneering works reported the formation of metallic QPC by creating contacts with atomic-scale control of the distance through mechanical positioning systems. This includes the mechanically controllable break-junction technique,[336] that involves breaking/recontacting a metallic wire through piezoelectric elements, and QPC preparation by contacting a metal surface with Scanning Tunneling Microscopy (STM).[337] As an alternative, Terabe *et al*.[338] reported the realization of atomic point contacts by exploiting solid-state electrochemical reactions, where atoms are manipulated through the applied electric field without the need of any mechanical positioning system. Similarly, quantum conductance effects arising from electrochemical phenomena responsible for the formation/rupture of nanofilaments were later observed in the nanogap of atomic-switch devices[22] and in gapless metal-insulating-metal memristive cells.[31, 333, 334] While the filament morphology in memristive cells can be manipulated through the applied voltage that drives the formation/rupture of the filaments, it is worth mentioning that also additional supportive and destructive forces including corrosion, Joule heating and oxidative/corrosive action of the surrounding matrix contribute to determining the filament morphology.[31] In this context, it is important to remark that point contact conduction related to the formation of voltage-induced

atomic-size conduction channels have been already observed in late 90s during the breakdown of MOS devices.[339]

In the framework of memristive devices, evidence of the formation of QPCs were observed during the processes of both formation and annihilation of the conductive nanofilaments, in both VCM and ECM-type devices. Notably, quantum effects were observed in devices based on various switching materials, including mixed-ionic conductors,[340, 341] polymers[342, 343] and a wide range of metal oxides,[34, 333, 344-350] during both SET and RESET operations. Figure 27c reports an example of quantum effects in an ECM cell based on silver iodide (AgI) under current sweep stimulation during the SET process, showing discrete levels of conductance that be identified as multiples of the fundamental quantum of conductance.[351] Similarly, quantum steps were observed in memristive cells under different operational conditions including voltage sweep stimulation,[348, 350, 352] constant voltage stimulation[346] and pulse operation mode.[341, 346, 353, 354] The difficulty of controlling the morphology of the filament at the atomic level leads to cell-to-cell and cycle-to-cycle variations in the device response that is usually analyzed through conductance-state histograms. In this context, conductance-state histogram extracted from experimental data have been reported to show peaks at integer multiples of G_0.[346] However, also half-integer multiples of G_0 have been widely reported in many memristive devices.[31, 34, 333, 334, 350] In this context, it is important to remark that only for *s*-electron metals the transmission probability leads to conductance channels close to integer multiples of G_0 and quantum steps in memristive devices are more properly described to be in the *quantum* rather than *quantized* regime of conductance.[31] Note also that quantum effects in memristive devices are not the result of pure electrical effects, as in conventional QPC systems, but are the result of structural modifications of the nanofilament due to movement of ions. Thus, quantization effects should be discerned from step-like changes

in the device conductance related to a limited number of accessible nanofilament atomic structures. For example, the fingerprint of the rearrangement of the nanofilament atomic structure was observed in single NW memristive devices with volatile resistive switching characteristics, where discrete steps in conductance (not multiples of G_0) have been observed during spontaneous dissolution of an Ag nanofilament to form spherical Ag nanoclusters (Figure 27d).[355]

A similar but even more mazy situation is found in self-organizing nanowire networks, where both quantum and criticality effects can take place. Nanowire networks are bottom-up devices where a complex emergent behavior is arising from the interplay between a multitude of ECM cells and their nonlinear interactions.[356] Thanks to their structural plasticity through reconnection and rewiring,[357] coupled with nonlinear dynamics and fading memory, such nanonetworks were recently proved to act as ideal devices for *in materia* implementations of brain-inspired computing paradigms, such as reservoir computing.[358-360] Nanowire networks are typically formed by random distributed nanowires characterized by electromigrating metallic cores and insulating shells (Ag and PVP polymer, or Ag and Ag sulfide/iodide are the most common choices), so that each contact region between the nanowires becomes an ECM cell. When electrically excited at macroscopic level, the collective interactions among the many memristive elements cause a continuous local redistribution of electrical potential and current.[361, 362] The system is thus characterized by metastable states, where several local metallic nanofilaments are concurrently bridged and broken.[363] The so-induced discrete changes in macroscopic conductivity are characterized by long-range spatio-temporal interactions over many scales, i.e. they are scale-invariant phenomena (the fingerprint of any critical system). Such systems are described by avalanches and the effective criticality or not can be established from mathematical analyses of exponents in power-law probability distributions.[364] Whatever if the nanowire network is quantitatively defined as critical

or it can be just represented as a percolating system, each nanowire junction can be in a quantum conductance regime, as for any standard ECM cell. As reported in Figure 27e, plotting the measured conductance versus the current compliance shows a similar power law dependence for standard ECM planar cells, as well as for single nanowire junctions and for complex nanowire networks. Furthermore, conductance plateau proportional to G_0 are experimentally found (Figure 27f) and attributed to macroscopic "winner takes all" conductive pathways, where the involved nanowire junctions are all set at G_0. Since such pathway is calculated by simulations as the lowest possible energy state in the network, the quantum conductance behavior in the network is expected to be much more stable respect to single nanowire junctions.[365]

As a matter of fact, while quantum levels can be exploited for multilevel data storage within a single memory cell,[366] quantum effects in memristive devices have been proposed for a wide range of applications spanning from emulation of synaptic functionalities and neuroinspired computing[346, 352] to logic applications,[22] plasmonic switches[367] and quantum metrology.[368] However, to achieve controllable and tailored quantum levels, as required for all these applications, further understanding on the relationship between quantum effects and involved switching materials, device configuration and operational conditions is necessary for a rational design of memristive devices working in the quantum regime.

6.4) Exploiting device non-ideality

Stochasticity is ubiquitous in the world around us. However, stochasticity in devices is seldom exploited in computation. Conventional digital logic computing is deterministic. The goal is to get rid of any variability or non-deterministic behavior and compute with high precision. However, this is not an energy efficient approach. Neuromorphic computing is an emerging paradigm inspired by the brain that can improve computational efficiency and computational density of next-

generation architectures. Devices including memristors, phase-change memory, ferroelectric transistors, magnetic tunnel junctions and organic materials-based devices are being actively considered to build neuromorphic architectures. These emerging devices have varying levels of integration capability with CMOS, and the primary candidate is yet to be picked.

Memristors are noisy and this hinders their widespread adoption. However, to utilize analog devices effectively, the noise needs to be embraced. The brain is a prime example of how noisy devices are utilized for cognition.[369] The brain is also the most efficient processing engine known. With ever increasing computational demands, it is needed to look to computational substrates that are energy efficient. This requires treating stochasticity as a feature, not a bug. Appropriate algorithms are required to exploit the noise in memristor devices for robust computation. Recent techniques show overcoming memristor device non-idealities like faulty devices, device-to-device variability, random telegraph noise and line resistance by ensemble averaging,[370] and incorporating noise during online training for better performance of large-scale memristor based neural networks.[215, 225, 371]

Current approaches to neuromorphic architectures focus primarily on a bottom-up approach to codesign, which does not effectively exploit the physics of the devices for computation. The challenge is that typically, low level device engineers are left with the burden of demonstrating the usability of their device. The theory, algorithm and architecture designers don't have to contend with unexplored device properties in the toolbox. Merely designing perfect devices is not sufficient to bring about a paradigm shift. AI-guided codesign can further alleviate the challenges of inter-disciplinary codesign and accelerate design.[372, 373] AI-enhanced codesign techniques can use the constraints from emerging devices to develop algorithmic solutions for a given application. This can not only accelerate design but also enable interactions between experts from different areas of

the microelectronics design stack including theory, algorithms, circuits, devices and materials. Creativity in hardware design will dominate future computing systems that leverage increasingly heterogeneous components. Probabilistic computing is an avenue to leverage neuromorphic devices.[372] Multiple applications can benefit from probabilistic computing including modeling complex problems like nuclear and high-energy physics events, complex biological systems, precise climate models, large-scale neuromorphic applications, and AI algorithms.

6.5) Dynamic memristors for spiking neural networks

To maximize the system efficiency, it is worth taking a look at the whole stack of the most efficient computing system known to date – the brain, and compare it with existing digital computing systems. As pointed out in a recent review article,[374] although both systems have similar stacks (*e.g.* systems – organs – cells – molecules in biological systems, and systems – circuits – devices – materials in computing systems), in biological systems every stack can perform computing *natively* utilizing the underlying molecular/ionic/network dynamics (yellow regions), while in computing systems the devices and circuits are largely static (grey regions) and functions are only implemented through programming at the system level (Figure 28). The ability to directly process information with internal dynamics, as in biological systems, will not only lead to more efficient computing hardware, but can also stimulate algorithm developments to leverage full stack computing based on this paradigm.

Recent developments on emerging devices, such as memristive devices, have indeed revealed interesting internal dynamics that make this approach feasible.[5, 9, 180, 181, 183, 205, 210, 225, 249, 278, 375-383] In fact, *by definition* a memristive device is controlled by a set of time-dependent differential equations,[180, 381, 384-388] whereas the device state evolves based on the current and previous inputs to the device. For example, in an oxide-based memristor, the device conductance is determined by

the internal distribution of oxygen vacancies (V_{OS}) after going through oxidation, migration and reduction processes. These processes are also affected by the internal temperature and local electric field, which in turn depend on the device conductance.[387] The coupled electronic, ionic and thermal processes, each evolving on its characteristic energy and time scales, can potentially provide the desired internal dynamics that enable information to be processed natively at the materials and devices level, and up to the system level stack.

This concept has been tested in memristor devices with multiple state-variables that evolve at different time scales, termed "2nd-order memristors",[379, 384, 385] and memristors that offer short-term memory properties. In a 2nd-order memristor, the activation (e.g. due to the application of an input voltage pulse) and the spontaneous decay of one state variable (i.e. local temperature) effectively create an internal timing mechanism and in turn affect how the conductance, which is controlled by another state variable (i.e. V_O distribution profile) respond to an input spike train.[385] Utilizing these internal dynamics, common learning rules such as spike-timing dependent plasticity (STDP) can be natively implemented using identical, non-overlapping spikes, driven naturally by the temporal pattern of the inputs.[385, 386] Preliminary studies have shown that networks based on these devices can be used to directly process temporal data to uncover the spatiotemporal patterns embedded in the spike trains.[389]

Devices that show only short-term memory (STM) effects can also potentially be useful in processing temporal data, in particular, in the "reservoir computing" (RC) system concept where memristors with STM effects act as nodes in the network.[390, 391] Here the reservoir effectively maps the streaming inputs (in the form of pulse trains with identical pulse shapes and information coded in the timing of the pulses) into different memristor states. A second network layer was then used to analyze the excited reservoir state (formed by the collective memristor states) and perform

functions such as classification and signal reconstruction.[390, 391] The ability to directly process information at the device level leaves plenty of room for imagination on how a system can be built to effectively utilize these devices. One vision is to implement a system comprising 2nd-order memristors and STM memristors in a hierarchical manner, inspired by the columnar structure in the neocortex.[392-394] The system can be directly integrated with sensors that produce spatiotemporal data, such as neural probes, touch sensors, or an event-based camera.[395-398] The memristors can directly receive the input spike trains from the sensors, and natively process spatiotemporal information hidden in the spike trains for tasks such as object detection or motor movement control.[399] Compared with conventional approaches that need to store and accumulate the sensor inputs, directly processing the streaming spikes can lead to faster response, lower power and more robust noise tolerance.

An early example of this approach, termed columnar learning network (CLN), has recently been analyzed for multi-sensory input recognition tasks. is inspired by the mammalian cortex that consists of cortical columns (Figure 29a) formed by pyramidal neurons (Figure 29b) and diverse dendritic connections.[392, 393, 400-402] Compared to networks with point neuron models, the different dendritic connections help encode spatio- and temporal-information received by different regions of a neuron, and allow neurons in the network to interact with each other to generate more robust outputs.[403-406] Particularly, feedforward, lateral and feedback connections (shown as different arrows in Figure 29c, d) are implemented through proximal basal dendrite (PBD), distal basal dendrite (DBD) and apical dendrite (AD) connections, respectively. The feedforward PBD connections extract local features from the input spikes, while the lateral DBD and the feedback AD connections help predict the next possible features based on the historical sequence tendency.[400] CLN has been shown (through simulation)[400] to enhance not only classification

accuracy, but also noise robustness. The absence of supervision during training also leads to much lower power consumption. The same network structure developed for the visual sensor was found to also perform well for audio classification tasks, and a network consisting of two identical sub-networks can be used to process multi-sensory inputs with improved classification accuracy and noise tolerance.[400]

7. Future perspective on memristive technologies

Extensive research has been conducted at the material level to develop memristive technology by improving conventional non-volatile memories and developing devices with innovative operating principles. Despite the advantages of computing-in-memory using memristive devices for parallel computation tasks, the reliable production of devices and systems that meet the requirements of commercially available processors remains challenging.[6, 7, 407] To date, no memristive device has been identified as a clear winner to replace CMOS-based electronics. To overcome the challenges of integrating memristive devices into commercial IC technologies, comprehensive investigations, including device optimization, algorithm-architectural studies, hardware-software codesign, and system-level design, are necessary (Figure 30a).

Among the many classes of memristive devices, ReRAM and PCM have reached the most mature development. Redox-based resistive switching memory is considered the most relevant for CIM applications due to its fast switching, low power consumption, and high scalability. However, device variability resulting from the inherent stochastic resistive switching mechanism poses a significant challenge. Despite the four-decade history of PCM as a storage medium and recent integration as MVM accelerators, power consumption, switching speed, and CMOS compatibility remain significant hurdles. Further studies on integrating other types of devices, such as

ferroelectric memories, ion-intercalation resistors, and MTJs, into crossbar arrays for CIM are necessary.

The primary challenge associated with memristive devices is from the limited control over ionic movement, which governs the switching characteristics. For decades, CMOS technologies have achieved an extremely high level of controllability in charge transport of electrons and holes, while ion-based electronic devices have seldom been incorporated into IC technologies. Therefore, more studies on ionic transport, as well as the associated redox reactions and rearrangements of composition in solid-state electronics, are necessary to advance analog memory technology. In that sense, memristive technology could be viewed as the starting point for the next chapter in the development of electronic devices.

While there are challenges associated with using analog memory-based crosspoint array devices, recent algorithm and architectural studies have shown promising results in overcoming non-idealities and achieving high-performance on-chip training. However, system-level designs must also be carefully considered to balance gains in performance with increased costs in chip area and energy efficiency. Further research is needed to fully realize the potential of analog memory-based CIM chips, and simulation frameworks are being developed to aid in this effort. Ultimately, the goal is to develop algorithms and architectural designs that can fully leverage the parallelism of these devices and fill the gap between device-level requirements and current technology offerings, leading to the widespread use of analog memory-based chips for CIM (Figure 30a).

In the future development for CIM chips, the scalability and extendibility issues should be addressed to enable real-world application of CIM technology. Scalability means that the memristor capacity on chip should be increased to the 1 Gb level in order to match mainstream AI models. This goal requires advancing integration technology to more advanced nodes in addition

to increasing the memristor resistance to avoid high current issues on the bitlines. Extendability means that memristive chips should incorporate more types and more complex AI models, including CNN, RNN, and LSTM (Figure 30b). In this case, the overhead of ADC/DAC and data buffers should be minimized, and some general-purpose digital circuits, such as special function units, RISC-V processor units, and reconfigurable computing units, should be embedded into memristive chips.

On the other hand, for certain applications, the requirements may not be as strict as those of CIM chips. For instance, instead of using a large-sized chip for general purposes, a small-sized analog memory array can be employed to perform small-sized MVM or kernel operations, while still maintaining higher computing efficiency than conventional CMOS circuits. While these specialized arrays may have limited functionality compared to general purpose chips, they can still execute critical tasks when integrated with other devices. In addition, the integration of biosensors and memristive devices for cognitive computing of biosignals can provide possibilities for on-skin computing and diagnosis. In addition, the random nature of memristive devices can be exploited in some applications, such as encryption techniques and stochastic computing, where device variability is not a flaw but rather an attribute that can be leveraged (Figure 30b). Overall, further improvements in memristive devices, algorithms, architectures, and systems will enable memristive techonology to reach its full potential in next-generation computing.

Vocabulary

Memristor, A non-linear passive component characterized by a pinched hysteresis loop in the voltage-current plane[408, 409]. Practically, the device 'remembers' its most recent resistance state even when power is removed, which is why it's referred to as a memristor, short for "memory resistor". In the context of this work, a memristive device is defined as any devices that exhibits memristive behavior, including two-terminal and three-terminal devices; **Compute-in-memory (CIM)**, a computer architecture in which computations such as multiplication and summation are carried out directly within the memory unit eliminating the need for external data transmission[410]; **Neuromorphic computing,** a branch of computer engineering where the structure and function are inspired by the brain. It involves the design of systems composed of artificial neurons and synapses, implemented using non-von Neumann architectures to mimic the computing mechanisms of biological brains[411, 412]; **Artificial neural network,** a computational model that is inspired by biological neural networks. It may consist of interconnected layers - namely input, hidden, and output layers - each containing a number of neurons, which collectively contribute to its ability to learn and process complex patterns of information[275]; **Potentiation and depression,** changes in weight resulting from repetitive stimuli (*e.g.,* voltage pulse). Typically, potentiation indicates an increase in conductance (or a decrease in resistance), representing a 'weight' increase, whereas depression signifies a decrease in conductance (or an increase in resistance), representing a 'weight' decrease. These opposite-direction weight updates emulate synaptic plasticity mechanisms found in biological neural systems[413]; **Linearity,** the property where the weight update (typically a change in resistance) is directly proportional to the applied stimulus (such as a voltage or current pulse)[413]; **Symmetry,** the property where the weight updates during potentiation

and depression exhibit a symmetrical behavior, typically with equal magnitudes but opposite directions[413];

Figures

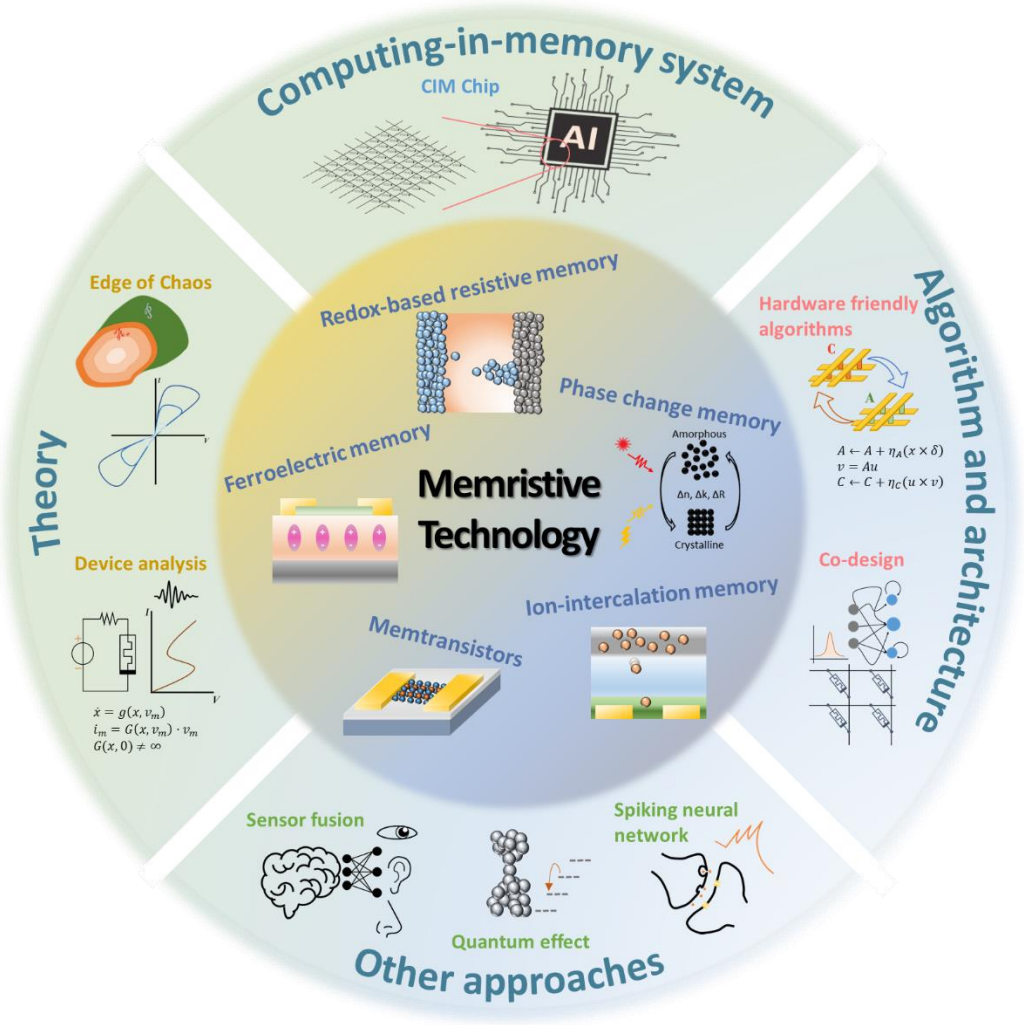

Figure 1. Development of memristive technology. Various memristive devices have been proposed including redox resistive switching memory, phase change memory, ferroelectric memory, ion-intercalation memory, and memtransistors. Four different categories of the research areas using such devices are presented in this review.

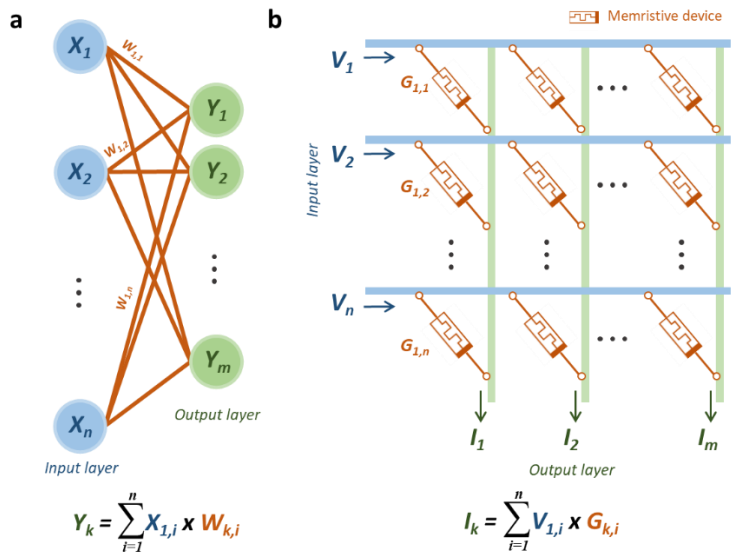

Figure 2. Artificial neural network (ANN) and hardware implementation using memristive devices of matrix-vector multiplication. (a) Single-layered perceptron model and (b) corresponding memristive crossbar array. Input, weight and output values in ANN are corresponding to input voltage, conductance of memristive devices, and output current, respectively.

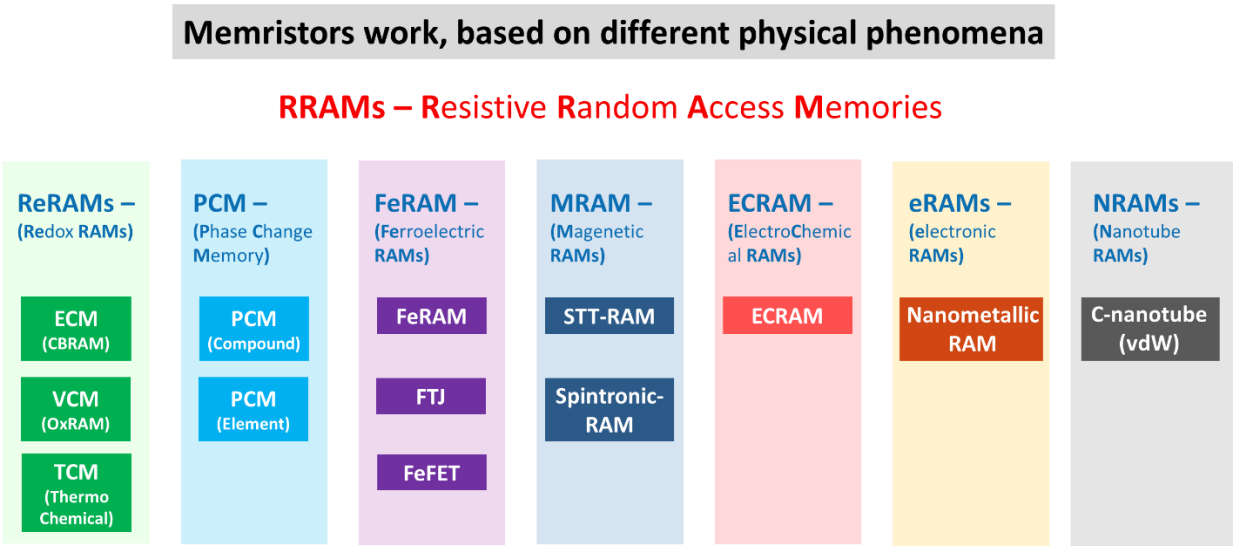

Figure 3. Memristive devices classified in accordance with different physical phenomena underlying their operational principle.

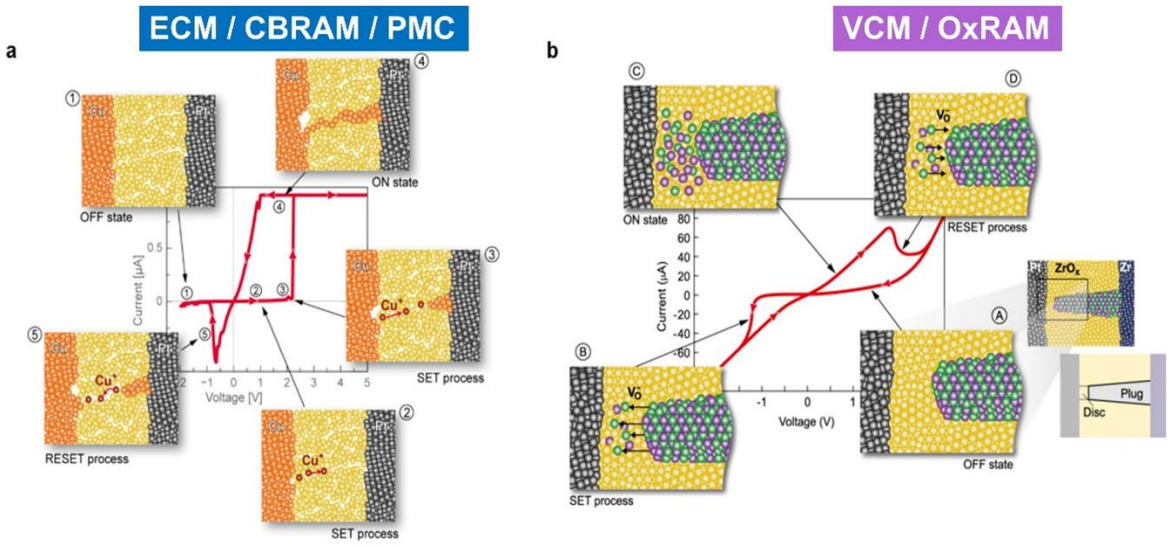

Figure 4. Different types of bipolar switching, depending on the predominant mobile species. (a) Electrochemical metallization memory and (b) Valence change memory. Main difference is that for the establishing of high resistive and low resistive states in VCM devices a formation and modulation of a Schottky barrier is essential. Reprinted with permission under a Creative Commons CC BY License from ref 31. Copyright 2022 John Wiley & Sons.

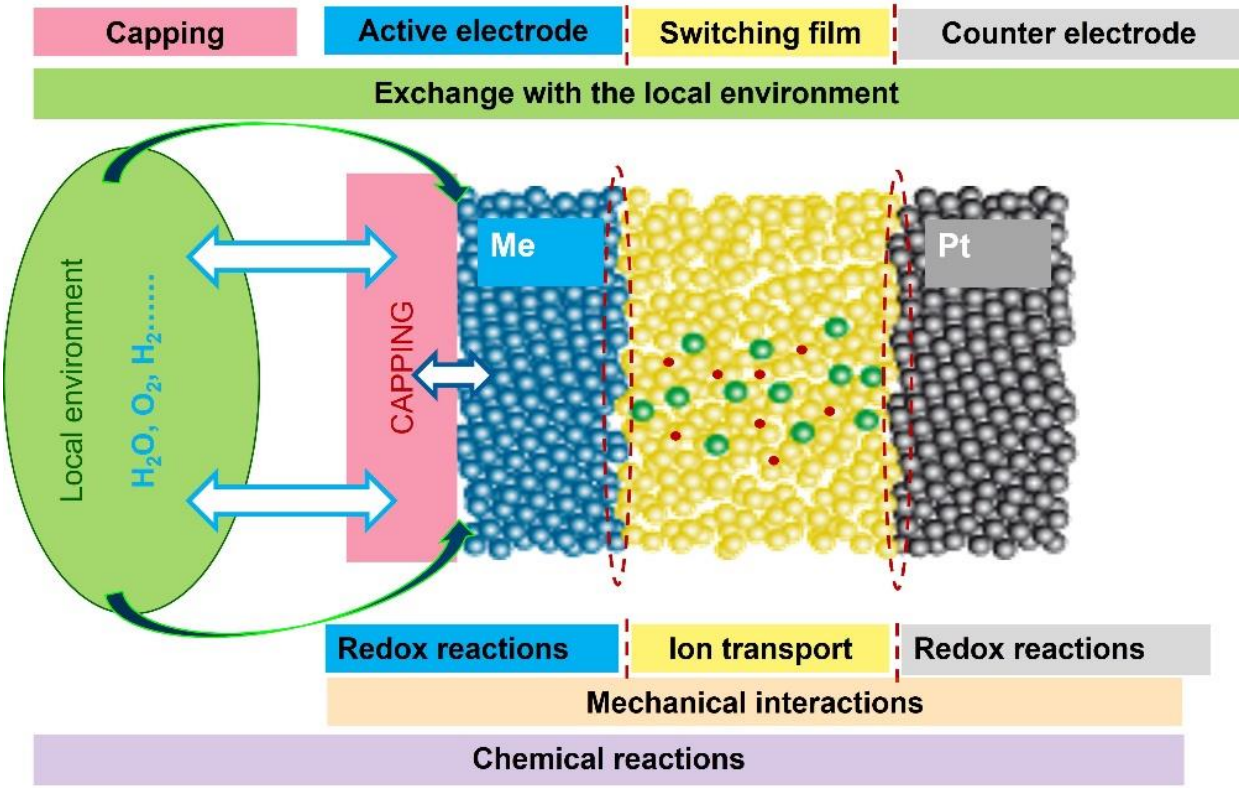

Figure 5. Structure of a memrsitive device and the processes and possible interactions between different components of the cell.

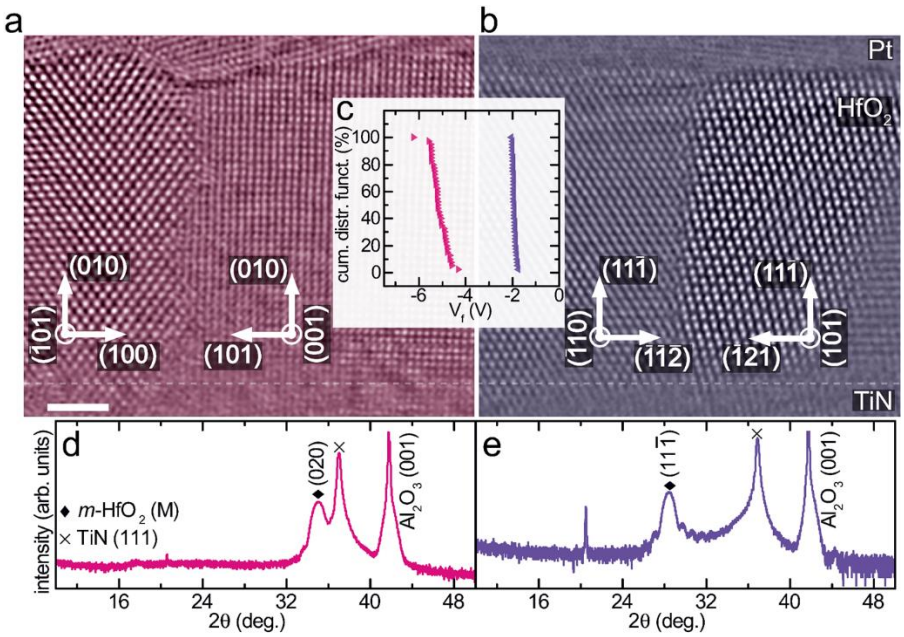

Figure 6. Cross-sectional high-angle annular dark-field (HAADF) scanning transmission electron microscopy (STEM) image of a metal-insulator-metal stack with (a) (010) and (b) (11$\bar{1}$) textured HfO_2. Scale bar is 2 nm. The TiN-HfO_2 interface is indicated by a dashed line. The X-ray diffraction (XRD) pattern reveals that a change in the growth temperature and rf-power for HfO_2 results in (d) (11$\bar{1}$), (purple) and (e) (020), (pink) orientation. (c) Devices with (010) textured HfO_2 (pink) have an increased average forming voltage of $\overline{V_f}$= -5.3 V compared to devices with (11$\bar{1}$) HfO_2 having $\overline{V_f}$= -1.9 V as shown by the cumulative distribution function measured from 50 30x30 μm^2 devices. Reprinted with permission under a Creative Commons CC BY License from ref 33. Copyright 2022 John Wiley & Sons.

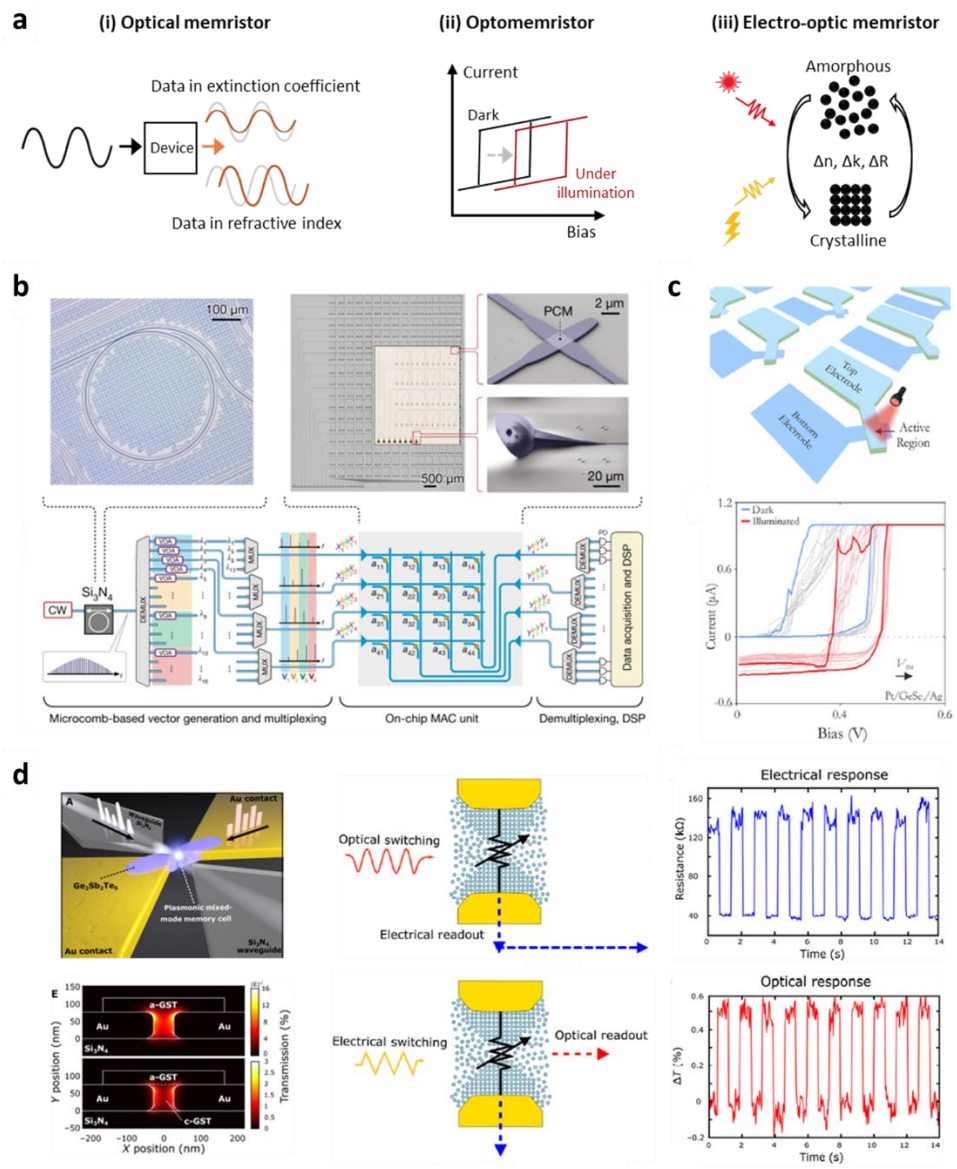

Figure 7. (a) Three categories of microscale and nanoscale integrated optical memory. (i) optical memristor. (ii) optomemristor. (iii) electro-optic memristor. (b) A photonic tensor core using PCM optical memristors for parallelized in-memory matrix-vector multiplication. Wavelength division multiplexing is achieved by integrating a Si_3N_4 microcomb. Reprinted with permission from ref 64. Copyright 2021 Springer Nature. (c) An optomemristor enabled by Pt/$GeSe_3$/Ag stack. This device is essentially an electrical memristor with its hysteresis controlled by optical illumination.

Reprinted with permission under a Creative Commons CC BY License from ref 65. Copyright 2022 Springer Nature. (d) An electro-optic memristor formed by integrating plasmonic nanogap with tapered dielectric waveguide. The electro-optic memristor features dual electrical-optical functionality, meaning the device can be programmed optically and read out electrically and vice versa. Reprinted with permission under a Creative Commons CC BY License from ref 66. Copyright 2023 American Association for the Advancement of Science.

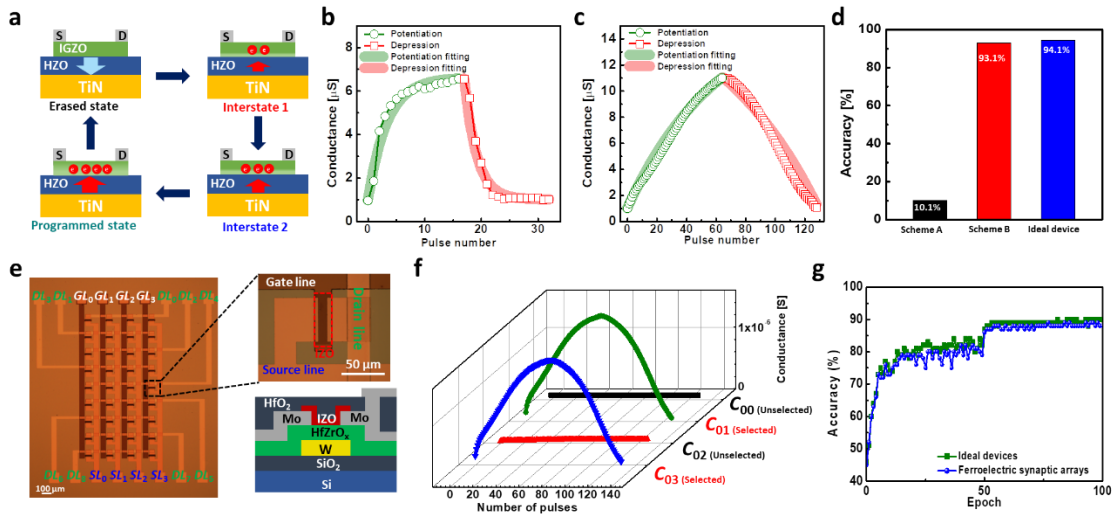

Figure 8. (a) Schematic illustration of conductance modulation mechanism in ferroelectric transistors. Reprinted with permission from ref 82. Copyright 2019 American Chemical Society. Conductance modulation characteristics of ferroelectric transistors using (b) identical and (c) incremental pulse schemes. (d) Classification accuracy of neural network based on the ferroelectric transistor and ideal synapses. Neural networks based on ferroelectric transistors showed higher classification accuracy when an incremental pulse scheme is used, compared to an identical pulse scheme. Reprinted with permission from ref 86. Copyright 2021 AIP Publishing LLC. (e) Optical image of ferroelectric synaptic transistor array (left). Optical image and schematic illustration of the ferroelectric transistor in the array (right). (f) Column-wise parallel weight update characteristics of the ferroelectric synaptic transistor array. The weight update was done simultaneously, but only the selected cells (C_{01} and C_{03}) were programmed, while the unselected cells (C_{00} and C_{02}) were kept in their original state. (g) Classification accuracy of neural networks based on ferroelectric transistor array and ideal synapse array. Reprinted with permission under a Creative Commons CC BY License from ref 91. Copyright 2022 American Association for the Advancement of Science.

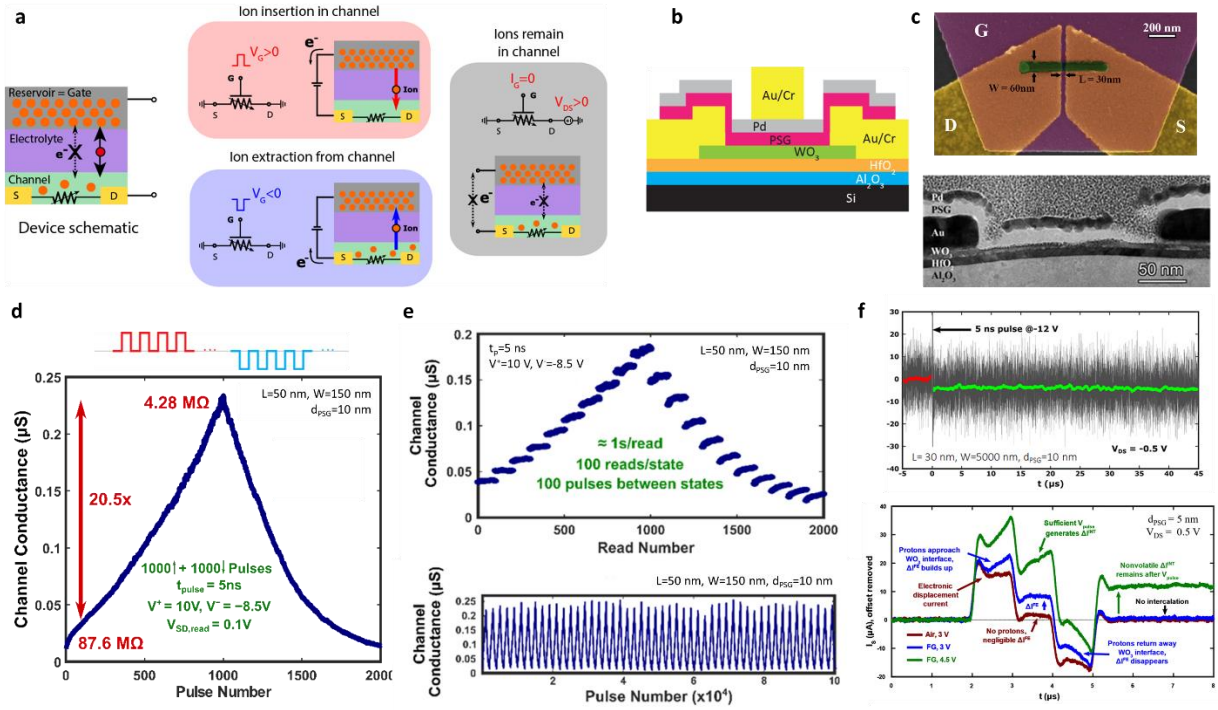

Figure 9. (a) Operation of a non-volatile ion-intercalation programmable resistor. Left: device structure featuring ion reservoir that also serves as gate, an electrolyte and a channel material where the ions behave as dopants. Center: device programming by ion insertion in (top) and extraction from (bottom) the channel. Right: Non-volatility when gate is open; ions remain in channel. Reprinted with permission from ref 111. Copyright 2022 IEEE. (b) Schematic diagram of Pd/PSG/WO_3 non-volatile protonic programmable resistor. (c) Top: False-color SEM micrograph of fabricated device. Bottom: TEM cross-section of fabricated device. (d) Modulation performance of a protonic device in response to a positive stream of 5 ns pulses followed by an identical negative stream. (e) Top: Conductance retention characteristics of protonic device during 100 s reads of the conductance state. Bottom: Endurance characterization of a protonic device displaying reproducible modulation over 10^5 pulses conducted over 30 hours. Reprinted with permission from ref 110. Copyright 2022 American Association for the Advancement of Science. (f) Top: Channel current before, during, and after a 5 ns pulse applied to a protonic device clearly displaying

impulse-like fast modulation characteristics. Bottom: Channel current under positive trapezoidal voltage pulse starting at t=2 μs with 1 μs rise time, 1 μs hold time and 1 μs fall time in air (red) and in forming gas under a low/high voltage pulse (blue/green). The experiment reveals the existence of transient displacement current, volatile field-effect current and non-volatile proton intercalation current. Reprinted with permission from ref 111. Copyright 2022 IEEE.

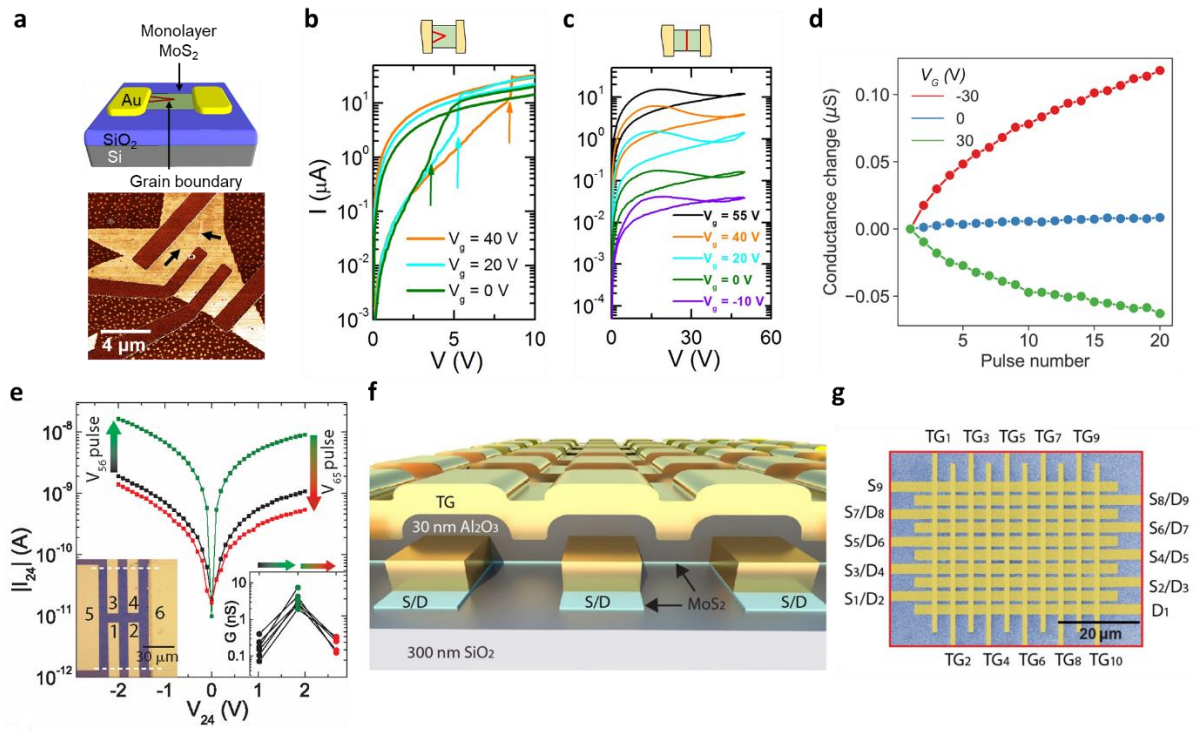

Figure 10. (a-c) Initial fundamental studies of memristive phenomena in 2D semiconducting MoS_2. (a) Schematic and atomic force microscopy image of a single-flake monolayer MoS_2 device with an intersecting grain boundary. (b) Gate-tunable memristive response in an intersecting grain boundary monolayer MoS_2 device. In this case, the set voltage between the high and low resistance states is controlled by the gate potential (V_g). (c) Gate-tunable memristive response in a bisecting grain boundary monolayer MoS_2 device. In this case, the current in both the high and low resistance states is modulated by the gate potential. Reprinted with permission from ref 132. Copyright 2015 Springer Nature. (d-g) Neuromorphic functionality in 2D memtransistors. (d) The unipolar synaptic response of 2D memtransistors can be tuned from potentiation to depression as a function of the gate bias (V_G). Reprinted with permission from ref 136. Copyright 2021 American Chemical Society. (e) Multi-terminal 2D memtransistors show heterosynaptic responses where the conductance state between two contacts (contacts 2 and 4) is modulated by voltage pulsing

between two orthogonal contacts (contacts 5 and 6). Reprinted with permission from ref 133. Copyright 2018 Springer Nature. (f) Schematic and (g) false-colored scanning electron micrograph of a dual-gated 2D memtransistor crossbar array. Reprinted with permission from ref 137. Copyright 2020 John Wiley & Sons.

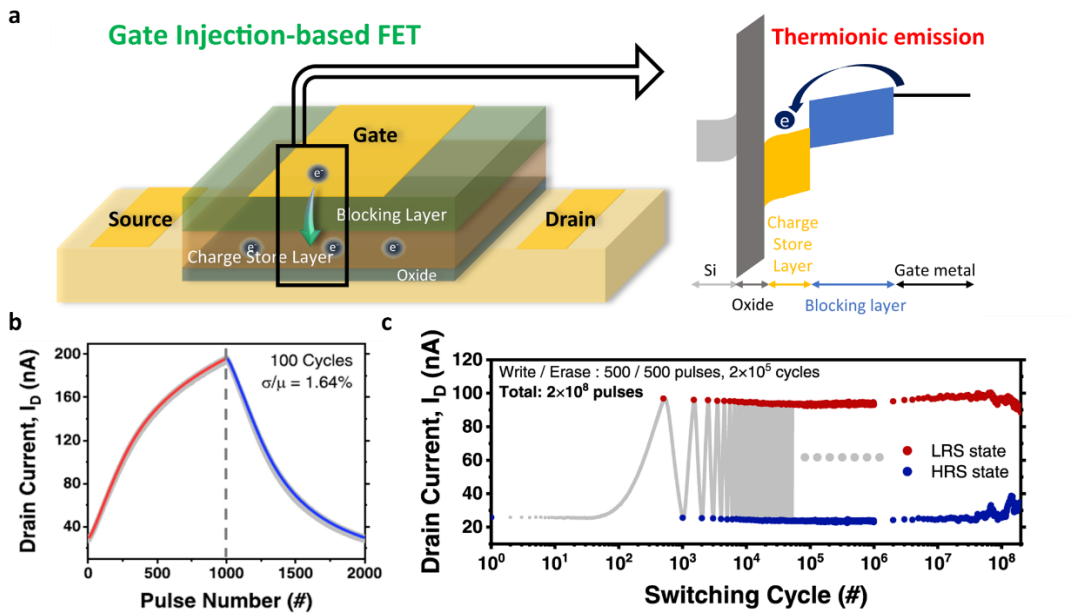

Figure 11. (a) Schematic and band diagram of GIFET. The charge is injected from or extracted to gate metal by thermionic emission. (b) Repeated LTP-LTD characteristics on a single GIFET device with 1000 potentiation–1000 depression (5 V/−3.3 V, 500 μs). (c) The endurance of GIFET over 2×10^5 switching cycles (2×10^8 pulses). Each switching cycle is composed of 500 write pulses with 5 V, 200 μs and 500 erase pulses with −5 V, 200 μs. Reprinted with permission under a Creative Commons CC BY License from ref 140. Copyright 2022 Springer Nature.

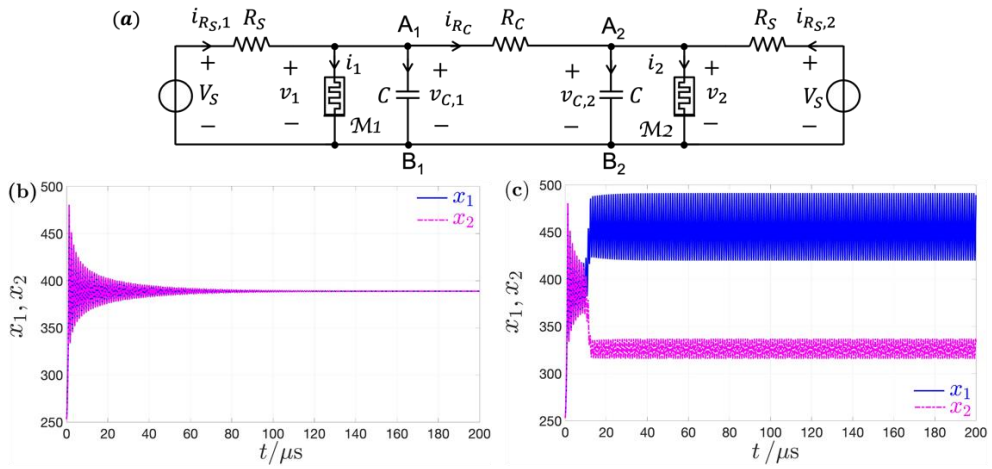

Figure 12. (a) The simplest ever-reported bio-inspired circuit,[148] which reproduces the counterintuitive diffusion-driven transition from quiescence to sustained oscillatory behavior, observed by Smale in an eight-order reaction-diffusion system from cellular biology,[149] through half the number of dynamical states, encoded in two capacitors and two locally-active NaMLab memristors,[146] on the proviso that the two identical memristive Pearson-Anson cells are preliminarily poised on some common Edge of Chaos operating point before the diffusion path, enabling their diffusive interaction, is activated. The four degrees of freedom of the proposed fourth-order RD-MCNN are encoded in the memristors' states x_1 and x_2, and in the capacitors' voltages v_1, and v_2 (refer to ref 148 for details). (b)-(c) Diffusion-induced transition of the bio-inspired array of (a) from silence to persistent oscillatory mode (all details on memristor model and cell parameters, chosen through a systematic methodology from Local Activity and Edge of Chaos Theory, are reported in ref 148). (b) Progressive approach of the states x_1 and x_2 of memristors M_1 and M_2 toward the same quiescent steady state, which they would independently attain, irrespective of the initial condition, in case they were let evolve in the respective uncoupled cell. Here $R_C = 50\,\Omega$, while the initial condition for the cell $i = 1(2)$ is arbitrarily set to $(x_i(0), v_{C,i}(0)) = (253.15,\ \zeta \cdot 0.1\ \text{mV})$. (c) Formation of one of the two admissible dynamic

patterns, across the homogeneous cellular medium, as determined by the choice of the initial conditions for the four RD-MCNN states, which are set as specified earlier for the simulation result in (b). Even though, here, $R_C = 25\ \Omega$, reducing the coupling resistance progressively from positive infinity, first the homogeneous solution of the array loses stability, allowing the two-cell array to enter a bi-stability regime with static pattern formation, out of a pitchfork supercritical bifurcation, for $R_C = R_{C,P} = 49.75\ \Omega$, and, later on, due to a Hopf supercritical bifurcation, the two admissible locally-stable inhomogeneous static solutions of the RD-MCNN become unstable, while, concurrently, two infinitesimal almost-sinusoidal locally-stable oscillations, increasing in amplitude as the coupling strength is further increased, are found to develop across them, for $R_C = R_{C,H} = 28.1\ \Omega$.[153] Reprinted with permission from ref 148. Copyright 2022 IEEE.

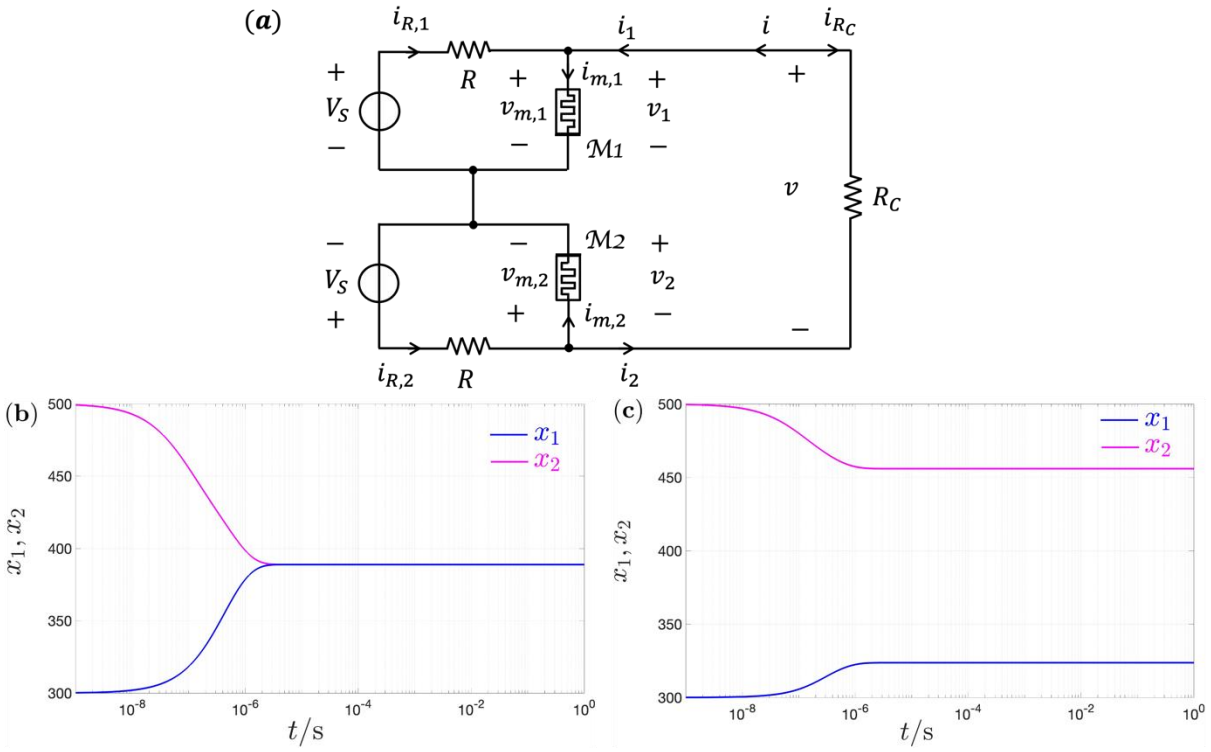

Figure 13. (a) The simplest ever-proposed reaction-diffusion system,[156] which captures the same unexpected dissipation-induced static pattern formation, appearing, together with the destabilization of the homogeneous solution, in Turing-based models,[155] through half the number of dynamical states, encoded in two niobium oxide-based threshold switches from NaMLab,[146] providing its two identical memristive cells are preliminarily poised on some common locally-active and asymptotically-stable operating point, before the inclusion of a linear resistor of resistance R_C between the memristors' positive terminals turns the diffusion process on. (b)-(c) Symmetry-breaking effects in the bio-inspired two-cell array of (a) (all details on memristor model and cell parameters, chosen through a systematic methodology from Local Activity and Edge of Chaos Theory, are reported in ref 156). (b) Convergence of the states x_1 and x_2 of memristors M_1 and M_2 toward the same asymptotic level, which they would independently approach, irrespective of the initial condition, if they were let evolve in the respective uncoupled cell. Here $R_C = 100\ \Omega$,

while the initial condition for the memristor M_1 (M_2) is arbitrarily set to $x_1(0) = 300$ ($x_1(0) = 500$). (c) Progressive formation across the homogeneous cellular medium of one of the two admissible Turing patterns, as dictated by the choice of the initial conditions for the two RD-MCNN states, which are set as specified earlier for the simulation result in (b). While, here, $R_C = 25\ \Omega$, reducing the coupling resistance progressively from positive infinity, the homogeneous solution of the two-cell array actually loses stability, and static pattern formation first occurs, due to a supercritical pitchfork bifurcation, for $R_C = R_{C,P} = 49.75\ \Omega$[156] (see ref 157 for insights on static pattern formation in large RD-MCNN arrays). Reprinted with permission from ref 156.

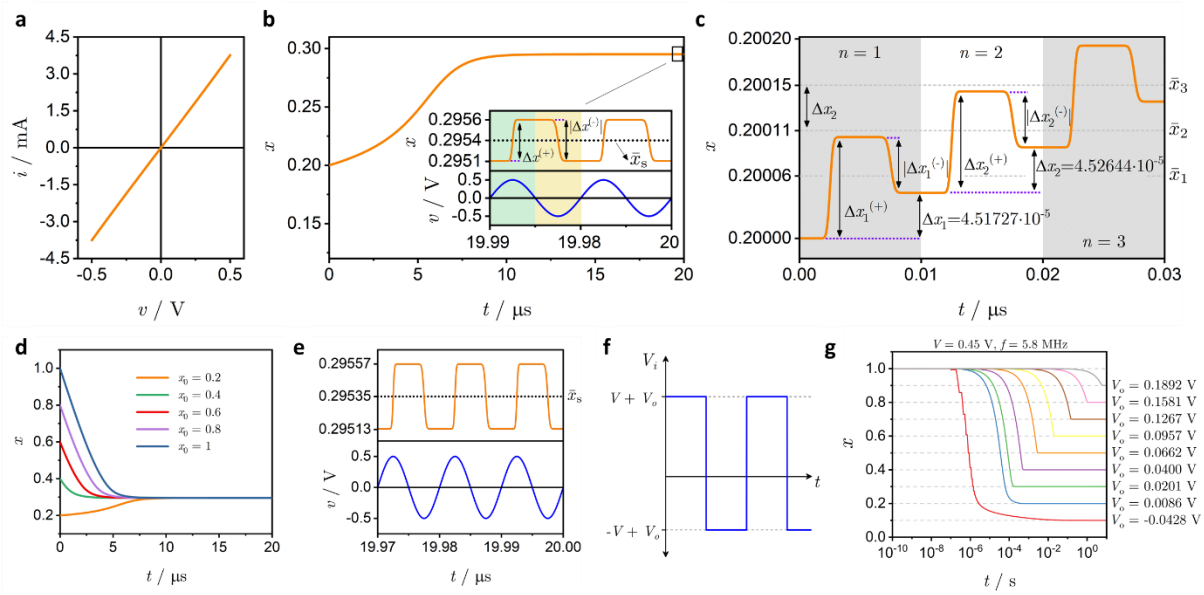

Figure 14. (a)-(c) Response of the TaO_x-based memristor from HP Labs in the high frequency limit. (a) Steady-state $i - v$ plot of the TaO_x-based memristor model,[160] under an AC voltage input of the form $v = (0.5\,V)\,sin(2\pi(100\,MHz)t)$. (b) Time-waveform of the device state variable x, preliminarily initialized at 0.2 ($x_0 = 0.2$), responding to the same periodic voltage input. The inset provides a detailed view of the x vs t response over the last two input cycles of the simulated time-response. (c) Zoomed-in view of the x vs t response in (b) over the first three input cycles. Reprinted with permission from ref 158. Copyright 2022 IEEE. (d)-(e) Fading memory in the TaO_x-based memristor model in the high frequency limit. (d) The time-waveforms of the state variable x, for the initial conditions defined in the set $x_0 \in \{0.2, 0.4, 0.6, 0.8, 1\}$. Each of the depicted time-responses was induced by a zero-mean sinusoidal voltage input with amplitude $\hat{v} = 0.5\,V$ and frequency $f = 100\,MHz$. At some finite time, all x vs t responses converge to the same steady-state oscillation, as illustrated in panel (e). (f)-(g) TaO_x memristor state tuning with high-frequency bipolar square-wave periodic voltage inputs (see exemplary plot in panel (f)). The illustrated x vs t plots in (g) were obtained by stimulating the nanodevice with a bipolar square-

wave periodic voltage input with amplitude $\hat{v} = V = 0.45\ V$ and frequency $f = 5.8\ MHz$, for the set of DC offset values V_o listed in the figure legend. Each V_o value was calculated by solving (6) with respect to $\bar{x}_s$, for $\bar{x}_s \in \{0.1,\ 0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9\}$. The above-mentioned frequency was chosen such so as to induce an infinitesimally-small periodic oscillation in the state x about $\bar{x}_s$, in each of the simulated x vs t responses. A mathematical method for determining the frequency value that characterizes an input as a high-frequency periodic signal, based on the definition provided towards the beginning of this section, can be found in ref 158. Reprinted with permission from ref 158. Copyright 2022 IEEE.

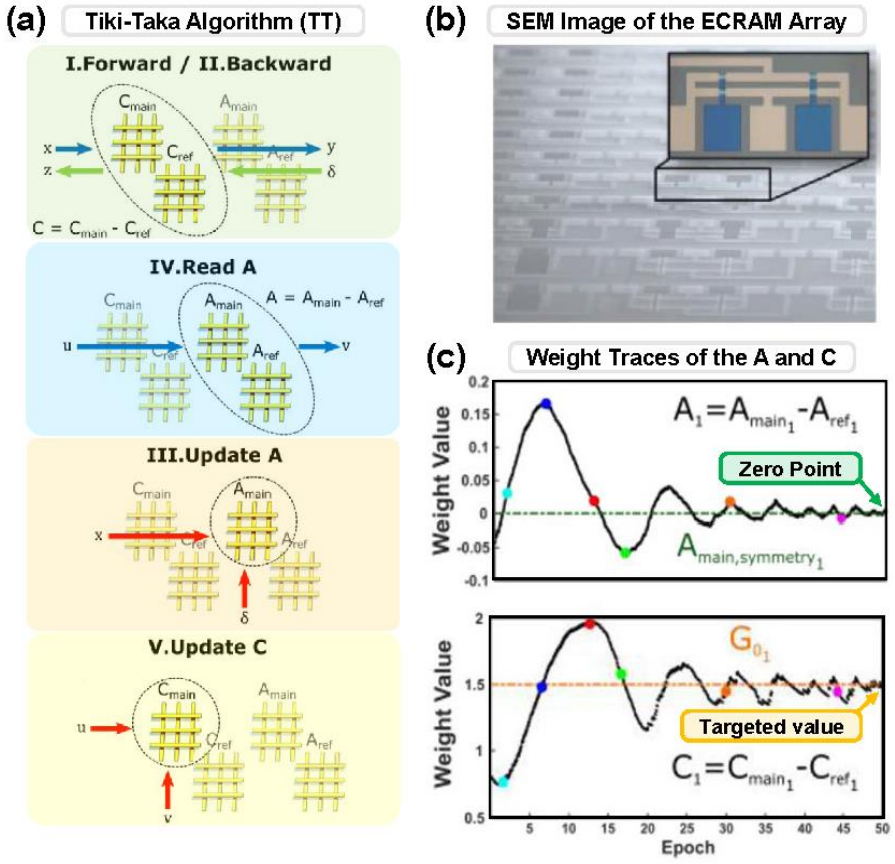

Figure 15. The hardware demonstration of the Tiki-Taka algorithm (TT) using the ECRAM array. (a) Schematic illustration of TT algorithm. Unlike SGD, the weight update is done on the auxiliary array, A, instead of the core array, C. Weight values in A are then transferred to C on a regular frequency. (b) SEM image of 2x2 ECRAM arrays. (c) Experimentally measured weight traces of A (top) and C (bottom) during the linear regression demonstration, converging to the zero point and targeted value as expected, respectively.[195] Reprinted with permission under a Creative Commons CC BY License from ref 195. Copyright 2022 IEEE.

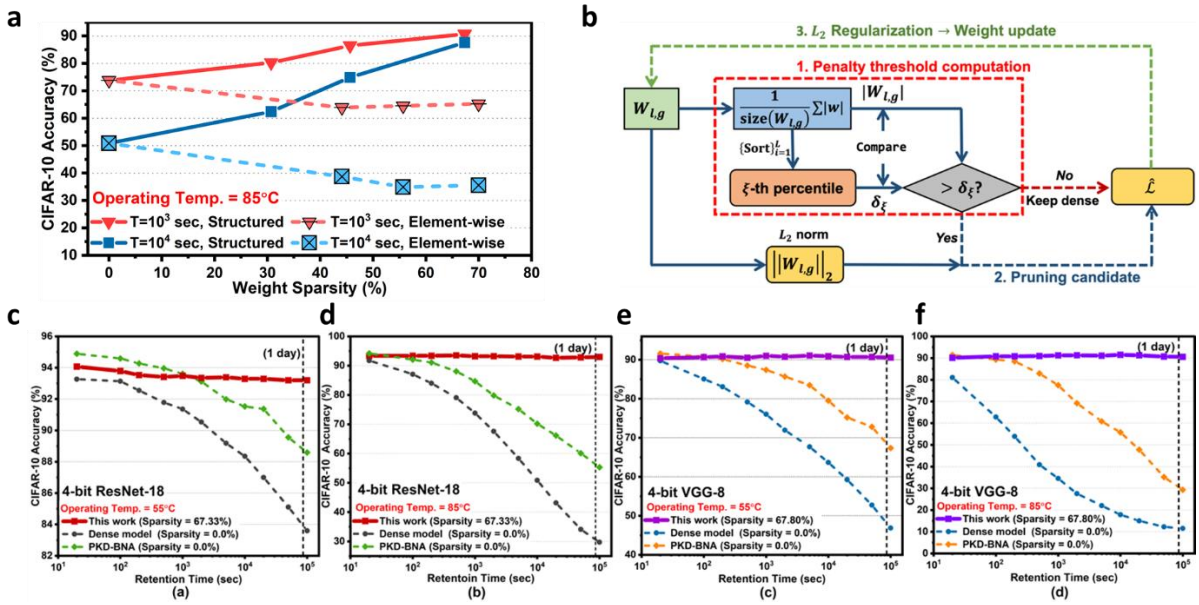

Figure 16. (a) NeuroSim-based hardware inference accuracy of 4-bit ResNet-18 with different fine-grained and coarse-grained sparsity. (b) Overview of the proposed robustness-aware structured pruning algorithm. (c-f) Mixed hardware-software experiment results of 4-bit DNN models on CIFAR-10 dataset with static 55ºC and 85 ºC operating temperature: (c-d) 4-bit ResNet-18 inference accuracy. (e-f) 4-bit VGG-8 inference accuracy. Reprinted with permission from ref 205. Copyright 2022 IEEE.

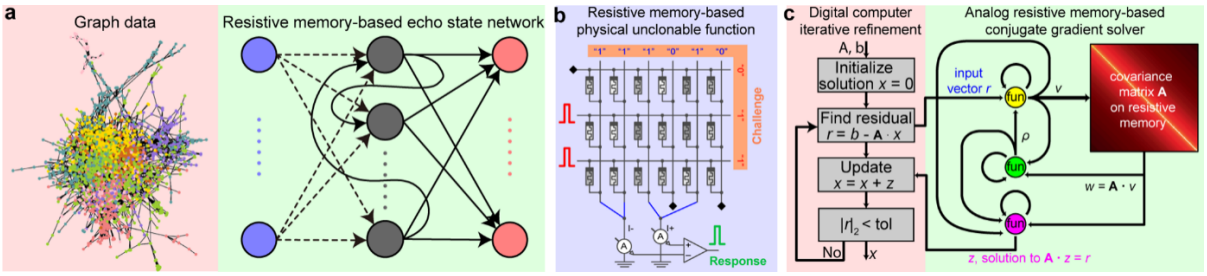

Figure 17. Hardware-software codesign for resistive memory-based analog computing. (a) Schematic illustration of the architecture of a resistive memory-based echo state graph neural network handling CORA dataset.[212] (b) Schematic illustration of the analog random resistive memory-based physical unclonable function for authentication.[216] (c) Schematic illustration of the iterative refinement algorithm with conjugate gradient solver implemented on the hybrid analog-digital computing system for solving systems of linear equations.[219] Reprinted with permission from ref 50. Copyright 2023 and 2018 Springer Nature, respectively.

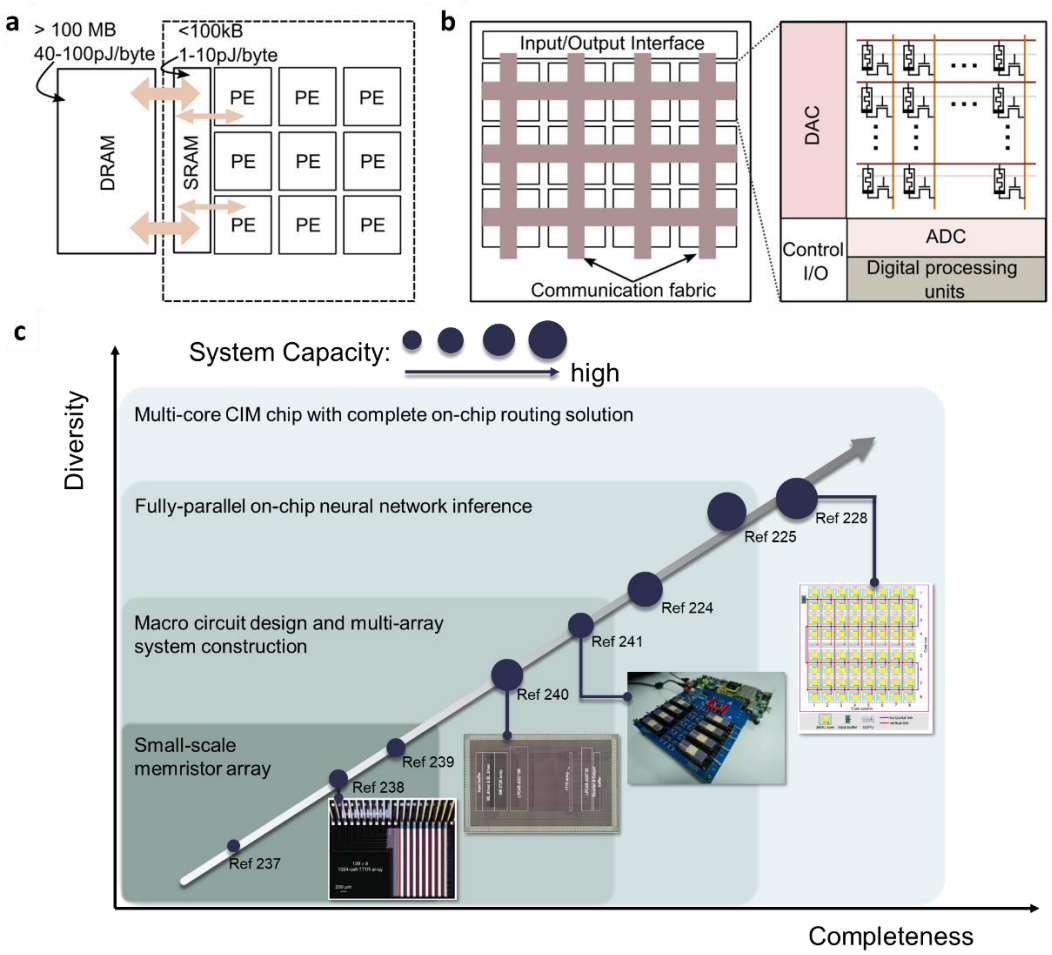

Figure 18. (a) The energy efficiency of conventional digital accelerators for DNN acceleration tends to plateau at hundreds of fJ/operation due to the energy costs associated with shuttling around the synaptic weights. (b) Analog in-memory computing with fixed-synaptic weights could address this challenge. Memristive devices with potential for ultra-dense integration density are promising candidates for AIMC. (c) The develop trend of system completeness and diversity of the CIM systems. Reprinted with permission under a Creative Commons CC BY License from ref 238. Copyright 2017 Springer Nature. Reprinted with permission from ref 240, 241. Copyright 2020 IEEE and 2020 Springer Nature, respectively.

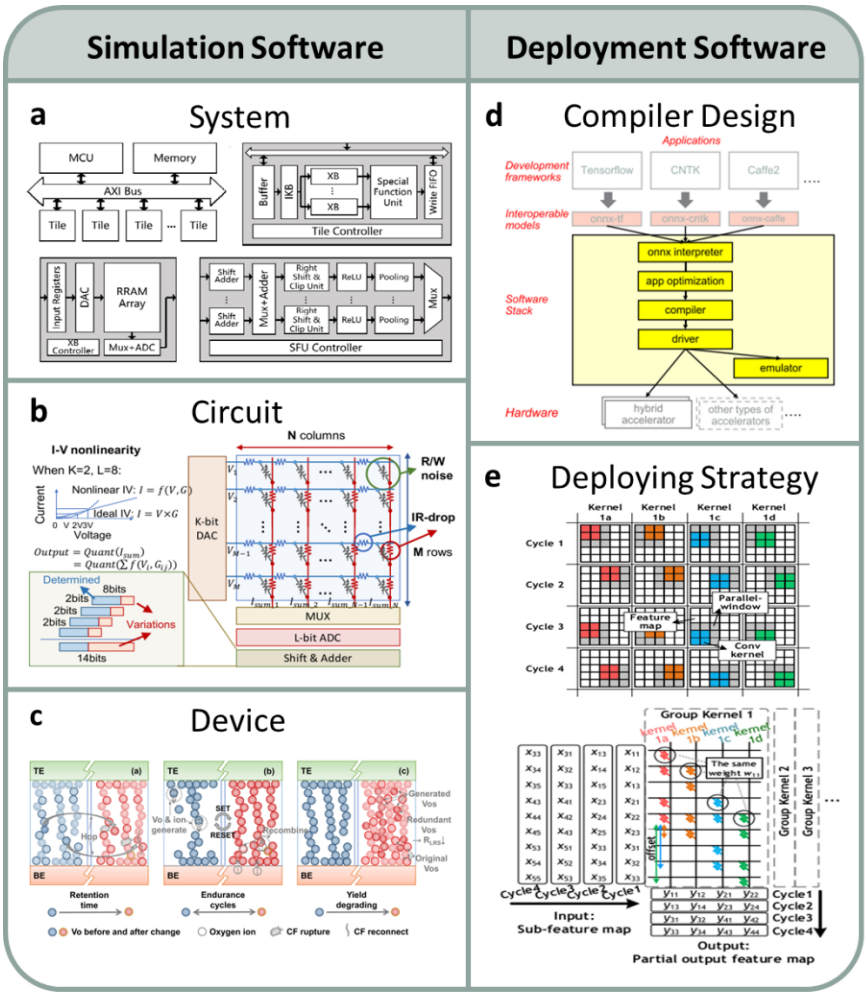

Figure 19. Software tools of CIM systems. (a) The typical hierarchy of the CIM architecture in the simulator. Reprinted with permission from ref 242. Copyright 2021 IEEE. (b) The non-ideal factors and functional modules of the memristor-based macro circuit. Reprinted with permission from ref 244. (c) The schematics of physical mechanisms of reliability degradation in the memristor device. Reprinted with permission from ref 245. Copyright 2021 IEEE. (d) The compilation flow of the CIM hardware. Reprinted with permission from ref 247. Copyright 2018 IEEE. (e) The deploying strategy for improving throughput. Reprinted with permission from ref 248. Copyright 2021 IEEE.

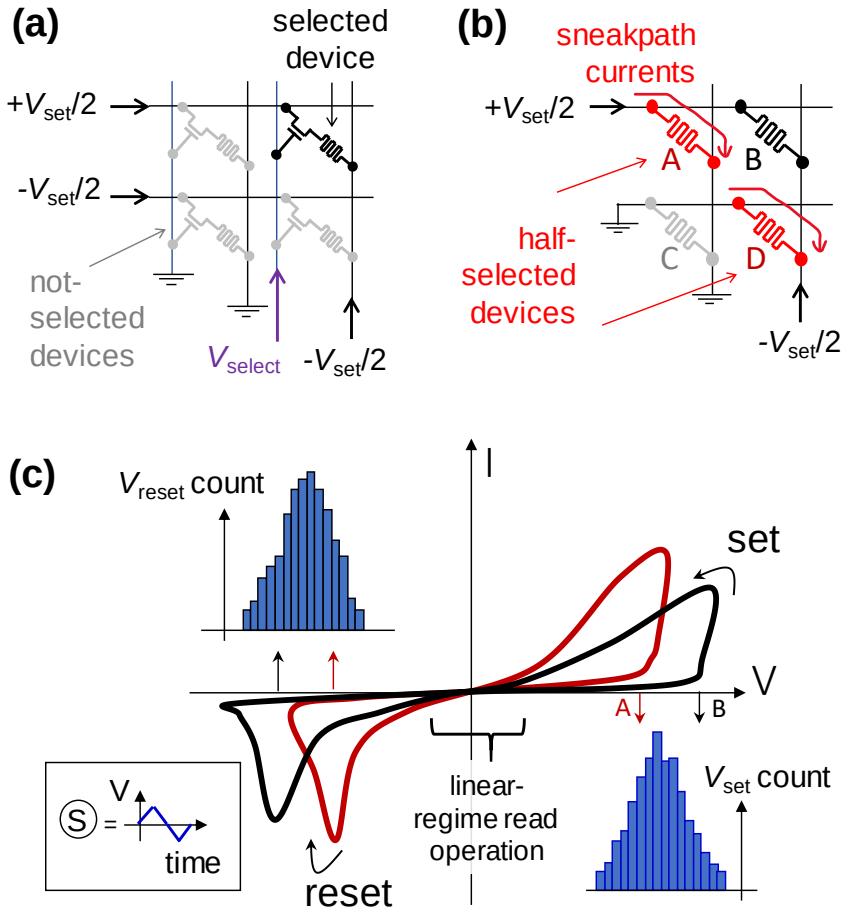

Figure 20. (a,b) A typical half-biasing scheme for applying write voltages to the selected memristor in (a) active ("1T1R") and (b) passive ("0T1R") crossbar circuits. For clarity, both panels show a four-device crossbar circuits. In 0T1R arrays, a half of the voltage applied to the crossbar array to write the selected ("B") device is dropped across the half-selected ("A" and "D") devices that share the same horizontal and vertical electrodes with the selected device. In 1T1R circuits, the select transistors and additional lines controlling them allow applying a nonzero voltage across the selected device only. (c) Typical I-V hysteretic curves for metal-oxide memristors for the symmetric voltage sweep (lower bottom inset). The inset histograms show schematically device-to-device variations in switching voltages at which conductance changes by more than a certain amount. (Note that there is no sharp threshold for switching; in principle, the

switching thresholds would depend on the voltage ramp rate. In that sense, the thresholds are instead “effective” values representative of the intended device usage.) To highlight half-select disturbance problem, the black I-V curve corresponds to the selected device “B” with a switching threshold at the higher end of the distribution. Setting such device can disturb the half-selected device “A” featuring red I-V curve with switching threshold at the lower end of the distribution. Reprinted with permission from ref 251, 274. Copyright 2018 Springer Nature.

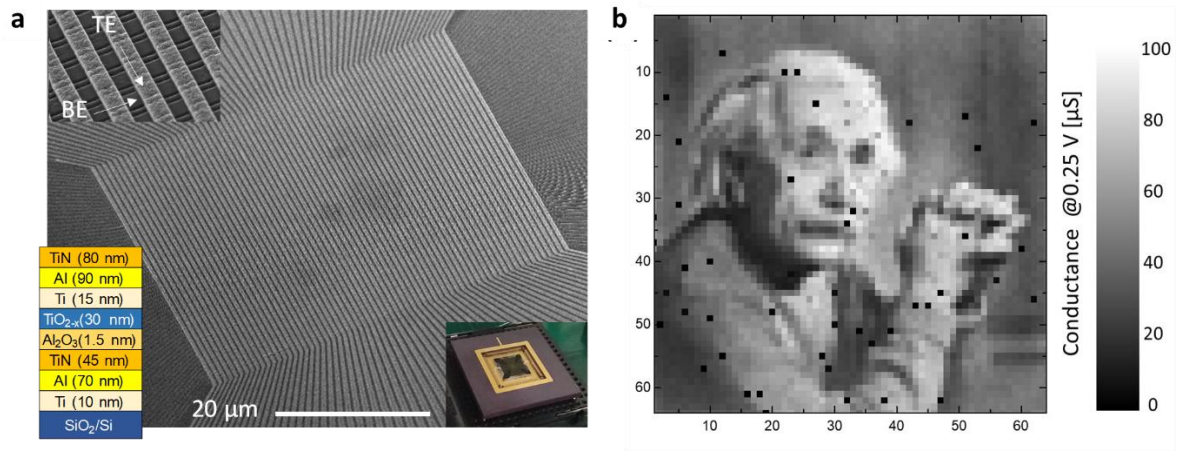

Figure 21. (a) Scanning electron microscope image of 64×64 crossbar array with passively integrated Al_2O_3 / TiO_{2-x} memristive devices. The top left inset shows a zoomed-in portion of the crossbar array. The bottom left and bottom right insets show material layers at the device cross-section with corresponding thicknesses in nanometers and the packaged chip. (b) The results of tuning crossbar memristors' conductance to the target values corresponding to gray-scale quantized image Einstein using a write-verify automated algorithm with a 5% relative error, defined as an absolute difference between target and desired conductances, normalized to the desired conductance. Black squares show devices that could not be switched with the chosen maximum write voltages. Reprinted with permission under a Creative Commons CC BY License from ref 251. Copyright 2021 Springer Nature.

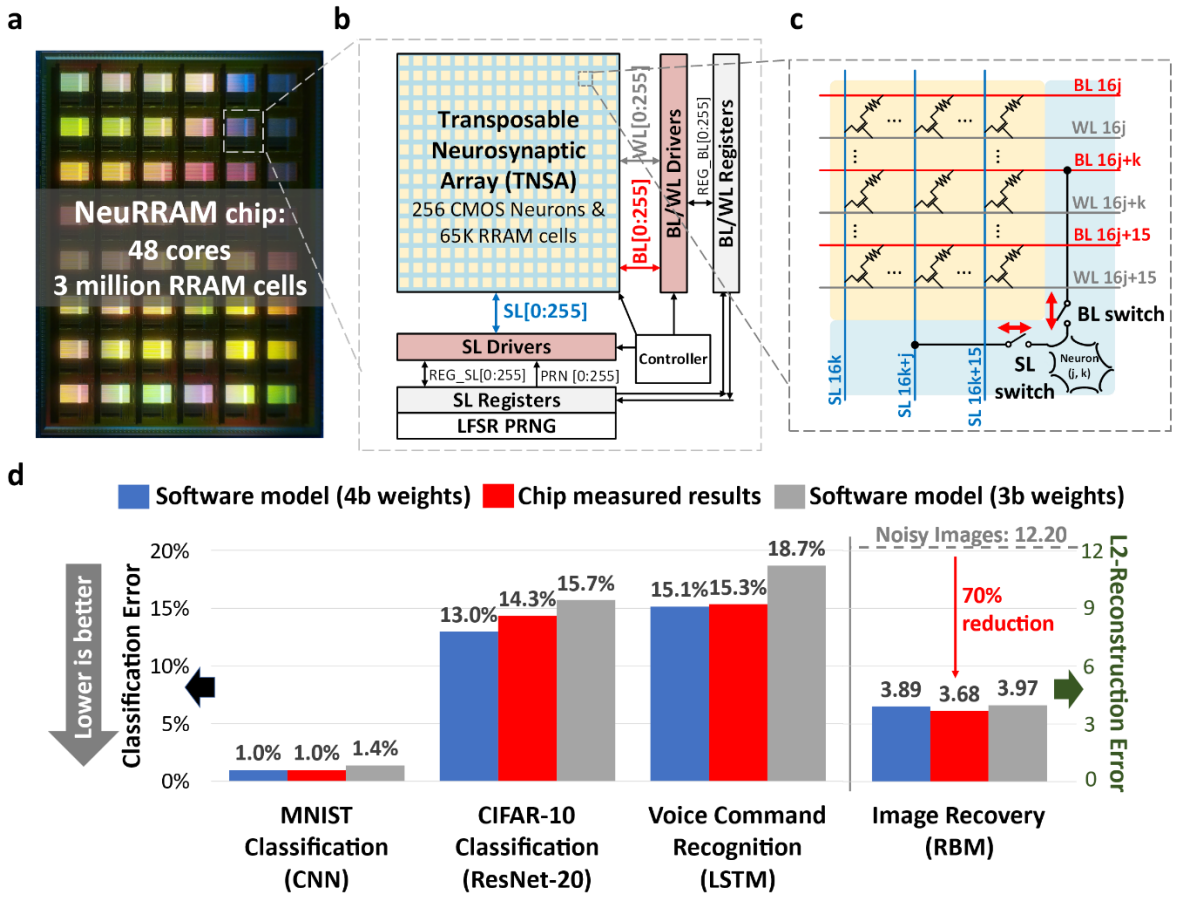

Figure 22. (a) NeuRRAM chip micrograph (b) NeuRRAM core block diagram. (c) Transposable Neurosynaptic Array architecture physically interleaves RRAM weights and CMOS neurons. (d) Measured inference accuracies on NeuRRAM are comparable to those achieved by software models with 4-bit weights. Reprinted with permission under a Creative Commons CC BY License from ref 225. Copyright 2022 Springer Nature

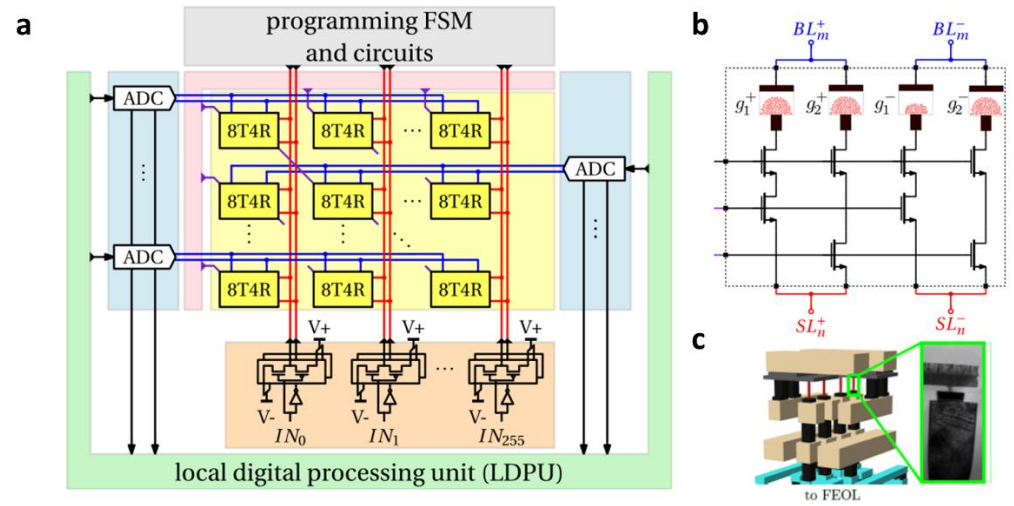

Figure 23. (a) Schematic of the PCM based crossbar array associated with the IBM HERMES project chip. It comprises 256x256 array of synaptic units.[227] (b) Each synaptic unit cell comprises 4 PCM devices onto which the synaptic weights are mapped in terms of the analog conductance value of the devices arising from the underlying phase configuration of the devices.[228] (c) Backend of the line integration of PCM devices.[227]

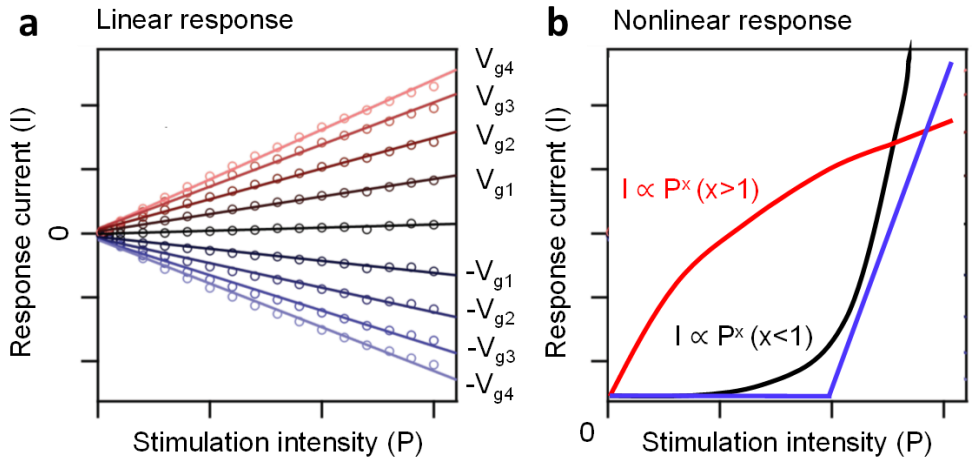

Figure 24. Different sensory response characteristics for in-sensor computing. (a) Linear response characteristics can provide high precision of in-sensor computing with artificial neural networks. (b) Nonlinear response characteristics of sensors, including superlinear (black), sublinear (red) and threshold (blue), can output intensity-dependent information, which allows to encode spatiotemporal information and enrich computation functions at sensory terminals. Reprinted with permission from ref 293. Copyright 2022 Springer Nature.

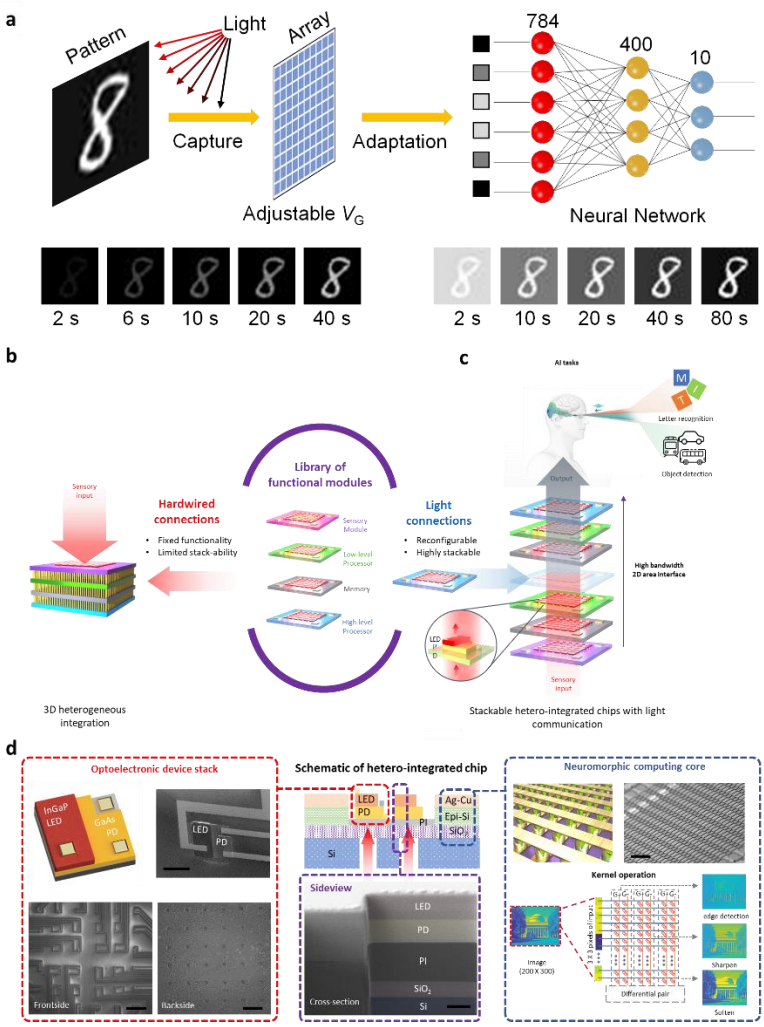

Figure 25. Examples of in-sensor computing and heterointegration with sensors (a) Bioinspired in-sensor vision adaptation using the sensor with sublinear response characteristics. Reprinted with permission from ref 290. Copyright 2022 Springer Nature. (b-d) Stackable Hetero-Integrated Chips for Reconfigurable Neuromorphic Computing. (b) A schematic illustration of the conventional 3D heterogenous integration (3DHI) for sensor-computing system with hardwired connections. (c) Schematics of stackable hetero-integrated chips with chip-to-chip light communication for reconfigurable sensor-computing system. (d) schematic representation of the stackable hetero-integrated chip. Illustration and SEM images of optoelectronic device stack (LED/PD, red dotted box), top-right scale bar: 100 μm, bottom row scale bars: 1 mm, and side

view of chips (LED/PD/substrate, purple dotted box), scale bar: 1 μm. Illustration and SEM images of neuromorphic computing core (Ag-Cu alloy-based Si memristor crossbar array, blue dotted box) scale bar: 100 μm. For the optoelectronic device stack, we fabricated an array of LED/PD stacks and aligned them with backside holes using deep reactive ion etching (DRIE). For the neuromorphic computing core, we fabricated 32 × 32 memristor crossbar arrays as shown in schematic and a SEM image. Three different types of 3 × 3 kernel operations (edge detection, sharpen, and soften) are performed. Reprinted with permission from ref 320. Copyright 2022 Springer Nature.

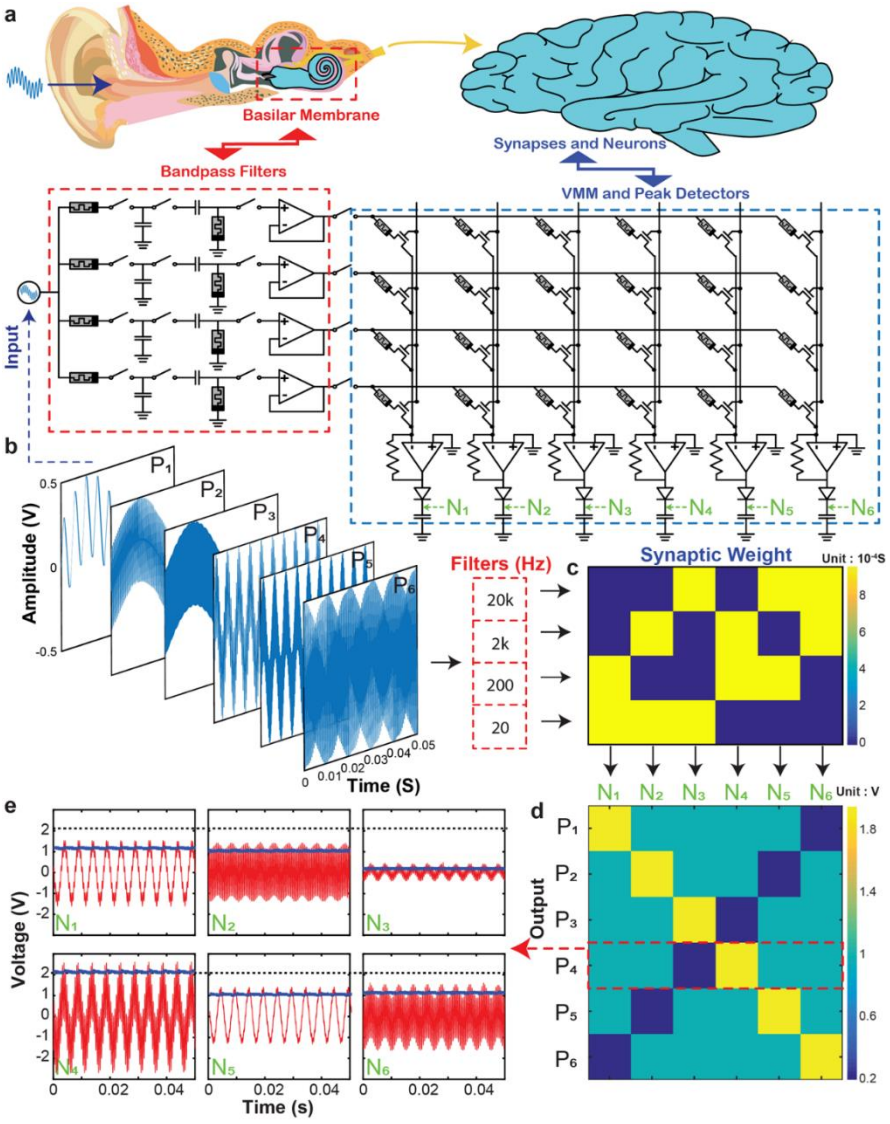

Figure 26. A memFPAA mixed-frequency signal classifier. (a) Schematic of the classifier circuit consisting of a bank of cascaded high and low pass filters, serving the role of the basilar membrane of the human hearing system, followed by a 4×6 VMM as synaptic array with 6 neurons implemented by TIAs and peak detectors. (b) 6 different temporal input patterns (P1~P6), each of which is the composition of two sinusoidal waves of different frequencies. (From the left, 20Hz+200Hz, 20Hz+2kHz, 20Hz+20kHz, 200Hz+2kHz, 200Hz+20kHz, and 2kHz+20kHz.) (c) Measured conductance map of the 4×6 VMM synaptic array. (d) Measured output peak voltages

of the post-synaptic neurons to the input patterns. Each individual input pattern is associated with a response of the output artificial neurons. (e) The temporal output responses of the 6 output neurons N1~N6 to input pattern P4. The red signals are voltages measured after the TIAs and the blue signals are voltages measured after the peak detectors. Reprinted with permission under a Creative Commons CC BY License from ref 328. Copyright 2022 John Wiley & Sons.

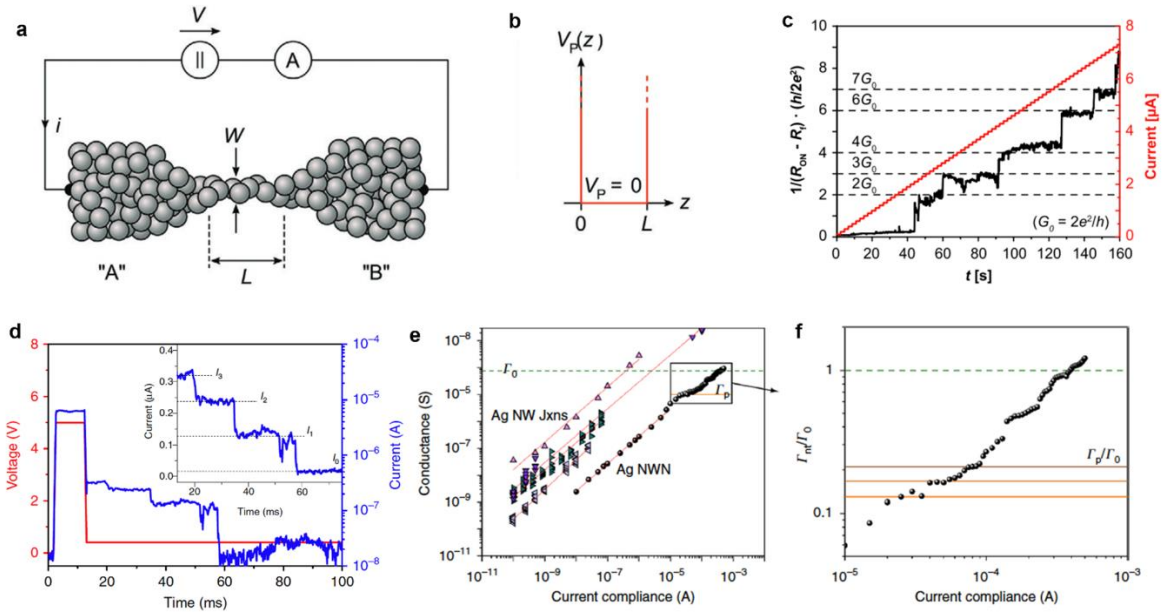

Figure 27. Quantum effects in memristive devices. (a) Schematic representation of ballistic electron transport in memristive devices, where two conductive reservoirs A and B representing metal electrodes are connected by a constriction with length L and width W and (b) potential energy of a particle in a 1D box. Reprinted with permission under a Creative Commons CC BY License from ref 31. Copyright 2022 John Wiley & Sons. (c) Quantum effects in gapless-type resistive switching devices observed by stimulating the device with a current sweep. Reproduced with permission from ref 351 Copyright 2012 IOP Publishing Ltd. (d) Discrete conductance steps during the spontaneous relaxation of a single NW memristive cell with volatile resistive switching mechanism. Reprinted with permission under a Creative Commons CC BY License from ref 355. Copyright 2018 Springer Nature. (e) Dependence of the conductance of Ag nanowire networks (NWN) and Ag nanowire junctions (NW Jxns) on the applied current compliance and (f) enlarged view showing conductance plateaus close to the quantum conductance. Reprinted with permission under a Creative Commons CC BY License from ref 365. Copyright 2018 Springer Nature.

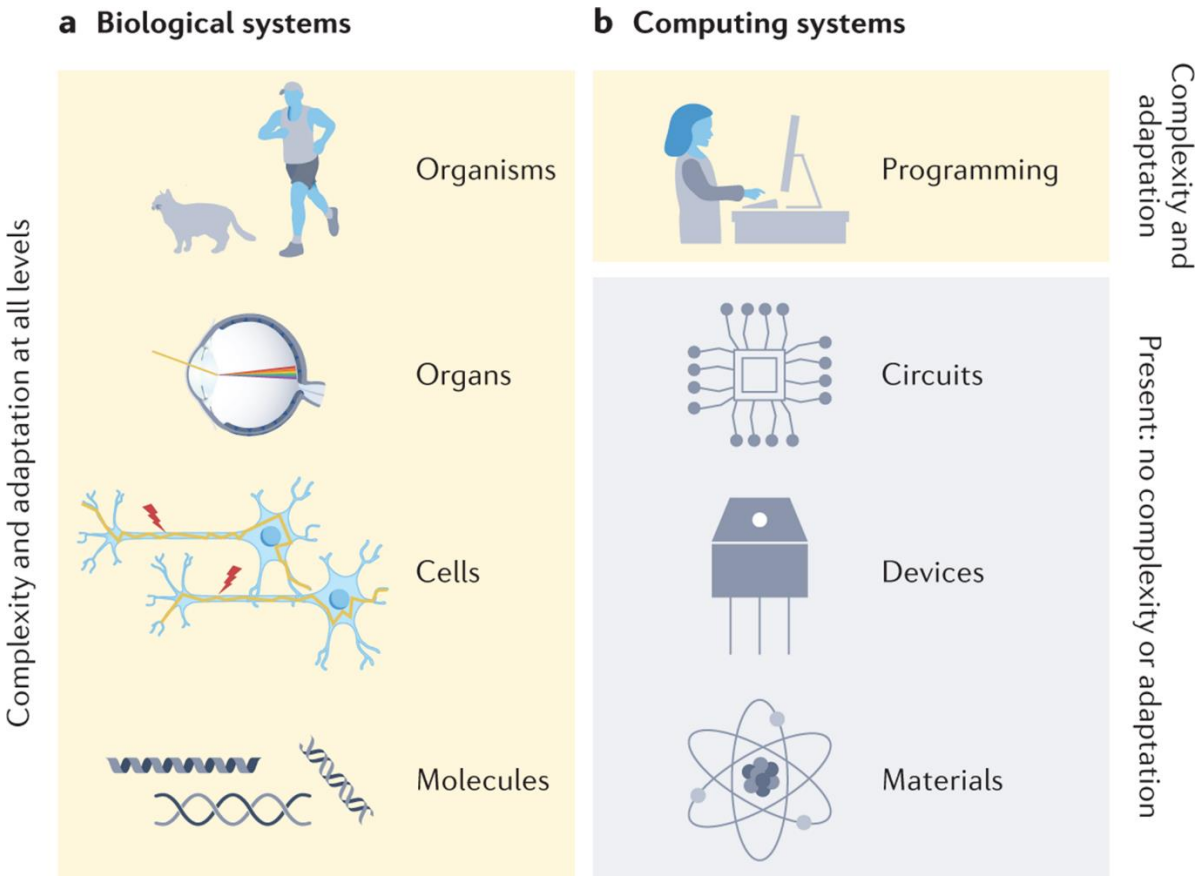

Figure 28. Even though there are parallels between biological systems and electronic systems, in biological systems, dynamics are leveraged at all levels to directly perform computing (yellow blocks), while in conventional electronic systems, dynamics are typically lacking and computing is only enabled by software. Reprinted with permission from ref 374. Copyright 2022 Springer Nature

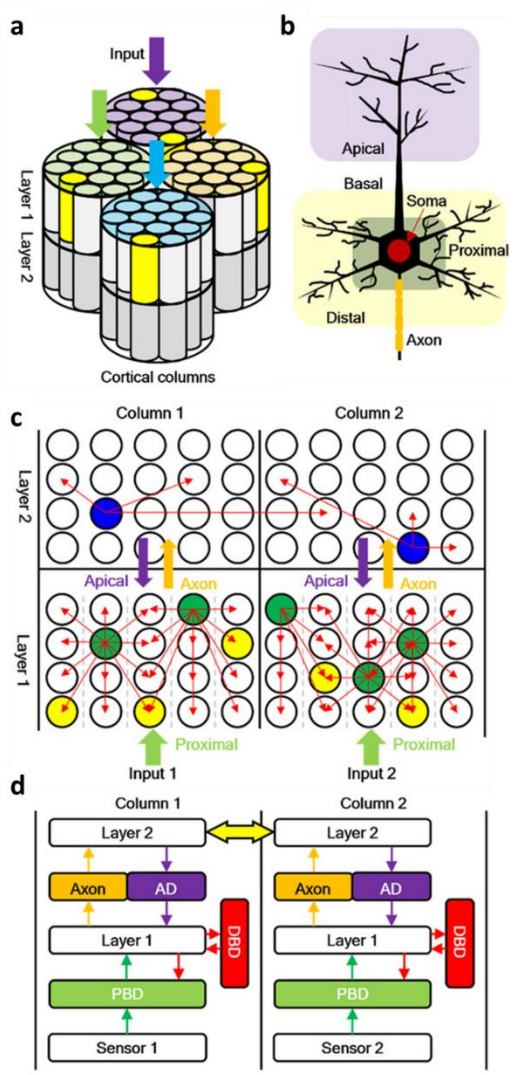

Figure 29. (a) Columnar structures in the cortex. A column of neurons processes the input by forming a receptive field (represented by colored circles), in which cells process a local region (indicated in yellow) (b) Schematic of a pyramidal neuron. The basal dendrites process feedforward and lateral signals, while the apical dendrites handle upper-layer feedback. (c) Basic structure of a columnar learning network (CLN). Various connections in the CLN include proximal basal dendrites (PBDs, green), distal basal dendrites (DBDs, red), axonal (orange), and apical dendrite (AD, purple) connections. (d) The connections are indicated in the same color as those in

(c). Reprinted with permission under a Creative Commons CC BY License from ref 400. Copyright 2022 John Wiley & Sons.

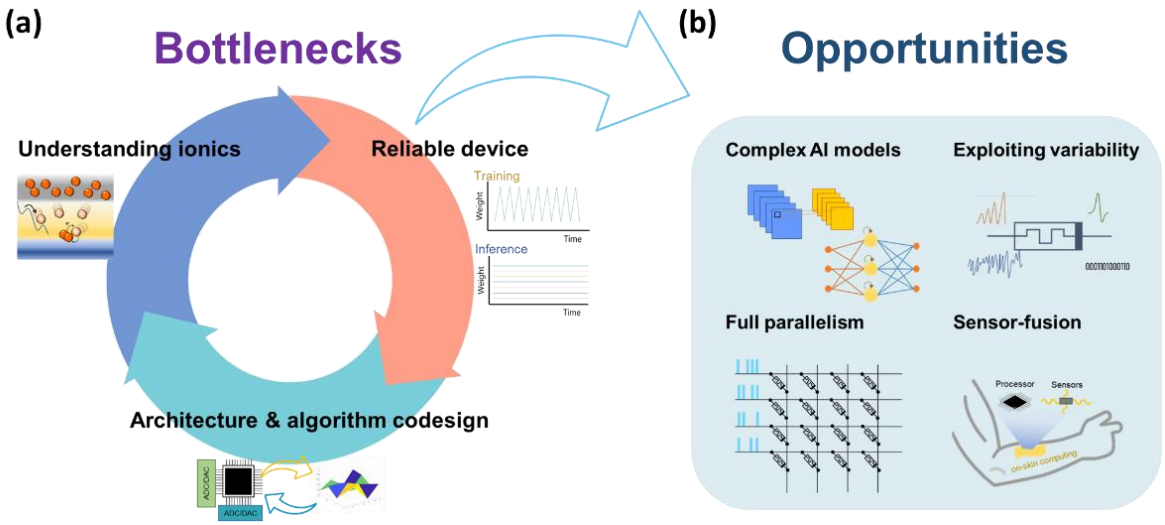

Figure 30. (a) Representative bottlenecks of current memristive technology. (b) Opportunities of future memristive technology.

Task	Combinatorial optimization	Neural networks of random features	Bayesian Inference	Cybersecurity	Hyperdimensional Computing	Linear system solver
Hardware	Resistive memory-based MAC with tunable read noise	Resistive memory-based MAC with intrinsic programming variation			Resistive memory-based CAM	Hybrid analog-digital MAC system
Software	Energy-based models such as (i) Hopfield neural network [207, 209](ii) Boltzmann machine[208]	(i) Echo state graph neural network[212, 213] (ii) Echo state network with random convolutional kernels[214]	Markov chain Monte Carlo using Metropolis–Hastings algorithm[215]	Memory-based physical unclonable function[216]	Encoder and associative memory[217, 218]	(i) Iterative refinement algorithm. (ii) Newton's method
Example	Max-cut[207-209]	(i) Graph learning[212, 213] (ii) Spatial-temporal signal learning[214]	(i) Supervised classification[215] (ii) Reinforcement learning classical control problem[215]	Authentication[216]	(i) Supervised image classification[217] (ii) Few shot learning of images[218]	(i) Partial correlation. (ii) Differential equation

Table 1. Summary of the recent development of hardware-software codesign using analog resistive memory. MAC: Multiply–accumulator. CAM: content-addressable memory.

AUTHOR INFORMATION

Corresponding Author

*Email: jeehwan@mit.edu

Author Contributions

‡M.-K.S and J.-H.K. contributed equally. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

H.B. holds shares in Salience Labs Ltd.

ACKNOWLEDGMENT

This review was developed out of the conference “The 5th International Conference on Memristive Materials, Devices & Systems (MEMRISYS 2022)”, held on November 30 – December 02, 2022. M.-K.S. acknowledges support from the National Research Foundation of Korea (NRF-2021R1A6A3A14044297). L.A. acknowledges support from Deutsche Forschungsgemeinschaft (DFG) under Project AL 560/21-1 and from the framework of the WAKeMeUP and StorAIge project which received funding from the Electronic Components and Systems for European Leadership Joint Undertaking in collaboration with the European Union’s H2020 Framework Programme (H2020/2014-2020) and National Authorities, under grant agreement No. 783176 and No. 101007321, respectively. H.B. acknowledges support from the European Union’s Horizon 2020 research and innovation programme (Grant No. 101017237, PHOENICS Project) and European Union's Innovation Council Pathfinder programme (Grant No. 101046878, HYBRAIN Project). J.A.d.A., M.O., J.L. and B.Y. acknowledge support from the

MIT-IBM Watson AI Lab. M.C.H. acknowledges support from the National Science Foundation Materials Research Science and Engineering Center at Northwestern University (NSF DMR-1720139). S.H. and S.C. acknowledge support from by the R&D programs of National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (2022M3F3A2A01072851 and 2022M3I7A2078273). S. K. acknowledges support from Samsung Science & Technology Foundation (Grant No. SRFC-IT2001-40206). J.-S.S. acknowledges support from JUMP COCOSYS, a Semiconductor Research Corporation program sponsored by Defense Advanced Research Projects Agency. Z.W. thanks the support from the Hong Kong Research Grant Council (Grant Nos. 27206321 and 17205922) and National Natural Science Foundation of China (Grant No. 62122004). Y.C acknowledges support from Research Grant Council of Hong Kong (PolyU502/22). G.M., C.R. and I.V. acknowledge support from the European project MEMQuD, code 20FUN06. This project (EMPIR 20FUN06 MEMQuD) has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme. S.G.C. acknowledges financial support from the DOE Office of Science (ASCR / BES) for the Microelectronics Co-Design project COINLFIPS. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology \& Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

REFERENCES

(1) Auth, C. P.; Plummer, J. D. Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFET's. *IEEE Electron Device Lett.* **1997**, *18*, 74-76.

(2) Yakimets, D.; Bardon, M. G.; Jang, D.; Schuddinck, P.; Sherazi, Y.; Weckx, P.; Miyaguchi, K.; Parvais, B.; Raghavan, P.; Spessot, A. Power aware FinFET and lateral nanosheet FET targeting for 3nm CMOS technology. In *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017; IEEE: pp 20.24. 21-20.24. 24.

(3) Das, S.; Chen, A.; Marinella, M. Beyond CMOS. In *2021 IEEE International Roadmap for Devices and Systems Outbriefs*, 2021; IEEE: pp 01-129.

(4) Chen, A. Beyond-CMOS roadmap—from Boolean logic to neuro-inspired computing. *Jpn. J. Appl. Phys.* **2022**, *61*, SM1003.

(5) Yang, J. J.; Strukov, D. B.; Stewart, D. R. Memristive devices for computing. *Nat. Nanotechnol.* **2013**, *8*, 13-24.

(6) Lanza, M.; Sebastian, A.; Lu, W. D.; Le Gallo, M.; Chang, M.-F.; Akinwande, D.; Puglisi, F. M.; Alshareef, H. N.; Liu, M.; Roldan, J. B. Memristive technologies for data storage, computation, encryption, and radio-frequency communication. *Science* **2022**, *376*, eabj9979.

(7) Christensen, D. V.; Dittmann, R.; Linares-Barranco, B.; Sebastian, A.; Le Gallo, M.; Redaelli, A.; Slesazeck, S.; Mikolajick, T.; Spiga, S.; Menzel, S. 2022 roadmap on neuromorphic computing and engineering. *Neuromorphic Comput. Eng.* **2022**, *2*, 022501.

(8) Song, M.-K.; Namgung, S. D.; Choi, D.; Kim, H.; Seo, H.; Ju, M.; Lee, Y. H.; Sung, T.; Lee, Y.-S.; Nam, K. T. Proton-enabled activation of peptide materials for biological bimodal memory. *Nat. Commun.* **2020**, *11*, 5896.

(9) Strukov, D. B.; Snider, G. S.; Stewart, D. R.; Williams, R. S. The missing memristor found. *Nature* **2008**, *453*, 80-83.

(10) Chua, L. Memristor-the missing circuit element. *IEEE Trans. Circuits Theory* **1971**, *18*, 507-519.

(11) Prezioso, M.; Merrikh-Bayat, F.; Hoskins, B. D.; Adam, G. C.; Likharev, K. K.; Strukov, D. B. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature* **2015**, *521*, 61-64.

(12) Xia, Q.; Yang, J. J. Memristive crossbar arrays for brain-inspired computing. *Nat. Mater.* **2019**, *18*, 309-323.

(13) Hickmott, T. Low-frequency negative resistance in thin anodic oxide films. *J. Appl. Phys.* **1962**, *33*, 2669-2682.

(14) Gibbons, J.; Beadle, W. Switching properties of thin NiO films. *Solid State Electron.* **1964**, *7*, 785-790.

(15) Terabe, K.; Tsuchiya, T.; Yang, R.; Aono, M. Nanoionic devices enabling a multitude of new features. *Nanoscale* **2016**, *8*, 13873-13879.

(16) Chua, L. O.; Kang, S. M. Memristive devices and systems. *Proc. IEEE* **1976**, *64*, 209-223.

(17) Valov, I.; Kozicki, M. N. Cation-based resistance change memory. *J. Phys. D: Appl. Phys* **2013**, *46*, 074005.

(18) Ielmini, D. Resistive switching memories based on metal oxides: mechanisms, reliability and scaling. *Semicond. Sci. Technol.* **2016**, *31*, 063002.

(19) Wong, H.-S. P.; Lee, H.-Y.; Yu, S.; Chen, Y.-S.; Wu, Y.; Chen, P.-S.; Lee, B.; Chen, F. T.; Tsai, M.-J. Metal–Oxide RRAM. *Proc. IEEE* **2012**, *100*, 1951-1970.

(20) Dittmann, R.; Menzel, S.; Waser, R. Nanoionic memristive phenomena in metal oxides: the valence change mechanism. *Adv. Phys.* **2021**, *70*, 155-349.

(21) Ielmini, D.; Bruchhaus, R.; Waser, R. Thermochemical resistive switching: materials, mechanisms, and scaling projections. *Phase Transit.* **2011**, *84*, 570-602.

(22) Terabe, K.; Hasegawa, T.; Nakayama, T.; Aono, M. Quantized conductance atomic switch. *Nature* **2005**, *433*, 47-50.

(23) Waser, R.; Aono, M. Nanoionics-based resistive switching memories. *Nat. Mater.* **2007**, *6*, 833-840.

(24) Chen, S.; Valov, I. Design of Materials Configuration for Optimizing Redox-Based Resistive Switching Memories. *Adv. Mater.* **2022**, *34*, 2105022.

(25) Lübben, M.; Cüppers, F.; Mohr, J.; von Witzleben, M.; Breuer, U.; Waser, R.; Neumann, C.; Valov, I. Design of defect-chemical properties and device performance in memristive systems. *Sci. Adv.* **2020**, *6*, eaaz9079.

(26) Valov, I.; Tsuruoka, T. Effects of moisture and redox reactions in VCM and ECM resistive switching memories. *J. Phys. D: Appl. Phys* **2018**, *51*, 413001.

(27) Yeon, H.; Lin, P.; Choi, C.; Tan, S. H.; Park, Y.; Lee, D.; Lee, J.; Xu, F.; Gao, B.; Wu, H. Alloying conducting channels for reliable neuromorphic computing. *Nat. Nanotechnol.* **2020**, *15*, 574-579.

(28) Waser, R.; Dittmann, R.; Staikov, G.; Szot, K. Redox-based resistive switching memories–nanoionic mechanisms, prospects, and challenges. *Adv. Mater.* **2009**, *21*, 2632-2663.

(29) Sharath, S. U.; Vogel, S.; Molina-Luna, L.; Hildebrandt, E.; Wenger, C.; Kurian, J.; Duerrschnabel, M.; Niermann, T.; Niu, G.; Calka, P.; et al. Control of switching modes and conductance quantization in oxygen engineered HfO_x based memristive devices. *Adv. Funct. Mater.* **2017**, *27*, 1700432.

(30) Kaiser, N.; Vogel, T.; Zintler, A.; Petzold, S.; Arzumanov, A.; Piros, E.; Eilhardt, R.; Molina-Luna, L.; Alff, L. Defect-stabilized substoichiometric polymorphs of hafnium oxide with semiconducting properties. *ACS Appl. Mater. Interfaces* **2021**, *14*, 1290-1303.

(31) Milano, G.; Aono, M.; Boarino, L.; Celano, U.; Hasegawa, T.; Kozicki, M.; Majumdar, S.; Menghini, M.; Miranda, E.; Ricciardi, C. Quantum conductance in memristive devices: fundamentals, developments, and applications. *Adv. Mater.* **2022**, *34*, 2201248.

(32) Petzold, S.; Zintler, A.; Eilhardt, R.; Piros, E.; Kaiser, N.; Sharath, S. U.; Vogel, T.; Major, M.; McKenna, K. P.; Molina-Luna, L.; et al. Forming-Free Grain Boundary Engineered Hafnium Oxide Resistive Random Access Memory Devices. *Adv. Electron. Mater.* **2019**, *5*, 1900484.

(33) Winkler, R.; Zintler, A.; Petzold, S.; Piros, E.; Kaiser, N.; Vogel, T.; Nasiou, D.; McKenna, K. P.; Molina-Luna, L.; Alff, L. Controlling the Formation of Conductive Pathways in Memristive Devices. *Adv. Sci.* **2022**, *9*, 2201806.

(34) Petzold, S.; Piros, E.; Eilhardt, R.; Zintler, A.; Vogel, T.; Kaiser, N.; Radetinac, A.; Komissinskiy, P.; Jalaguier, E.; Nolot, E.; et al. Tailoring the Switching Dynamics in Yttrium Oxide-Based RRAM Devices by Oxygen Engineering: From Digital to Multi-Level Quantization toward Analog Switching. *Adv. Electron. Mater.* **2020**, *6*, 2000439.

(35) Piros, E.; Lonsky, M.; Petzold, S.; Zintler, A.; Sharath, S.; Vogel, T.; Kaiser, N.; Eilhardt, R.; Molina-Luna, L.; Wenger, C.; et al. Role of oxygen defects in conductive-filament formation in Y_2O_3-based analog RRAM devices as revealed by fluctuation spectroscopy. *Phys. Rev. Appl.* **2020**, *14*, 034029.

(36) Liu, C.; Chen, H.; Wang, S.; Liu, Q.; Jiang, Y.-G.; Zhang, D. W.; Liu, M.; Zhou, P. Two-dimensional materials for next-generation computing technologies. *Nat. Nanotechnol.* **2020**, *15*, 545-557.

(37) Wang, M.; Cai, S.; Pan, C.; Wang, C.; Lian, X.; Zhuo, Y.; Xu, K.; Cao, T.; Pan, X.; Wang, B. Robust memristors based on layered two-dimensional materials. *Nat. Electron.* **2018**, *1*, 130-136.

(38) Zhou, Z.; Yang, F.; Wang, S.; Wang, L.; Wang, X.; Wang, C.; Xie, Y.; Liu, Q. Emerging of two-dimensional materials in novel memristor. *Front. Phys.* **2022**, *17*, 1-14.

(39) Song, Y.-W.; Song, M.-K.; Choi, D.; Kwon, J.-Y. Encapsulation-enhanced switching stability of MoS_2 memristors. *J. Alloys Compd.* **2021**, *885*, 161016.

(40) Chen, S.; Mahmoodi, M. R.; Shi, Y.; Mahata, C.; Yuan, B.; Liang, X.; Wen, C.; Hui, F.; Akinwande, D.; Strukov, D. B.; et al. Wafer-scale integration of two-dimensional materials in high-density memristive crossbar arrays for artificial neural networks. *Nat. Electron.* **2020**, *3*, 638-645.

(41) Shi, Y.; Liang, X.; Yuan, B.; Chen, V.; Li, H.; Hui, F.; Yu, Z.; Yuan, F.; Pop, E.; Wong, H.-S. P.; et al. Electronic synapses made of layered two-dimensional materials. *Nat. Electron.* **2018**, *1*, 458-465.

(42) Zhu, K.; Pazos, S.; Aguirre, F.; Shen, Y.; Yuan, Y.; Zheng, W.; Alharbi, O.; Villena, M. A.; Fang, B.; Li, X.; et al. Hybrid 2D/CMOS microchips for memristive applications. *Nature* **2023**. DOI: 10.1038/s41586-023-05973-1.

(43) Shen, Y.; Zheng, W.; Zhu, K.; Xiao, Y.; Wen, C.; Liu, Y.; Jing, X.; Lanza, M. Variability and Yield in h-BN-Based Memristive Circuits: The Role of Each Type of Defect. *Adv. Mater.* **2021**, *33*, 2103656.

(44) Wen, C.; Li, X.; Zanotti, T.; Puglisi, F. M.; Shi, Y.; Saiz, F.; Antidormi, A.; Roche, S.; Zheng, W.; Liang, X.; et al. Advanced Data Encryption using 2D Materials. *Adv. Mater.* **2021**, *33*, 2100185.

(45) Zheng, W.; Saiz, F.; Shen, Y.; Zhu, K.; Liu, Y.; McAleese, C.; Conran, B.; Wang, X.; Lanza, M. Defect-Free Metal Deposition on 2D Materials via Inkjet Printing Technology. *Adv. Mater.* **2022**, *34*, 2104138.

(46) Lanza, M.; Wong, H. S. P.; Pop, E.; Ielmini, D.; Strukov, D.; Regan, B. C.; Larcher, L.; Villena, M. A.; Yang, J. J.; Goux, L. Recommended methods to study resistive switching devices. *Adv. Electron. Mater.* **2019**, *5*, 1800143.

(47) Lanza, M.; Hui, F.; Wen, C.; Ferrari, A. C. Resistive switching crossbar arrays based on layered materials. *Adv. Mater.* **2022**, 2205402.

(48) Wuttig, M.; Yamada, N. Phase-change materials for rewriteable data storage. *Nat. Mater.* **2007**, *6*, 824-832.

(49) Wuttig, M.; Bhaskaran, H.; Taubner, T. Phase-change materials for non-volatile photonic applications. *Nat. Photon.* **2017**, *11*, 465-476.

(50) Wang, Z.; Wu, H.; Burr, G. W.; Hwang, C. S.; Wang, K. L.; Xia, Q.; Yang, J. J. Resistive switching materials for information processing. *Nat. Rev. Mater.* **2020**, *5*, 173-195.

(51) Dong, W.; Liu, H.; Behera, J. K.; Lu, L.; Ng, R. J.; Sreekanth, K. V.; Zhou, X.; Yang, J. K.; Simpson, R. E. Wide bandgap phase change material tuned visible photonics. *Adv. Funct. Mater.* **2019**, *29*, 1806181.

(52) Delaney, M.; Zeimpekis, I.; Lawson, D.; Hewak, D. W.; Muskens, O. L. A new family of ultralow loss reversible phase-change materials for photonic integrated circuits: Sb_2S_3 and Sb_2Se_3. *Adv. Funct. Mater.* **2020**, *30*, 2002447.

(53) Ríos, C.; Stegmaier, M.; Hosseini, P.; Wang, D.; Scherer, T.; Wright, C. D.; Bhaskaran, H.; Pernice, W. H. Integrated all-photonic non-volatile multi-level memory. *Nat. Photon.* **2015**, *9*, 725-732.

(54) Rios, C.; Stegmaier, M.; Cheng, Z.; Youngblood, N.; Wright, C. D.; Pernice, W. H.; Bhaskaran, H. Controlled switching of phase-change materials by evanescent-field coupling in integrated photonics. *Opt. Mater. Express* **2018**, *8*, 2455-2470.

(55) Li, X.; Youngblood, N.; Ríos, C.; Cheng, Z.; Wright, C. D.; Pernice, W. H.; Bhaskaran, H. Fast and reliable storage using a 5 bit, nonvolatile photonic memory cell. *Optica* **2019**, *6*, 1-6.

(56) Ríos, C.; Youngblood, N.; Cheng, Z.; Le Gallo, M.; Pernice, W. H.; Wright, C. D.; Sebastian, A.; Bhaskaran, H. In-memory computing on a photonic platform. *Sci. Adv.* **2019**, *5*, eaau5759.

(57) Cheng, Z.; Ríos, C.; Pernice, W. H.; Wright, C. D.; Bhaskaran, H. On-chip photonic synapse. *Sci. Adv.* **2017**, *3*, e1700160.

(58) Feldmann, J.; Youngblood, N.; Wright, C. D.; Bhaskaran, H.; Pernice, W. H. All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature* **2019**, *569*, 208-214.

(59) Tan, J. Y.; Cheng, Z.; Feldmann, J.; Li, X.; Youngblood, N.; Ali, U. E.; Wright, C. D.; Pernice, W. H.; Bhaskaran, H. Monadic Pavlovian associative learning in a backpropagation-free photonic network. *Optica* **2022**, *9*, 792-802.

(60) Fang, Z.; Chen, R.; Zheng, J.; Khan, A. I.; Neilson, K. M.; Geiger, S. J.; Callahan, D. M.; Moebius, M. G.; Saxena, A.; Chen, M. E. Ultra-low-energy programmable non-volatile silicon photonics based on phase-change materials with graphene heaters. *Nat. Nanotechnol.* **2022**, *17*, 842-848.

(61) Ríos, C.; Zhang, Y.; Shalaginov, M. Y.; Deckoff-Jones, S.; Wang, H.; An, S.; Zhang, H.; Kang, M.; Richardson, K. A.; Roberts, C. Multi-level electro-thermal switching of optical phase-change materials using graphene. *Adv. Photonics Res.* **2021**, *2*, 2000034.

(62) Zheng, J.; Fang, Z.; Wu, C.; Zhu, S.; Xu, P.; Doylend, J. K.; Deshmukh, S.; Pop, E.; Dunham, S.; Li, M. Nonvolatile electrically reconfigurable integrated photonic switch enabled by a silicon PIN diode heater. *Adv. Mater.* **2020**, *32*, 2001218.

(63) Li, X.; Youngblood, N.; Zhou, W.; Feldmann, J.; Swett, J.; Aggarwal, S.; Sebastian, A.; Wright, C. D.; Pernice, W.; Bhaskaran, H. On-chip phase change optical matrix multiplication core. In *2020 IEEE International Electron Devices Meeting (IEDM)*, 2020; IEEE: pp 7.5. 1-7.5. 4.

(64) Feldmann, J.; Youngblood, N.; Karpov, M.; Gehring, H.; Li, X.; Stappers, M.; Le Gallo, M.; Fu, X.; Lukashchuk, A.; Raja, A. S. Parallel convolutional processing using an integrated photonic tensor core. *Nature* **2021**, *589*, 52-58.

(65) Sarwat, S. G.; Moraitis, T.; Wright, C. D.; Bhaskaran, H. Chalcogenide optomemristors for multi-factor neuromorphic computation. *Nat. Commun.* **2022**, *13*, 2247.

(66) Farmakidis, N.; Youngblood, N.; Li, X.; Tan, J.; Swett, J. L.; Cheng, Z.; Wright, C. D.; Pernice, W. H.; Bhaskaran, H. Plasmonic nanogap enhanced phase-change devices with dual electrical-optical functionality. *Sci. Adv.* **2019**, *5*, eaaw2687.

(67) Farmakidis, N.; Youngblood, N.; Lee, J. S.; Feldmann, J.; Lodi, A.; Li, X.; Aggarwal, S.; Zhou, W.; Bogani, L.; Pernice, W. H. Electronically reconfigurable photonic switches incorporating plasmonic structures and phase change materials. *Adv. Sci.* **2022**, *9*, 2200383.

(68) Shen, J.; Jia, S.; Shi, N.; Ge, Q.; Gotoh, T.; Lv, S.; Liu, Q.; Dronskowski, R.; Elliott, S. R.; Song, Z. Elemental electrical switch enabling phase segregation–free operation. *Science* **2021**, *374*, 1390-1394.

(69) Cheng, Z.; Milne, T.; Salter, P.; Kim, J. S.; Humphrey, S.; Booth, M.; Bhaskaran, H. Antimony thin films demonstrate programmable optical nonlinearity. *Sci. Adv.* **2021**, *7*, eabd7097.

(70) Aggarwal, S.; Milne, T.; Farmakidis, N.; Feldmann, J.; Li, X.; Shu, Y.; Cheng, Z.; Salinga, M.; Pernice, W. H.; Bhaskaran, H. Antimony as a Programmable Element in Integrated Nanophotonics. *Nano Lett.* **2022**, *22*, 3532-3538.

(71) Salinga, M.; Kersting, B.; Ronneberger, I.; Jonnalagadda, V. P.; Vu, X. T.; Le Gallo, M.; Giannopoulos, I.; Cojocaru-Mirédin, O.; Mazzarello, R.; Sebastian, A. Monatomic phase change memory. *Nat. Mater.* **2018**, *17*, 681-685.

(72) Zuliani, P.; Varesi, E.; Palumbo, E.; Borghi, M.; Tortorelli, I.; Erbetta, D.; Dalla Libera, G.; Pessina, N.; Gandolfo, A.; Prelini, C. Overcoming Temperature Limitations in Phase Change Memories With Optimized $Ge_xSb_yTe_z$. *IEEE Trans. Electron Devices* **2013**, *60*, 4020-4026.

(73) Müller, J.; Böscke, T.; Bräuhaus, D.; Schröder, U.; Böttger, U.; Sundqvist, J.; Kücher, P.; Mikolajick, T.; Frey, L. Ferroelectric $Zr_{0.5}Hf_{0.5}O_2$ thin films for nonvolatile memory applications. *Appl. Phys. Lett.* **2011**, *99*, 112901.

(74) Li, B.; Li, S.; Wang, H.; Chen, L.; Liu, L.; Feng, X.; Li, Y.; Chen, J.; Gong, X.; Ang, K. W. An electronic synapse based on 2D ferroelectric $CuInP_2S_6$. *Adv. Electron. Mater.* **2020**, *6*, 2000760.

(75) Kim, D.; Heo, S. J.; Pyo, G.; Choi, H. S.; Kwon, H.-J.; Jang, J. E. PZT ferroelectric synapse TFT with multi-level of conductance state for neuromorphic applications. *IEEE Access* **2021**, *9*, 140975-140982.

(76) Lee, M.; Park, W.; Son, H.; Seo, J.; Kwon, O.; Oh, S.; Hahm, M.; Kim, U.; Cho, B. Brain-inspired ferroelectric Si nanowire synaptic device. *APL Mater.* **2021**, *9*, 031103.

(77) Lee, D. H.; Park, G. H.; Kim, S. H.; Park, J. Y.; Yang, K.; Slesazeck, S.; Mikolajick, T.; Park, M. H. Neuromorphic devices based on fluorite-structured ferroelectrics. *InfoMat* **2022**, *4*, e12380.

(78) Mulaosmanovic, H.; Breyer, E. T.; Dünkel, S.; Beyer, S.; Mikolajick, T.; Slesazeck, S. Ferroelectric field-effect transistors based on HfO_2: a review. *Nanotechnology* **2021**, *32*, 502002.

(79) Jerry, M.; Chen, P.-Y.; Zhang, J.; Sharma, P.; Ni, K.; Yu, S.; Datta, S. Ferroelectric FET analog synapse for acceleration of deep neural network training. In *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017; IEEE: pp 6.2. 1-6.2. 4.

(80) Khan, A. I.; Keshavarzi, A.; Datta, S. The future of ferroelectric field-effect transistor technology. *Nat. Electron.* **2020**, *3*, 588-597.

(81) Kim, I. J.; Lee, J. S. Ferroelectric Transistors for Memory and Neuromorphic Device Applications. *Adv. Mater.* **2022**, 2206864.

(82) Kim, M.-K.; Lee, J.-S. Ferroelectric analog synaptic transistors. *Nano Lett.* **2019**, *19*, 2044-2050.

(83) Oh, S.; Kim, T.; Kwak, M.; Song, J.; Woo, J.; Jeon, S.; Yoo, I. K.; Hwang, H. $HfZrO_x$-based ferroelectric synapse device with 32 levels of conductance states for neuromorphic applications. *IEEE Electron Device Lett.* **2017**, *38*, 732-735.

(84) Yang, S. T.; Li, X. Y.; Yu, T. L.; Wang, J.; Fang, H.; Nie, F.; He, B.; Zhao, L.; Lü, W. M.; Yan, S. S. High-Performance Neuromorphic Computing Based on Ferroelectric Synapses with Excellent Conductance Linearity and Symmetry. *Adv. Funct. Mater.* **2022**, *32*, 2202366.

(85) Kim, M. K.; Lee, J. S. Synergistic Improvement of Long-Term Plasticity in Photonic Synapses Using Ferroelectric Polarization in Hafnia-Based Oxide-Semiconductor Transistors. *Adv. Mater.* **2020**, *32*, 1907826.

(86) Kim, M.-K.; Kim, I.-J.; Lee, J.-S. Oxide semiconductor-based ferroelectric thin-film transistors for advanced neuromorphic computing. *Appl. Phys. Lett.* **2021**, *118*, 032902.

(87) Jerry, M.; Dutta, S.; Kazemi, A.; Ni, K.; Zhang, J.; Chen, P.-Y.; Sharma, P.; Yu, S.; Hu, X. S.; Niemier, M. A ferroelectric field effect transistor based synaptic weight cell. *J. Phys. D: Appl. Phys* **2018**, *51*, 434001.

(88) Kim, D.; Jeon, Y.-R.; Ku, B.; Chung, C.; Kim, T. H.; Yang, S.; Won, U.; Jeong, T.; Choi, C. Analog synaptic transistor with Al-doped HfO_2 ferroelectric thin film. *ACS Appl. Mater. Interfaces* **2021**, *13*, 52743-52753.

(89) Cheema, S. S.; Kwon, D.; Shanker, N.; Dos Reis, R.; Hsu, S.-L.; Xiao, J.; Zhang, H.; Wagner, R.; Datar, A.; McCarter, M. R. Enhanced ferroelectricity in ultrathin films grown directly on silicon. *Nature* **2020**, *580*, 478-482.

(90) Lyu, X.; Si, M.; Shrestha, P.; Cheung, K.; Ye, P. First direct measurement of sub-nanosecond polarization switching in ferroelectric hafnium zirconium oxide. In *2019 IEEE International Electron Devices Meeting (IEDM)*, 2019; IEEE: pp 15.12. 11-15.12. 14.

(91) Kim, M.-K.; Kim, I.-J.; Lee, J.-S. CMOS-compatible ferroelectric NAND flash memory for high-density, low-power, and high-speed three-dimensional memory. *Sci. Adv.* **2021**, *7*, eabe1341.

(92) Chen, P.-Y.; Peng, X.; Yu, S. NeuroSim: A circuit-level macro model for benchmarking neuro-inspired architectures in online learning. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2018**, *37*, 3067-3080.

(93) Peng, X.; Huang, S.; Jiang, H.; Lu, A.; Yu, S. DNN+ NeuroSim V2. 0: An end-to-end benchmarking framework for compute-in-memory accelerators for on-chip training. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2020**, *40*, 2306-2319.

(94) Oh, S.; Hwang, H.; Yoo, I. Ferroelectric materials for neuromorphic computing. *APL Mater.* **2019**, *7*, 091109.

(95) Kim, G.; Ko, D. H.; Kim, T.; Lee, S.; Jung, M.; Lee, Y. K.; Lim, S.; Jo, M.; Eom, T.; Shin, H. Power-Delay Area-Efficient Processing-In-Memory Based on Nanocrystalline Hafnia Ferroelectric Field-Effect Transistors. *ACS Appl. Mater. Interfaces* **2022**, *15*, 1463-1474.

(96) Lee, G. H.; Song, M. S.; Kim, S.; Yim, J.; Hwang, S.; Yu, J.; Kwon, D.; Kim, H. Ferroelectric Field-Effect Transistors for Binary Neural Network With 3-D NAND Architecture. *IEEE Trans. Electron Devices* **2022**, *69*, 6438-6445.

(97) Huang, M.; Schwacke, M.; Onen, M.; Del Alamo, J.; Li, J.; Yildiz, B. Electrochemical Ionic Synapses: Progress and Perspectives. *Adv. Mater.* **2022**, 2205169.

(98) Fuller, E. J.; Gabaly, F. E.; Léonard, F.; Agarwal, S.; Plimpton, S. J.; Jacobs-Gedrim, R. B.; James, C. D.; Marinella, M. J.; Talin, A. A. Li-ion synaptic transistor for low power analog computing. *Adv. Mater.* **2017**, *29*, 1604310.

(99) Tang, J.; Bishop, D.; Kim, S.; Copel, M.; Gokmen, T.; Todorov, T.; Shin, S.; Lee, K.-T.; Solomon, P.; Chan, K. ECRAM as scalable synaptic cell for high-speed, low-power neuromorphic

computing. In *2018 IEEE International Electron Devices Meeting (IEDM)*, 2018; IEEE: pp 13.11. 11-13.11. 14.

(100) Yang, C. S.; Shang, D. S.; Liu, N.; Fuller, E. J.; Agrawal, S.; Talin, A. A.; Li, Y. Q.; Shen, B. G.; Sun, Y. All-solid-state synaptic transistor with ultralow conductance for neuromorphic computing. *Adv. Funct. Mater.* **2018**, *28*, 1804170.

(101) Nikam, R. D.; Kwak, M.; Lee, J.; Rajput, K. G.; Banerjee, W.; Hwang, H. Near ideal synaptic functionalities in Li ion synaptic transistor using $Li_3PO_xSe_x$ electrolyte with high ionic conductivity. *Sci. Rep.* **2019**, *9*, 18883.

(102) Kim, S.; Todorov, T.; Onen, M.; Gokmen, T.; Bishop, D.; Solomon, P.; Lee, K.-T.; Copel, M.; Farmer, D. B.; Ott, J. A. Metal-oxide based, CMOS-compatible ECRAM for Deep Learning Accelerator. In *2019 IEEE International Electron Devices Meeting (IEDM)*, 2019; IEEE: pp 35.37. 31-35.37. 34.

(103) Li, Y.; Fuller, E. J.; Sugar, J. D.; Yoo, S.; Ashby, D. S.; Bennett, C. H.; Horton, R. D.; Bartsch, M. S.; Marinella, M. J.; Lu, W. D. Filament-free bulk resistive memory enables deterministic analogue switching. *Adv. Mater.* **2020**, *32*, 2003984.

(104) Nikam, R. D.; Kwak, M.; Hwang, H. All-Solid-State Oxygen Ion Electrochemical Random-Access Memory for Neuromorphic Computing. *Adv. Electron. Mater.* **2021**, *7*, 2100142.

(105) Van De Burgt, Y.; Lubberman, E.; Fuller, E. J.; Keene, S. T.; Faria, G. C.; Agarwal, S.; Marinella, M. J.; Alec Talin, A.; Salleo, A. A non-volatile organic electrochemical device as a low-voltage artificial synapse for neuromorphic computing. *Nat. Mater.* **2017**, *16*, 414-418.

(106) Melianas, A.; Quill, T.; LeCroy, G.; Tuchman, Y.; Loo, H. v.; Keene, S.; Giovannitti, A.; Lee, H.; Maria, I.; McCulloch, I. Temperature-resilient solid-state organic artificial synapses for neuromorphic computing. *Sci. Adv.* **2020**, *6*, eabb2958.

(107) Yao, X.; Klyukin, K.; Lu, W.; Onen, M.; Ryu, S.; Kim, D.; Emond, N.; Waluyo, I.; Hunt, A.; Del Alamo, J. A. Protonic solid-state electrochemical synapse for physical neural networks. *Nat. Commun.* **2020**, *11*, 3134.

(108) Nikam, R. D.; Lee, J.; Choi, W.; Banerjee, W.; Kwak, M.; Yadav, M.; Hwang, H. Ionic Sieving Through One-Atom-Thick 2D Material Enables Analog Nonvolatile Memory for Neuromorphic Computing. *Small* **2021**, *17*, 2103543.

(109) Onen, M.; Emond, N.; Li, J.; Yildiz, B.; Del Alamo, J. A. CMOS-compatible protonic programmable resistor based on phosphosilicate glass electrolyte for analog deep learning. *Nano Lett.* **2021**, *21*, 6111-6116.

(110) Onen, M.; Emond, N.; Wang, B.; Zhang, D.; Ross, F. M.; Li, J.; Yildiz, B.; Del Alamo, J. A. Nanosecond protonic programmable resistors for analog deep learning. *Science* **2022**, *377*, 539-543.

(111) Onen, M.; Li, J.; Yildiz, B.; Del Alamo, J. Dynamics of PSG-Based Nanosecond Protonic Programmable Resistors for Analog Deep Learning. In *2022 International Electron Devices Meeting (IEDM)*, 2022; IEEE: pp 2.6. 1-2.6. 4.

(112) Li, H.; Jin, D.; Kong, X.; Tu, H.; Yu, Q.; Jiang, F. High proton-conducting monolithic phosphosilicate glass membranes. *Microporous Mesoporous Mater.* **2011**, *138*, 63-67.

(113) Bhusari, D.; Li, J.; Jayachandran, P. J.; Moore, C.; Kohl, P. A. Development of P-doped SiO_2 as proton exchange membrane for microfuel cells. *Electrochem. Solid-State Lett.* **2005**, *8*, A588.

(114) Yang, Y.; Lu, W. Nanoscale resistive switching devices: mechanisms and modeling. *Nanoscale* **2013**, *5*, 10076-10092.

(115) Choi, S.; Tan, S. H.; Li, Z.; Kim, Y.; Choi, C.; Chen, P.-Y.; Yeon, H.; Yu, S.; Kim, J. SiGe epitaxial memory for neuromorphic computing with reproducible high performance based on engineered dislocations. *Nat. Mater.* **2018**, *17*, 335-340.

(116) Zhu, J.; Zhang, T.; Yang, Y.; Huang, R. A comprehensive review on emerging artificial neuromorphic devices. *Appl. Phys. Rev.* **2020**, *7*, 011312.

(117) Liyanagedera, C. M.; Sengupta, A.; Jaiswal, A.; Roy, K. Stochastic spiking neural networks enabled by magnetic tunnel junctions: From nontelegraphic to telegraphic switching regimes. *Phys. Rev. Appl.* **2017**, *8*, 064017.

(118) Huang, Y.; Kang, W.; Zhang, X.; Zhou, Y.; Zhao, W. Magnetic skyrmion-based synaptic devices. *Nanotechnology* **2017**, *28*, 08LT02.

(119) Schneider, M. L.; Donnelly, C. A.; Russek, S. E.; Baek, B.; Pufall, M. R.; Hopkins, P. F.; Dresselhaus, P. D.; Benz, S. P.; Rippard, W. H. Ultralow power artificial synapses using nanotextured magnetic Josephson junctions. *Sci. Adv.* **2018**, *4*, e1701329.

(120) Slonczewski, J. C. Current-driven excitation of magnetic multilayers. *J. Magn. Magn. Mater.* **1996**, *159*, L1-L7.

(121) Berger, L. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B* **1996**, *54*, 9353.

(122) Ikegawa, S.; Mancoff, F. B.; Aggarwal, S. Commercialization of mram–historical and future perspective. In *2021 IEEE International Interconnect Technology Conference (IITC)*, 2021; IEEE: pp 1-3.

(123) Kan, J. J.; Park, C.; Ching, C.; Ahn, J.; Xie, Y.; Pakala, M.; Kang, S. H. A study on practically unlimited endurance of STT-MRAM. *IEEE Trans. Electron Devices* **2017**, *64*, 3639-3646.

(124) Min, T.; Kar, G. S.; Swerts, J.; Mertens, S.; Coseman, S.; Bekaert, J.; Xu, K.; Souriau, L.; Radisic, D.; Okuyama, H. Status and outlook of STT-MRAM development. *Spintronics VII* **2014**, *9167*, 187-194.

(125) Jinnai, B.; Watanabe, K.; Fukami, S.; Ohno, H. Scaling magnetic tunnel junction down to single-digit nanometers—Challenges and prospects. *Appl. Phys. Lett.* **2020**, *116*, 160501.

(126) Jung, S.; Lee, H.; Myung, S.; Kim, H.; Yoon, S. K.; Kwon, S.-W.; Ju, Y.; Kim, M.; Yi, W.; Han, S. A crossbar array of magnetoresistive memory devices for in-memory computing. *Nature* **2022**, *601*, 211-216.

(127) Rzeszut, P.; Chęciński, J.; Brzozowski, I.; Ziętek, S.; Skowroński, W.; Stobiecki, T. Multi-state MRAM cells for hardware neuromorphic computing. *Sci. Rep.* **2022**, *12*, 7178.

(128) Sangwan, V. K.; Hersam, M. C. Electronic transport in two-dimensional materials. *Annu. Rev. Phys. Chem.* **2018**, *69*, 299-325.

(129) Azizi, A.; Zou, X.; Ercius, P.; Zhang, Z.; Elías, A. L.; Perea-López, N.; Stone, G.; Terrones, M.; Yakobson, B. I.; Alem, N. Dislocation motion and grain boundary migration in two-dimensional tungsten disulphide. *Nat. Commun.* **2014**, *5*, 4867.

(130) Yu, Z. G.; Zhang, Y.-W.; Yakobson, B. I. An anomalous formation pathway for dislocation-sulfur vacancy complexes in polycrystalline monolayer MoS_2. *Nano Lett.* **2015**, *15*, 6855-6861.

(131) Sangwan, V. K.; Hersam, M. C. Neuromorphic nanoelectronic materials. *Nat. Nanotechnol.* **2020**, *15*, 517-528.

(132) Sangwan, V. K.; Jariwala, D.; Kim, I. S.; Chen, K.-S.; Marks, T. J.; Lauhon, L. J.; Hersam, M. C. Gate-tunable memristive phenomena mediated by grain boundaries in single-layer MoS_2. *Nat. Nanotechnol.* **2015**, *10*, 403-406.

(133) Sangwan, V. K.; Lee, H.-S.; Bergeron, H.; Balla, I.; Beck, M. E.; Chen, K.-S.; Hersam, M. C. Multi-terminal memtransistors from polycrystalline monolayer molybdenum disulfide. *Nature* **2018**, *554*, 500-504.

(134) Jadwiszczak, J.; Keane, D.; Maguire, P.; Cullen, C. P.; Zhou, Y.; Song, H.; Downing, C.; Fox, D.; McEvoy, N.; Zhu, R. MoS_2 memtransistors fabricated by localized helium ion beam irradiation. *ACS Nano* **2019**, *13*, 14262-14273.

(135) Wang, L.; Liao, W.; Wong, S. L.; Yu, Z. G.; Li, S.; Lim, Y. F.; Feng, X.; Tan, W. C.; Huang, X.; Chen, L. Artificial synapses based on multiterminal memtransistors for neuromorphic application. *Adv. Funct. Mater.* **2019**, *29*, 1901106.

(136) Yuan, J.; Liu, S. E.; Shylendra, A.; Gaviria Rojas, W. A.; Guo, S.; Bergeron, H.; Li, S.; Lee, H.-S.; Nasrin, S.; Sangwan, V. K. Reconfigurable MoS_2 memtransistors for continuous learning in spiking neural networks. *Nano Lett.* **2021**, *21*, 6432-6440.

(137) Lee, H. S.; Sangwan, V. K.; Rojas, W. A. G.; Bergeron, H.; Jeong, H. Y.; Yuan, J.; Su, K.; Hersam, M. C. Dual-Gated MoS_2 Memtransistor Crossbar Array. *Adv. Funct. Mater.* **2020**, *30*, 2003683.

(138) Yan, X.; Qian, J. H.; Sangwan, V. K.; Hersam, M. C. Progress and challenges for memtransistors in neuromorphic circuits and systems. *Adv. Mater.* **2022**, *34*, 2108025.

(139) Sangwan, V. K.; Liu, S. E.; Trivedi, A. R.; Hersam, M. C. Two-dimensional materials for bio-realistic neuronal computing networks. *Matter* **2022**, *5*, 4133-4152.

(140) Seo, S.; Kim, B.; Kim, D.; Park, S.; Kim, T. R.; Park, J.; Jeong, H.; Park, S.-O.; Park, T.; Shin, H. The gate injection-based field-effect synapse transistor with linear conductance update for online training. *Nat. Commun.* **2022**, *13*, 6431.

(141) Chua, L. O. Local activity is the origin of complexity. *Int. J. Bifurc. Chaos* **2005**, *15*, 3435-3456.

(142) Mainzer, K.; Chua, L. *Local Activity Principle: The Cause of Complexity and Symmetry Breaking*; Imperial College Press, **2013**.

(143) Chua, L.; Ascoli, A.; Tetzlaff, R. *Edge of Chaos: The Elan Vital of Complex Phenomena*. https://cmc-dresden.org/media/edge-of-chaos-the-elan-vital-of-complex-phenomena/ (accessed 05-27-2023).

(144) Chua, L.; Sbitnev, V.; Kim, H. Neurons are poised near the edge of chaos. *Int. J. Bifurc. Chaos* **2012**, *22*, 1250098.

(145) Ascoli, A.; Slesazeck, S.; Mähne, H.; Tetzlaff, R.; Mikolajick, T. Nonlinear dynamics of a locally-active memristor. *IEEE Trans. Circuits Syst. I: Reg. Papers* **2015**, *62*, 1165-1174.

(146) Ascoli, A.; Demirkol, A. S.; Tetzlaff, R.; Slesazeck, S.; Mikolajick, T.; Chua, L. O. On local activity and edge of chaos in a NaMLab memristor. *Front. Neurosci.* **2021**, *15*, 651452.

(147) Messaris, I.; Brown, T. D.; Demirkol, A. S.; Ascoli, A.; Al Chawa, M. M.; Williams, R. S.; Tetzlaff, R.; Chua, L. O. NbO_2-Mott memristor: A circuit-theoretic investigation. *IEEE Trans. Circuits Syst. I: Reg. Papers* **2021**, *68*, 4979-4992.

(148) Ascoli, A.; Demirkol, A. S.; Tetzlaff, R.; Chua, L. Edge of chaos theory resolves smale paradox. *IEEE Trans. Circuits Syst. I: Reg. Papers* **2022**, *69*, 1252-1265.

(149) Smale, S. A Mathematical Model of Two Cells via Turing's Equation. In *Lectures in American Mathematical Society*, 1974; Vol. 6, pp 15-26.

(150) Pogromsky, A.; Glad, T.; Nijmeijer, H. On diffusion driven oscillations in coupled dynamical systems. *Int. J. Bifurc. Chaos* **1999**, *9*, 629-644.

(151) Wiggins, S.; Wiggins, S.; Golubitsky, M. *Introduction to applied nonlinear dynamical systems and chaos*; 2nd ed.; Springer Newyork, **2003**.

(152) Andronov, A. A.; Chaikin, C. E.; Khaĭkin, S. Ė.; Khaikin, S. E. *Theory of oscillations*; Princeton University Press, **1949**.

(153) Ascoli, A.; Demirkol, A.; Schmitt, N.; Tetzlaff, R.; Chua, L. Edge of chaos behind bistability of the inhomogeneous in homogeneous cellular media. In *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, 2022; IEEE: pp 193-198.

(154) Turing, A. M. The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond.* **1952**, *237*, 37-72.

(155) Kondo, S.; Asai, R. A reaction–diffusion wave on the skin of the marine angelfish Pomacanthus. *Nature* **1995**, *376*, 765-768.

(156) Ascoli, A.; Demirkol, A. S.; Tetzlaff, R.; Chua, L. Edge of Chaos is sine Qua non for Turing instability. *IEEE Trans. Circuits Syst. I: Reg. Papers* **2022**, *69*, 4596-4609.

(157) Demirkol, A. S.; Ascoli, A.; Messaris, I.; Tetzlaff, R. Pattern formation dynamics in a Memristor Cellular Nonlinear Network structure with a numerically stable VO_2 memristor model. *Jpn. J. Appl. Phys.* **2022**, *61*, SM0807.

(158) Messaris, I.; Ascoli, A.; Demirkol, A. S.; Tetzlaff, R. High Frequency Response of Non-Volatile Memristors. *IEEE Trans. Circuits Syst. I: Reg. Papers* **2022**, *70*, 566-578.

(159) Messaris, I.; Ascoli, A.; Demirkol, A.; Ntinas, V.; Tetzlaff, R. Analytical Study of the Fading Memory Phenomenon in a TaO_x Memristor Model. In *2022 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 2022; IEEE: pp 1-4.

(160) Strachan, J. P.; Torrezan, A. C.; Miao, F.; Pickett, M. D.; Yang, J. J.; Yi, W.; Medeiros-Ribeiro, G.; Williams, R. S. State dynamics and modeling of tantalum oxide memristors. *IEEE Trans. Electron Devices* **2013**, *60*, 2194-2202.

(161) Chua, L. Everything you wish to know about memristors but are afraid to ask. In *Handbook of Memristor Networks*, 1st ed.; Vol. 24; Springer Cham, 2019; DOI: 10.1007/978-3-319-76375-0_3.

(162) Ascoli, A.; Tetzlaff, R.; Chua, L. O.; Strachan, J. P.; Williams, R. S. History erase effect in a non-volatile memristor. *IEEE Trans. Circuits Syst. I: Reg. Papers* **2016**, *63*, 389-400.

(163) Ascoli, A.; Menzel, S.; Rana, V.; Kempen, T.; Messaris, I.; Demirkol, A. S.; Schulten, M.; Siemon, A.; Tetzlaff, R. A Deep Study of Resistance Switching Phenomena in TaO_x ReRAM Cells: System-Theoretic Dynamic Route Map Analysis and Experimental Verification. *Adv. Electron. Mater.* **2022**, *8*, 2200182.

(164) Ascoli, A.; Tetzlaff, R.; Menzel, S. Exploring the dynamics of real-world memristors on the basis of circuit theoretic model predictions. *IEEE Circuits Syst. Mag.* **2018**, *18*, 48-76.

(165) Hardtdegen, A.; La Torre, C.; Cüppers, F.; Menzel, S.; Waser, R.; Hoffmann-Eifert, S. Improved switching stability and the effect of an internal series resistor in HfO_2/TiO_x bilayer ReRAM cells. *IEEE Trans. Electron Devices* **2018**, *65*, 3229-3236.

(166) Messaris, I.; Serb, A.; Stathopoulos, S.; Khiat, A.; Nikolaidis, S.; Prodromakis, T. A data-driven verilog-a reram model. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2018**, *37*, 3151-3162.

(167) Ntinas, V.; Ascoli, A.; Messaris, I.; Wang, Y.; Rana, V.; Menzel, S.; Tetzlaff, R. Toward simplified physics-based memristor modeling of valence change mechanism devices. *IEEE Trans. Circuits Syst. II: Express Br.* **2022**, *69*, 2473-2477.

(168) Misra, J.; Saha, I. Artificial neural networks in hardware: A survey of two decades of progress. *Neurocomputing* **2010**, *74*, 239-255.

(169) Gokmen, T.; Vlasov, Y. Acceleration of deep neural network training with resistive cross-point devices: Design considerations. *Front. Neurosci.* **2016**, *10*, 333.

(170) Burr, G. W.; Narayanan, P.; Shelby, R. M.; Sidler, S.; Boybat, I.; Di Nolfo, C.; Leblebici, Y. Large-scale neural networks implemented with non-volatile memory as the synaptic weight element: Comparative performance analysis (accuracy, speed, and power). In *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015; IEEE: pp 4.4. 1-4.4. 4.

(171) Steinbuch, K. Die lernmatrix. *Kybernetik* **1961**, *1*, 36-45.

(172) Burr, G. W.; Shenoy, R. S.; Virwani, K.; Narayanan, P.; Padilla, A.; Kurdi, B.; Hwang, H. Access devices for 3D crosspoint memory. *J Vac. Sci. Technol. B Nanotechnol. Microelectron: Mater. Process. Meas. Phenom.* **2014**, *32*, 040802.

(173) Li, Y.; Kim, S.; Sun, X.; Solomon, P.; Gokmen, T.; Tsai, H.; Koswatta, S.; Ren, Z.; Mo, R.; Yeh, C. C. Capacitor-based cross-point array for analog neural network with record symmetry and linearity. In *2018 IEEE Symposium on VLSI Technology*, 2018; IEEE: pp 25-26.

(174) Chen, P.-Y.; Peng, X.; Yu, S. NeuroSim+: An integrated device-to-algorithm framework for benchmarking synaptic devices and array architectures. In *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017; IEEE: pp 6.1. 1-6.1. 4.

(175) Roy, S.; Sridharan, S.; Jain, S.; Raghunathan, A. Txsim: Modeling training of deep neural networks on resistive crossbar systems. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **2021**, *29*, 730-738.

(176) Chen, P.-Y.; Lin, B.; Wang, I.-T.; Hou, T.-H.; Ye, J.; Vrudhula, S.; Seo, J.-s.; Cao, Y.; Yu, S. Mitigating effects of non-ideal synaptic device characteristics for on-chip learning. In *2015 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2015; IEEE: pp 194-199.

(177) Burr, G. W.; Shelby, R. M.; Sidler, S.; Di Nolfo, C.; Jang, J.; Boybat, I.; Shenoy, R. S.; Narayanan, P.; Virwani, K.; Giacometti, E. U. Experimental demonstration and tolerancing of a large-scale neural network (165 000 synapses) using phase-change memory as the synaptic weight element. *IEEE Trans. Electron Devices* **2015**, *62*, 3498-3507.

(178) Agarwal, S.; Plimpton, S. J.; Hughart, D. R.; Hsia, A. H.; Richter, I.; Cox, J. A.; James, C. D.; Marinella, M. J. Resistive memory device requirements for a neural algorithm accelerator. In *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016; IEEE: pp 929-938.

(179) Sun, X.; Yu, S. Impact of non-ideal characteristics of resistive synaptic devices on implementing convolutional neural networks. *IEEE J. Emerg. Sel. Top. Circuits* **2019**, *9*, 570-579.

(180) Chang, T.; Jo, S.-H.; Lu, W. Short-term memory to long-term memory transition in a nanoscale memristor. *ACS Nano* **2011**, *5*, 7669-7676.

(181) Yu, S.; Wu, Y.; Jeyasingh, R.; Kuzum, D.; Wong, H.-S. P. An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation. *IEEE Trans. Electron Devices* **2011**, *58*, 2729-2737.

(182) Wu, Y.; Yu, S.; Wong, H.-S. P.; Chen, Y.-S.; Lee, H.-Y.; Wang, S.-M.; Gu, P.-Y.; Chen, F.; Tsai, M.-J. AlO_x-based resistive switching device with gradual resistance modulation for

neuromorphic device application. In *2012 4th IEEE International Memory Workshop*, 2012; IEEE: pp 1-4.

(183) Yu, S.; Gao, B.; Fang, Z.; Yu, H.; Kang, J.; Wong, H. S. P. A low energy oxide-based electronic synaptic device for neuromorphic visual systems with tolerance to device variation. *Adv. Mater.* **2013**, *25*, 1774-1779.

(184) Ohno, T.; Hasegawa, T.; Tsuruoka, T.; Terabe, K.; Gimzewski, J. K.; Aono, M. Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. *Nat. Mater.* **2011**, *10*, 591-595.

(185) Song, M.-K.; Namgung, S. D.; Lee, H.; Yoon, J. H.; Song, Y.-W.; Cho, K. H.; Lee, Y.-S.; Lee, J.-S.; Nam, K. T.; Kwon, J.-Y. Tyrosine-mediated analog resistive switching for artificial neural networks. *Nano Res.* **2023**, *16*, 858-864.

(186) Kuzum, D.; Jeyasingh, R. G.; Lee, B.; Wong, H.-S. P. Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. *Nano Lett.* **2012**, *12*, 2179-2186.

(187) Suri, M.; Bichler, O.; Querlioz, D.; Cueto, O.; Perniola, L.; Sousa, V.; Vuillaume, D.; Gamrat, C.; DeSalvo, B. Phase change memory as synapse for ultra-dense neuromorphic systems: Application to complex visual pattern extraction. In *2011 International Electron Devices Meeting*, 2011; IEEE: pp 4.4. 1-4.4. 4.

(188) Nishitani, Y.; Kaneko, Y.; Ueda, M.; Morie, T.; Fujii, E. Three-terminal ferroelectric synapse device with concurrent learning function for artificial neural networks. *J. Appl. Phys.* **2012**, *111*, 124108.

(189) Kim, S.; Gokmen, T.; Lee, H.-M.; Haensch, W. E. Analog CMOS-based resistive processing unit for deep neural network training. In *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2017; IEEE: pp 422-425.

(190) Kohda, Y.; Li, Y.; Hosokawa, K.; Kim, S.; Khaddam-Aljameh, R.; Ren, Z.; Solomon, P.; Gokmen, T.; Rajalingam, S.; Baks, C. Unassisted true analog neural network training chip. In *2020 IEEE International Electron Devices Meeting (IEDM)*, 2020; IEEE: pp 36.32. 31-36.32. 34.

(191) Chang, C.-C.; Liu, J.-C.; Shen, Y.-L.; Chou, T.; Chen, P.-C.; Wang, I.-T.; Su, C.-C.; Wu, M.-H.; Hudec, B.; Chang, C.-C. Challenges and opportunities toward online training acceleration using RRAM-based hardware neural network. In *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017; IEEE: pp 11.16. 11-11.16. 14.

(192) Gokmen, T.; Haensch, W. Algorithm for training neural networks on resistive device arrays. *Front. Neurosci.* **2020**, *14*, 103.

(193) Lee, C.; Noh, K.; Ji, W.; Gokmen, T.; Kim, S. Impact of Asymmetric Weight Update on Neural Network Training With Tiki-Taka Algorithm. *Front. Neurosci.* **2022**, 1554.

(194) Gokmen, T. Enabling training of neural networks on noisy hardware. *Front. Artif. Intell.* **2021**, *4*, 699148.

(195) Onen, M.; Gokmen, T.; Todorov, T. K.; Nowicki, T.; Del Alamo, J. A.; Rozen, J.; Haensch, W.; Kim, S. Neural Network Training with Asymmetric Crosspoint Elements. *Front. Artif. Intell.* **2022**, *5*, 891624.

(196) Gong, N.; Rasch, M.; Seo, S.-C.; Gasasira, A.; Solomon, P.; Bragaglia, V.; Consiglio, S.; Higuchi, H.; Park, C.; Brew, K. Deep learning acceleration in 14nm CMOS compatible ReRAM

array: device, material and algorithm co-optimization. In 2022 International Electron Devices Meeting (IEDM), San Francisco, CA, USA; 2022.

(197) Ambrogio, S.; Narayanan, P.; Tsai, H.; Shelby, R. M.; Boybat, I.; Di Nolfo, C.; Sidler, S.; Giordano, M.; Bodini, M.; Farinha, N. C. Equivalent-accuracy accelerated neural-network training using analogue memory. *Nature* **2018**, *558*, 60-67.

(198) Shim, W.; Meng, J.; Peng, X.; Seo, J.-s.; Yu, S. Impact of multilevel retention characteristics on RRAM based DNN inference engine. In *2021 IEEE International Reliability Physics Symposium (IRPS)*, 2021; IEEE: pp 1-4.

(199) Xiang, Y.; Huang, P.; Zhao, Y.; Zhao, M.; Gao, B.; Wu, H.; Qian, H.; Liu, X.; Kang, J. Impacts of state instability and retention failure of filamentary analog RRAM on the performance of deep neural network. *IEEE Trans. Electron Devices* **2019**, *66*, 4517-4522.

(200) Wen, W.; Wu, C.; Wang, Y.; Chen, Y.; Li, H. Learning structured sparsity in deep neural networks. In *Advances in Neural Information Processing Systems (NIPS 2016)*, Barcelona, Spain, 2016; Vol. 29, pp 2074-2082.

(201) Yang, L.; He, Z.; Fan, D. Harmonious coexistence of structured weight pruning and ternarization for deep neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020; Vol. 34, pp 6623-6630.

(202) Meng, J.; Yang, L.; Peng, X.; Yu, S.; Fan, D.; Seo, J.-S. Structured pruning of RRAM crossbars for efficient in-memory computing acceleration of deep neural networks. *IEEE Trans. Circuits Syst. II: Express Br.* **2021**, *68*, 1576-1580.

(203) Peng, X.; Huang, S.; Luo, Y.; Sun, X.; Yu, S. DNN+ NeuroSim: An end-to-end benchmarking framework for compute-in-memory accelerators with versatile device technologies. In *2019 IEEE international electron devices meeting (IEDM)*, 2019; IEEE: pp 32.35. 31-32.35. 34.

(204) Yin, S.; Sun, X.; Yu, S.; Seo, J.-s. High-throughput in-memory computing for binary deep neural networks with monolithically integrated RRAM and 90-nm CMOS. *IEEE Trans. Electron Devices* **2020**, *67*, 4185-4192.

(205) Meng, J.; Yeo, I.; Shim, W.; Yang, L.; Fan, D.; Yu, S.; Seo, J.-S. Sparse and Robust RRAM-based Efficient In-memory Computing for DNN Inference. In *2022 IEEE International Reliability Physics Symposium (IRPS)*, 2022; IEEE: pp 3C. 1-1-3C. 1-6.

(206) Meng, J.; Shim, W.; Yang, L.; Yeo, I.; Fan, D.; Yu, S.; Seo, J.-s. Temperature-resilient rram-based in-memory computing for dnn inference. *IEEE Micro* **2021**, *42*, 89-98.

(207) Cai, F.; Kumar, S.; Van Vaerenbergh, T.; Sheng, X.; Liu, R.; Li, C.; Liu, Z.; Foltin, M.; Yu, S.; Xia, Q. Power-efficient combinatorial optimization using intrinsic noise in memristor Hopfield neural networks. *Nat. Electron.* **2020**, *3*, 409-418.

(208) Mahmoodi, M.; Prezioso, M.; Strukov, D. Versatile stochastic dot product circuits based on nonvolatile memories for high performance neurocomputing and neurooptimization. *Nat. Commun.* **2019**, *10*, 5113.

(209) Yang, K.; Duan, Q.; Wang, Y.; Zhang, T.; Yang, Y.; Huang, R. Transiently chaotic simulated annealing based on intrinsic nonlinearity of memristors for efficient solution of optimization problems. *Sci. Adv.* **2020**, *6*, eaba9901.

(210) Wang, Z.; Joshi, S.; Savel'ev, S. E.; Jiang, H.; Midya, R.; Lin, P.; Hu, M.; Ge, N.; Strachan, J. P.; Li, Z. Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nat. Mater.* **2017**, *16*, 101-108.

(211) Wang, Z.; Joshi, S.; Savel'ev, S.; Song, W.; Midya, R.; Li, Y.; Rao, M.; Yan, P.; Asapu, S.; Zhuo, Y. Fully memristive neural networks for pattern classification with unsupervised learning. *Nat. Electron.* **2018**, *1*, 137-145.

(212) Wang, S.; Li, Y.; Wang, D.; Zhang, W.; Chen, X.; Dong, D.; Wang, S.; Zhang, X.; Lin, P.; Gallicchio, C. Echo state graph neural networks with analogue random resistive memory arrays. *Nat. Mach. Intell.* **2023**, *5*, 104-113.

(213) Zhang, W.; Wang, S.; Li, Y.; Xu, X.; Dong, D.; Jiang, N.; Wang, F.; Guo, Z.; Fang, R.; Dou, C. Few-shot graph learning with robust and energy-efficient memory-augmented graph neural network (MAGNN) based on homogeneous computing-in-memory. In *2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, 2022; IEEE: pp 224-225.

(214) Wang, S.; Chen, H.; Zhang, W.; Li, Y.; Wang, D.; Shi, S.; Zhao, Y.; Loong, K. C.; Chen, X.; Dong, Y. Convolutional Echo-State Network with Random Memristors for Spatiotemporal Signal Classification. *Adv. Intell. Syst.* **2022**, *4*, 2200027.

(215) Dalgaty, T.; Castellani, N.; Turck, C.; Harabi, K.-E.; Querlioz, D.; Vianello, E. In situ learning using intrinsic memristor variability via Markov chain Monte Carlo sampling. *Nat. Electron.* **2021**, *4*, 151-161.

(216) Nili, H.; Adam, G. C.; Hoskins, B.; Prezioso, M.; Kim, J.; Mahmoodi, M. R.; Bayat, F. M.; Kavehei, O.; Strukov, D. B. Hardware-intrinsic security primitives enabled by analogue state and nonlinear conductance variations in integrated memristors. *Nat. Electron.* **2018**, *1*, 197-202.

(217) Karunaratne, G.; Le Gallo, M.; Cherubini, G.; Benini, L.; Rahimi, A.; Sebastian, A. In-memory hyperdimensional computing. *Nat. Electron.* **2020**, *3*, 327-337.

(218) Karunaratne, G.; Schmuck, M.; Le Gallo, M.; Cherubini, G.; Benini, L.; Sebastian, A.; Rahimi, A. Robust high-dimensional memory-augmented neural networks. *Nat. Commun.* **2021**, *12*, 2468.

(219) Le Gallo, M.; Sebastian, A.; Mathis, R.; Manica, M.; Giefers, H.; Tuma, T.; Bekas, C.; Curioni, A.; Eleftheriou, E. Mixed-precision in-memory computing. *Nat. Electron.* **2018**, *1*, 246-253.

(220) Zidan, M. A.; Jeong, Y.; Lee, J.; Chen, B.; Huang, S.; Kushner, M. J.; Lu, W. D. A general memristor-based partial differential equation solver. *Nat. Electron.* **2018**, *1*, 411-420.

(221) Reuther, A.; Michaleas, P.; Jones, M.; Gadepally, V.; Samsi, S.; Kepner, J. AI and ML Accelerator Survey and Trends. In *2022 IEEE High Performance Extreme Computing Conference (HPEC)*, 2022; IEEE: pp 1-10.

(222) Murmann, B. Mixed-signal computing for deep neural network inference. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **2020**, *29*, 3-13.

(223) Sebastian, A.; Le Gallo, M.; Khaddam-Aljameh, R.; Eleftheriou, E. Memory devices and applications for in-memory computing. *Nat. Nanotechnol.* **2020**, *15*, 529-544.

(224) Narayanan, P.; Ambrogio, S.; Okazaki, A.; Hosokawa, K.; Tsai, H.; Nomura, A.; Yasuda, T.; Mackin, C.; Lewis, S. C.; Friz, A. Fully on-chip MAC at 14nm enabled by accurate row-wise programming of PCM-based weights and parallel vector-transport in duration-format. In *2021 Symposium on VLSI Technology*, 2021; IEEE: pp 1-2.

(225) Wan, W.; Kubendran, R.; Schaefer, C.; Eryilmaz, S. B.; Zhang, W.; Wu, D.; Deiss, S.; Raina, P.; Qian, H.; Gao, B. A compute-in-memory chip based on resistive random-access memory. *Nature* **2022**, *608*, 504-512.

(226) Fick, L.; Skrzyniarz, S.; Parikh, M.; Henry, M. B.; Fick, D. Analog matrix processor for edge AI real-time video analytics. In *2022 IEEE International Solid-State Circuits Conference (ISSCC)*, 2022; IEEE: Vol. 65, pp 260-262.

(227) Khaddam-Aljameh, R.; Stanisavljevic, M.; Mas, J. F.; Karunaratne, G.; Brändli, M.; Liu, F.; Singh, A.; Müller, S. M.; Egger, U.; Petropoulos, A. HERMES-Core—A 1.59-TOPS/mm^2 PCM on 14-nm CMOS in-memory compute core using 300-ps/LSB linearized CCO-based ADCs. *IEEE J. Solid-State Circuits* **2022**, *57*, 1027-1038.

(228) Gallo, M. L.; Khaddam-Aljameh, R.; Stanisavljevic, M.; Vasilopoulos, A.; Kersting, B.; Dazzi, M.; Karunaratne, G.; Braendli, M.; Singh, A.; Mueller, S. M. A 64-core mixed-signal in-memory compute chip based on phase-change memory for deep neural network inference. **2022**, 2212.02872, ver.1, *arXiv (Emerging Technologies)*. DOI: 10.48550/arXiv.2212.02872 (accessed 05-27-2023).

(229) Jain, S.; Tsai, H.; Chen, C.-T.; Muralidhar, R.; Boybat, I.; Frank, M. M.; Woźniak, S.; Stanisavljevic, M.; Adusumilli, P.; Narayanan, P. A Heterogeneous and Programmable Compute-

In-Memory Accelerator Architecture for Analog-AI Using Dense 2-D Mesh. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **2022**, *31*, 114-127.

(230) Kösters, D. J.; Kortman, B. A.; Boybat, I.; Ferro, E.; Dolas, S.; Ruiz de Austri, R.; Kwisthout, J.; Hilgenkamp, H.; Rasing, T.; Riel, H. Benchmarking energy consumption and latency for neuromorphic computing in condensed matter and particle physics. *APL Mach. Learn.* **2023**, *1*, 016101.

(231) Garofalo, A.; Ottavi, G.; Conti, F.; Karunaratne, G.; Boybat, I.; Benini, L.; Rossi, D. A heterogeneous in-memory computing cluster for flexible end-to-end inference of real-world deep neural networks. *IEEE J. Emerg. Sel. Top. Circuits* **2022**, *12*, 422-435.

(232) Klein, J.; Boybat, I.; Qureshi, Y.; Dazzi, M.; Levisse, A.; Ansaloni, G.; Zapater, M.; Sebastian, A.; Atienza, D. ALPINE: Analog In-Memory Acceleration with Tight Processor Integration for Deep Learning. *IEEE Trans. Comput.* **2022**, Early access. DOI: 10.1109/TC.2022.3230285 (accessed 2023-05-27).

(233) Ghazi Sarwat, S.; Philip, T. M.; Chen, C. T.; Kersting, B.; Bruce, R. L.; Cheng, C. W.; Li, N.; Saulnier, N.; BrightSky, M.; Sebastian, A. Projected Mushroom Type Phase-Change Memory. *Adv. Funct. Mater.* **2021**, *31*, 2106547.

(234) Büchel, J.; Vasilopoulos, A.; Kersting, B.; Odermatt, F.; Brew, K.; Ok, I.; Choi, S.; Saraf, I.; Chan, V.; Philip, T. Gradient descent-based programming of analog in-memory computing cores. In *2022 International Electron Devices Meeting (IEDM)*, 2022; IEEE: pp 33.31. 31-33.31. 34.

(235) Boybat, I.; Le Gallo, M.; Nandakumar, S.; Moraitis, T.; Parnell, T.; Tuma, T.; Rajendran, B.; Leblebici, Y.; Sebastian, A.; Eleftheriou, E. Neuromorphic computing with multi-memristive synapses. *Nat. Commun.* **2018**, *9*, 2514.

(236) Khwa, W.-S.; Chiu, Y.-C.; Jhang, C.-J.; Huang, S.-P.; Lee, C.-Y.; Wen, T.-H.; Chang, F.-C.; Yu, S.-M.; Lee, T.-Y.; Chang, M.-F. A 40-nm, 2M-cell, 8b-precision, hybrid SLC-MLC PCM computing-in-memory macro with 20.5-65.0 TOPS/W for tiny-Al edge devices. In *2022 IEEE International Solid-State Circuits Conference (ISSCC)*, 2022; IEEE: Vol. 65, pp 1-3.

(237) Sheridan, P. M.; Cai, F.; Du, C.; Ma, W.; Zhang, Z.; Lu, W. D. Sparse coding with memristor networks. *Nat. Nanotechnol.* **2017**, *12*, 784-789.

(238) Yao, P.; Wu, H.; Gao, B.; Eryilmaz, S. B.; Huang, X.; Zhang, W.; Zhang, Q.; Deng, N.; Shi, L.; Wong, H.-S. P. Face classification using electronic synapses. *Nat. Commun.* **2017**, *8*, 15199.

(239) Li, C.; Hu, M.; Li, Y.; Jiang, H.; Ge, N.; Montgomery, E.; Zhang, J.; Song, W.; Dávila, N.; Graves, C. E. Analogue signal and image processing with large memristor crossbars. *Nat. Electron.* **2018**, *1*, 52-59.

(240) Liu, Q.; Gao, B.; Yao, P.; Wu, D.; Chen, J.; Pang, Y.; Zhang, W.; Liao, Y.; Xue, C.-X.; Chen, W.-H. 33.2 A fully integrated analog ReRAM based 78.4 TOPS/W compute-in-memory chip with fully parallel MAC computing. In *2020 ieee international solid-state circuits conference-(isscc)*, 2020; IEEE: pp 500-502.

(241) Yao, P.; Wu, H.; Gao, B.; Tang, J.; Zhang, Q.; Zhang, W.; Yang, J. J.; Qian, H. Fully hardware-implemented memristor convolutional neural network. *Nature* **2020**, *577*, 641-646.

(242) Jiang, Y.; Gao, B.; Tang, J.; Wu, D.; He, H.; Qian, H.; Wu, H. HARNS: High-level architectural model of RRAM based computing-in-memory NPU. In *2021 IEEE International Conference on Integrated Circuits, Technologies and Applications (ICTA)*, 2021; IEEE: pp 35-36.

(243) Lee, M. K. F.; Cui, Y.; Somu, T.; Luo, T.; Zhou, J.; Tang, W. T.; Wong, W.-F.; Goh, R. S. M. A system-level simulator for RRAM-based neuromorphic computing chips. *ACM Trans. Archit. Code Optim.* **2019**, *15*, 1-24.

(244) Zhang, W.; Peng, X.; Wu, H.; Gao, B.; He, H.; Zhang, Y.; Yu, S.; Qian, H. Design guidelines of RRAM based neural-processing-unit: A joint device-circuit-algorithm analysis. In *Proceedings of the 56th Annual Design Automation Conference 2019*, 2019; pp 1-6.

(245) Liu, Y.; Zhao, M.; Gao, B.; Hu, R.; Zhang, W.; Yang, S.; Yao, P.; Xu, F.; Xi, Y.; Zhang, Q. Compact reliability model of analog rram for computation-in-memory device-to-system codesign and benchmark. *IEEE Trans. Electron Devices* **2021**, *68*, 2686-2692.

(246) Gao, B.; Wu, H.; Wu, W.; Wang, X.; Yao, P.; Xi, Y.; Zhang, W.; Deng, N.; Huang, P.; Liu, X. Modeling disorder effect of the oxygen vacancy distribution in filamentary analog RRAM for neuromorphic computing. In *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017; IEEE: pp 4.4. 1-4.4. 4.

(247) Ambrosi, J.; Ankit, A.; Antunes, R.; Chalamalasetti, S. R.; Chatterjee, S.; El Hajj, I.; Fachini, G.; Faraboschi, P.; Foltin, M.; Huang, S. Hardware-software co-design for an analog-digital accelerator for machine learning. In *2018 IEEE International Conference on Rebooting Computing (ICRC)*, 2018; IEEE: pp 1-13.

(248) Zhang, Y.; He, G.; Wang, G.; Li, Y. Efficient and robust RRAM-based convolutional weight mapping with shifted and duplicated kernel. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2020**, *40*, 287-300.

(249) Peng, X.; Liu, R.; Yu, S. Optimizing weight mapping and data flow for convolutional neural networks on processing-in-memory architectures. *IEEE Trans. Circuits Syst. I: Reg. Papers* **2019**, *67*, 1333-1343.

(250) Govoreanu, B.; Redolfi, A.; Zhang, L.; Adelmann, C.; Popovici, M.; Clima, S.; Hody, H.; Paraschiv, V.; Radu, I.; Franquet, A. Vacancy-modulated conductive oxide resistive RAM (VMCO-RRAM): An area-scalable switching current, self-compliant, highly nonlinear and wide on/off-window resistive switching cell. In *2013 IEEE International Electron Devices Meeting*, 2013; IEEE: pp 10.12. 11-10.12. 14.

(251) Kim, H.; Mahmoodi, M.; Nili, H.; Strukov, D. B. 4K-memristor analog-grade passive crossbar circuit. *Nat. Commun.* **2021**, *12*, 5198.

(252) Chevallier, C. J.; Siau, C. H.; Lim, S. F.; Namala, S. R.; Matsuoka, M.; Bateman, B. L.; Rinerson, D. A 0.13 μm 64Mb multi-layered conductive metal-oxide memory. In *2010 IEEE International Solid-State Circuits Conference-(ISSCC)*, 2010; IEEE: pp 260-261.

(253) Kawahara, A.; Azuma, R.; Ikeda, Y.; Kawai, K.; Katoh, Y.; Hayakawa, Y.; Tsuji, K.; Yoneda, S.; Himeno, A.; Shimakawa, K. An 8 Mb multi-layered cross-point ReRAM macro with 443 MB/s write throughput. *IEEE J. Solid-State Circuits* **2012**, *48*, 178-185.

(254) Liu, T.-y.; Yan, T. H.; Scheuerlein, R.; Chen, Y.; Lee, J. K.; Balakrishnan, G.; Yee, G.; Zhang, H.; Yap, A.; Ouyang, J. A 130.7-mm^2 2-Layer 32-Gb ReRAM Memory Device in 24-nm Technology. *IEEE J. Solid-State Circuits* **2013**, *49*, 140-153.

(255) Adam, G. C.; Hoskins, B. D.; Prezioso, M.; Merrikh-Bayat, F.; Chakrabarti, B.; Strukov, D. B. 3-D memristor crossbars for analog and neuromorphic computing applications. *IEEE Trans. Electron Devices* **2016**, *64*, 312-318.

(256) Lin, P.; Li, C.; Wang, Z.; Li, Y.; Jiang, H.; Song, W.; Rao, M.; Zhuo, Y.; Upadhyay, N. K.; Barnell, M. Three-dimensional memristor circuits as complex neural networks. *Nat. Electron.* **2020**, *3*, 225-232.

(257) Cai, F.; Correll, J. M.; Lee, S. H.; Lim, Y.; Bothra, V.; Zhang, Z.; Flynn, M. P.; Lu, W. D. A fully integrated reprogrammable memristor–CMOS system for efficient multiply–accumulate operations. *Nat. Electron.* **2019**, *2*, 290-299.

(258) Miyamura, M.; Nakaya, S.; Tada, M.; Sakamoto, T.; Okamoto, K.; Banno, N.; Ishida, S.; Ito, K.; Hada, H.; Sakimura, N. Programmable cell array using rewritable solid-electrolyte switch integrated in 90nm CMOS. In *2011 IEEE International Solid-State Circuits Conference*, 2011; IEEE: pp 228-229.

(259) Kim, K.-H.; Gaba, S.; Wheeler, D.; Cruz-Albrecht, J. M.; Hussain, T.; Srinivasa, N.; Lu, W. A functional hybrid memristor crossbar-array/CMOS system for data storage and neuromorphic applications. *Nano Lett.* **2012**, *12*, 389-395.

(260) Kim, G. H.; Lee, J. H.; Ahn, Y.; Jeon, W.; Song, S. J.; Seok, J. Y.; Yoon, J. H.; Yoon, K. J.; Park, T. J.; Hwang, C. S. 32× 32 crossbar array resistive memory composed of a stacked Schottky diode and unipolar resistive memory. *Adv. Funct. Mater.* **2013**, *23*, 1440-1449.

(261) Kumar, S.; Agarwal, A.; Mukherjee, S. Electrical Performance of Large-Area Y_2O_3 Memristive Crossbar Array With Ultralow C2C Variability. *IEEE Trans. Electron Devices* **2022**, *69*, 3660-3666.

(262) Prezioso, M.; Kataeva, I.; Merrikh-Bayat, F.; Hoskins, B.; Adam, G.; Sota, T.; Likharev, K.; Strukov, D. Modeling and implementation of firing-rate neuromorphic-network classifiers with bilayer Pt/Al_2O_3/TiO_{2-x}/Pt memristors. In *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015; IEEE: pp 17.14. 11-17.14. 14.

(263) Bayat, F. M.; Prezioso, M.; Chakrabarti, B.; Nili, H.; Kataeva, I.; Strukov, D. Implementation of multilayer perceptron network with highly uniform passive memristive crossbar circuits. *Nat. Commun.* **2018**, *9*, 2331.

(264) Kim, T.-H.; Lee, J.; Kim, S.; Park, J.; Park, B.-G.; Kim, H. 3-bit multilevel operation with accurate programming scheme in TiO_x/Al_2O_3 memristor crossbar array for quantized neuromorphic system. *Nanotechnology* **2021**, *32*, 295201.

(265) Kim, S.; Park, J.; Kim, T.-H.; Hong, K.; Hwang, Y.; Park, B.-G.; Kim, H. 4-bit Multilevel Operation in Overshoot Suppressed Al_2O_3/TiO_x Resistive Random-Access Memory Crossbar Array. *Adv. Intell. Syst.* **2022**, *4*, 2100273.

(266) El Mesoudy, A.; Lamri, G.; Dawant, R.; Arias-Zapata, J.; Gliech, P.; Beilliard, Y.; Ecoffey, S.; Ruediger, A.; Alibart, F.; Drouin, D. Fully CMOS-compatible passive TiO_2-based memristor crossbars for in-memory computing. *Microelectron. Eng.* **2022**, *255*, 111706.

(267) Indiveri, G.; Linares-Barranco, B.; Legenstein, R.; Deligeorgis, G.; Prodromakis, T. Integration of nanoscale memristor synapses in neuromorphic computing architectures. *Nanotechnology* **2013**, *24*, 384010.

(268) Chakrabarti, B.; Lastras-Montaño, M. A.; Adam, G.; Prezioso, M.; Hoskins, B.; Payvand, M.; Madhavan, A.; Ghofrani, A.; Theogarajan, L.; Cheng, K.-T. A multiply-add engine with monolithically integrated 3D memristor crossbar/CMOS hybrid circuit. *Sci. Rep.* **2017**, *7*, 42429.

(269) Bavandpour, M.; Mahmoodi, M.; Nili, H.; Bayat, F. M.; Prezioso, M.; Vincent, A.; Strukov, D.; Likharev, K. Mixed-signal neuromorphic inference accelerators: Recent results and future prospects. In *2018 IEEE International Electron Devices Meeting (IEDM)*, 2018; IEEE: pp 20.24. 21-20.24. 24.

(270) Kau, D.; Tang, S.; Karpov, I. V.; Dodge, R.; Klehn, B.; Kalb, J. A.; Strand, J.; Diaz, A.; Leung, N.; Wu, J. A stackable cross point phase change memory. In *2009 IEEE International Electron Devices Meeting (IEDM)*, 2009; IEEE: pp 1-4.

(271) Fuller, E. J.; Keene, S. T.; Melianas, A.; Wang, Z.; Agarwal, S.; Li, Y.; Tuchman, Y.; James, C. D.; Marinella, M. J.; Yang, J. J. Parallel programming of an ionic floating-gate memory array for scalable neuromorphic computing. *Science* **2019**, *364*, 570-574.

(272) Alibart, F.; Zamanidoost, E.; Strukov, D. B. Pattern classification by memristive crossbar circuits using ex situ and in situ training. *Nat. Commun.* **2013**, *4*, 2072.

(273) Lee, J.; Du, C.; Sun, K.; Kioupakis, E.; Lu, W. D. Tuning ionic transport in memristive devices by graphene with engineered nanopores. *ACS Nano* **2016**, *10*, 3571-3579.

(274) Strukov, D. B. Tightening grip. *Nat. Mater.* **2018**, *17*, 293-295.

(275) LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436-444.

(276) Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735-1780.

(277) Koller, D.; Friedman, N. *Probabilistic graphical models: principles and techniques*; MIT press, **2009**.

(278) Wan, W.; Kubendran, R.; Gao, B.; Joshi, S.; Raina, P.; Wu, H.; Cauwenberghs, G.; Wong, H. P. A voltage-mode sensing scheme with differential-row weight mapping for energy-efficient RRAM-based in-memory computing. In *2020 IEEE Symposium on VLSI Technology*, 2020; IEEE: pp 1-2.

(279) Krizhevsky, A. *Learning multiple layers of features from tiny images*; Technical Report TR-2009, University of Toronto, Toronto, ON, Canada, 2009.

(280) LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278-2324.

(281) Warden, P. Speech commands: A dataset for limited-vocabulary speech recognition. **2018**, 1804.03209, ver.1, *arXiv (Computation and Language)*. DOI: 10.48550/arXiv.1804.03209 (accessed 05-27-2023).

(282) Semiconductor Research Corporation. *The Decadal Plan for Semiconductors*. Semiconductor Research Corporation, **2021**. https://www.src.org/about/decadal-plan/decadal-plan-full-report.pdf (accessed 05-27-2023).

(283) Zhou, F.; Chai, Y. Near-sensor and in-sensor computing. *Nat. Electron.* **2020**, *3*, 664-671.

(284) Zhou, F.; Liu, Y.; Shen, X.; Wang, M.; Yuan, F.; Chai, Y. Low-voltage, optoelectronic $CH_3NH_3PbI_{3-x}Cl_x$ memory with integrated sensing and logic operations. *Adv. Funct. Mater.* **2018**, *28*, 1800080.

(285) Zhou, F.; Zhou, Z.; Chen, J.; Choy, T. H.; Wang, J.; Zhang, N.; Lin, Z.; Yu, S.; Kang, J.; Wong, H.-S. P. Optoelectronic resistive random access memory for neuromorphic vision sensors. *Nat. Nanotechnol.* **2019**, *14*, 776-782.

(286) Chai, Y. In-sensor computing for machine vision. *Nature* **2020**, *579*, 32-33.

(287) Choi, D.; Song, M.-K.; Sung, T.; Jang, S.; Kwon, J.-Y. Energy scavenging artificial nervous system for detecting rotational movement. *Nano Energy* **2020**, *74*, 104912.

(288) Mennel, L.; Symonowicz, J.; Wachter, S.; Polyushkin, D. K.; Molina-Mendoza, A. J.; Mueller, T. Ultrafast machine vision with 2D material neural network image sensors. *Nature* **2020**, *579*, 62-66.

(289) Song, M.-K.; Namgung, S. D.; Song, Y.-W.; Sung, T.; Ji, W.; Lee, Y.-S.; Nam, K. T.; Kwon, J.-Y. Fully degradable memristors and humidity sensors based on a tyrosine-rich peptide. *ACS Appl. Electron. Mater.* **2021**, *3*, 3372-3378.

(290) Liao, F.; Zhou, Z.; Kim, B. J.; Chen, J.; Wang, J.; Wan, T.; Zhou, Y.; Hoang, A. T.; Wang, C.; Kang, J. Bioinspired in-sensor visual adaptation for accurate perception. *Nat. Electron.* **2022**, *5*, 84-91.

(291) Song, M.-K.; Lee, H.; Yoon, J. H.; Song, Y.-W.; Namgung, S. D.; Sung, T.; Lee, Y.-S.; Lee, J.-S.; Nam, K. T.; Kwon, J.-Y. Humidity-induced synaptic plasticity of ZnO artificial synapses using peptide insulator for neuromorphic computing. *J. Mater. Sci. Technol.* **2022**, *119*, 150-155.

(292) Namgung, S. D.; Kim, R. M.; Lim, Y.-C.; Lee, J. W.; Cho, N. H.; Kim, H.; Huh, J.-S.; Rhee, H.; Nah, S.; Song, M.-K. Circularly polarized light-sensitive, hot electron transistor with chiral plasmonic nanoparticles. *Nat. Commun.* **2022**, *13*, 5081.

(293) Chai, Y. Silicon photodiodes that multiply. *Nat. Electron.* **2022**, *5*, 483-484.

(294) Wan, T.; Shao, B.; Ma, S.; Zhou, Y.; Li, Q.; Chai, Y. In-Sensor Computing: Materials, Devices, and Integration Technologies. *Adv. Mater.* **2022**, 2203830.

(295) Jang, H.; Hinton, H.; Jung, W.-B.; Lee, M.-H.; Kim, C.; Park, M.; Lee, S.-K.; Park, S.; Ham, D. In-sensor optoelectronic computing using electrostatically doped silicon. *Nat. Electron.* **2022**, *5*, 519-525.

(296) Wan, T.; Ma, S.; Liao, F.; Fan, L.; Chai, Y. Neuromorphic sensory computing. *Sci. China Inf. Sci.* **2022**, *65*, 1-14.

(297) Liao, F.; Zhou, F.; Chai, Y. Neuromorphic vision sensors: Principle, progress and perspectives. *J. Semicond.* **2021**, *42*, 013105.

(298) Yoon, J. H.; Wang, Z.; Kim, K. M.; Wu, H.; Ravichandran, V.; Xia, Q.; Hwang, C. S.; Yang, J. J. An artificial nociceptor based on a diffusive memristor. *Nat. Commun.* **2018**, *9*, 417.

(299) Kumar, M.; Kim, H. S.; Kim, J. A highly transparent artificial photonic nociceptor. *Adv. Mater.* **2019**, *31*, 1900021.

(300) Zhang, W.; Gao, B.; Tang, J.; Yao, P.; Yu, S.; Chang, M.-F.; Yoo, H.-J.; Qian, H.; Wu, H. Neuro-inspired computing chips. *Nat. Electron.* **2020**, *3*, 371-382.

(301) Chen, W.-H.; Dou, C.; Li, K.-X.; Lin, W.-Y.; Li, P.-Y.; Huang, J.-H.; Wang, J.-H.; Wei, W.-C.; Xue, C.-X.; Chiu, Y.-C. CMOS-integrated memristive non-volatile computing-in-memory for AI edge processors. *Nat. Electron.* **2019**, *2*, 420-428.

(302) Hills, G.; Lau, C.; Wright, A.; Fuller, S.; Bishop, M. D.; Srimani, T.; Kanhaiya, P.; Ho, R.; Amer, A.; Stein, Y. Modern microprocessor built from complementary carbon nanotube transistors. *Nature* **2019**, *572*, 595-602.

(303) Bishop, M. D.; Hills, G.; Srimani, T.; Lau, C.; Murphy, D.; Fuller, S.; Humes, J.; Ratkovich, A.; Nelson, M.; Shulaker, M. M. Fabrication of carbon nanotube field-effect transistors in commercial silicon manufacturing facilities. *Nat. Electron.* **2020**, *3*, 492-501.

(304) Mukhopadhyay, S.; Long, Y.; Mudassar, B.; Nair, C.; DeProspo, B. H.; Torun, H. M.; Kathaperumal, M.; Smet, V.; Kim, D.; Yalamanchili, S. Heterogeneous integration for artificial intelligence: Challenges and opportunities. *IBM J. Res. Dev.* **2019**, *63*, 4: 1-4: 1.

(305) Kum, H. S.; Lee, H.; Kim, S.; Lindemann, S.; Kong, W.; Qiao, K.; Chen, P.; Irwin, J.; Lee, J. H.; Xie, S. Heterogeneous integration of single-crystalline complex-oxide membranes. *Nature* **2020**, *578*, 75-81.

(306) Ohara, Y.; Lee, K. W.; Kiyoyama, K.; Konno, S.; Sato, Y.; Watanabe, S.; Yabata, A.; Kobayashi, H.; Kamada, T.; Bea, J. Chip-based hetero-integration technology for high-performance 3D stacked image sensor. In *2012 2nd IEEE CPMT Symposium Japan*, 2012; IEEE: pp 1-4.

(307) Amir, M. F.; Ko, J. H.; Na, T.; Kim, D.; Mukhopadhyay, S. 3-D stacked image sensor with deep neural network computation. *IEEE Sens. J.* **2018**, *18*, 4187-4199.

(308) Aly, M. M. S.; Wu, T. F.; Bartolo, A.; Malviya, Y. H.; Hwang, W.; Hills, G.; Markov, I.; Wootters, M.; Shulaker, M. M.; Wong, H.-S. P. The N3XT approach to energy-efficient abundant-data computing. *Proc. IEEE* **2018**, *107*, 19-48.

(309) Bhansali, S.; Chapman, G. H.; Friedman, E. G.; Ismail, Y.; Mukund, P.; Tebbe, D.; Jain, V. K. 3D heterogeneous sensor system on a chip for defense and security applications. In *Unattended/Unmanned Ground, Ocean, and Air Sensor Technologies and Applications VI*, 2004; SPIE: Vol. 5417, pp 413-424.

(310) Ng, K. W.; Ko, W. S.; Tran, T.-T. D.; Chen, R.; Nazarenko, M. V.; Lu, F.; Dubrovskii, V. G.; Kamp, M.; Forchel, A.; Chang-Hasnain, C. J. Unconventional growth mechanism for monolithic integration of III–V on silicon. *ACS Nano* **2013**, *7*, 100-107.

(311) Koma, A. Van der Waals epitaxy for highly lattice-mismatched systems. *J. Cryst. Growth* **1999**, *201*, 236-241.

(312) Liau, Z.; Mull, D. Wafer fusion: A novel technique for optoelectronic device fabrication and monolithic integration. *Appl. Phys. Lett.* **1990**, *56*, 737-739.

(313) Benwadih, M.; Coppard, R.; Bonrad, K.; Klyszcz, A.; Vuillaume, D. High mobility flexible amorphous IGZO thin-film transistors with a low thermal budget ultra-violet pulsed light process. *ACS Appl. Mater. Interfaces* **2016**, *8*, 34513-34519.

(314) Vinet, M.; Batude, P.; Tabone, C.; Previtali, B.; LeRoyer, C.; Pouydebasque, A.; Clavelier, L.; Valentian, A.; Thomas, O.; Michaud, S. 3D monolithic integration: Technological challenges and electrical results. *Microelectron. Eng.* **2011**, *88*, 331-335.

(315) Bao, S.; Wang, Y.; Lina, K.; Zhang, L.; Wang, B.; Sasangka, W. A.; Lee, K. E. K.; Chua, S. J.; Michel, J.; Fitzgerald, E. A review of silicon-based wafer bonding processes, an approach to realize the monolithic integration of Si-CMOS and III–V-on-Si wafers. *J. Semicond.* **2021**, *42*, 023106.

(316) Lee, S.-M.; Kwong, A.; Jung, D.; Faucher, J.; Biswas, R.; Shen, L.; Kang, D.; Lee, M. L.; Yoon, J. High performance ultrathin GaAs solar cells enabled with heterogeneously integrated dielectric periodic nanostructures. *ACS Nano* **2015**, *9*, 10356-10365.

(317) Shulaker, M. M.; Hills, G.; Patil, N.; Wei, H.; Chen, H.-Y.; Wong, H.-S. P.; Mitra, S. Carbon nanotube computer. *Nature* **2013**, *501*, 526-530.

(318) Wang, S.; Wang, C.-Y.; Wang, P.; Wang, C.; Li, Z.-A.; Pan, C.; Dai, Y.; Gao, A.; Liu, C.; Liu, J. Networking retinomorphic sensor with memristive crossbar for brain-inspired visual perception. *Natl. Sci. Rev.* **2021**, *8*, nwaa172.

(319) Wang, C.; Liang, S.-J.; Wang, C.-Y.; Yang, Z.-Z.; Ge, Y.; Pan, C.; Shen, X.; Wei, W.; Zhao, Y.; Zhang, Z. Scalable massively parallel computing using continuous-time data representation in nanoscale crossbar array. *Nat. Nanotechnol.* **2021**, *16*, 1079-1085.

(320) Choi, C.; Kim, H.; Kang, J.-H.; Song, M.-K.; Yeon, H.; Chang, C. S.; Suh, J. M.; Shin, J.; Lu, K.; Park, B.-I. Reconfigurable heterogeneous integration using stackable chips with embedded artificial intelligence. *Nat. Electron.* **2022**, *5*, 386-393.

(321) Choi, M.; Koh, H.; Yoon, E.; Shin, K.; Song, K. Self-aligning silicon groove technology platform for the low cost optical module. In *1999 Proceedings. 49th Electronic Components and Technology Conference (Cat. No. 99CH36299)*, 1999; IEEE: pp 1140-1144.

(322) Barwicz, T.; Peng, B.; Leidy, R.; Janta-Polczynski, A.; Houghton, T.; Khater, M.; Kamlapurkar, S.; Engelmann, S.; Fortier, P.; Boyer, N. Integrated metamaterial interfaces for self-aligned fiber-to-chip coupling in volume manufacturing. *IEEE J. Sel. Top. Quantum Electron.* **2018**, *25*, 1-13.

(323) Sasaki, A.; Baba, T.; Iga, K. Put-in microconnectors for alignment-free coupling of optical fiber arrays. *IEEE Photon. Technol. Lett.* **1992**, *4*, 908-911.

(324) Barwicz, T.; Taira, Y.; Lichoulas, T. W.; Boyer, N.; Martin, Y.; Numata, H.; Nah, J.-W.; Takenobu, S.; Janta-Polczynski, A.; Kimbrell, E. L. A novel approach to photonic packaging leveraging existing high-throughput microelectronic facilities. *IEEE J. Sel. Top. Quantum Electron.* **2016**, *22*, 455-466.

(325) Wan, C.; Gonzalez, J. L.; Fan, T.; Adibi, A.; Gaylord, T. K.; Bakir, M. S. Fiber-interconnect silicon chiplet technology for self-aligned fiber-to-chip assembly. *IEEE Photon. Technol. Lett.* **2019**, *31*, 1311-1314.

(326) Zhang, Y.; Wang, Z.; Zhu, J.; Yang, Y.; Rao, M.; Song, W.; Zhuo, Y.; Zhang, X.; Cui, M.; Shen, L. Brain-inspired computing with memristors: Challenges in devices, circuits, and systems. *Appl. Phys. Rev.* **2020**, *7*, 011308.

(327) Li, C.; Graves, C. E.; Sheng, X.; Miller, D.; Foltin, M.; Pedretti, G.; Strachan, J. P. Analog content-addressable memories with memristors. *Nat. Commun.* **2020**, *11*, 1638.

(328) Li, Y.; Song, W.; Wang, Z.; Jiang, H.; Yan, P.; Lin, P.; Li, C.; Rao, M.; Barnell, M.; Wu, Q. Memristive Field-Programmable Analog Arrays for Analog Computing. *Adv. Mater.* **2022**, 2206648.

(329) Yang, J. J.; Miao, F.; Pickett, M. D.; Ohlberg, D. A.; Stewart, D. R.; Lau, C. N.; Williams, R. S. The mechanism of electroforming of metal oxide memristive switches. *Nanotechnology* **2009**, *20*, 215201.

(330) Yang, J. J.; Pickett, M. D.; Li, X.; Ohlberg, D. A.; Stewart, D. R.; Williams, R. S. Memristive switching mechanism for metal/oxide/metal nanodevices. *Nat. Nanotechnol.* **2008**, *3*, 429-433.

(331) Choi, B. J.; Torrezan, A. C.; Strachan, J. P.; Kotula, P.; Lohn, A.; Marinella, M. J.; Li, Z.; Williams, R. S.; Yang, J. J. High-speed and low-energy nitride memristors. *Adv. Funct. Mater.* **2016**, *26*, 5290-5296.

(332) Rao, M.; Tang, H.; Wu, J.; Song, W.; Zhang, M.; Yin, W.; Zhuo, Y.; Kiani, F.; Chen, B.; Jiang, X. Thousands of conductance levels in memristors integrated on CMOS. *Nature* **2023**, *615*, 823-829.

(333) Xue, W.; Gao, S.; Shang, J.; Yi, X.; Liu, G.; Li, R. W. Recent advances of quantum conductance in memristors. *Adv. Electron. Mater.* **2019**, *5*, 1800854.

(334) Li, Y.; Long, S.; Liu, Y.; Hu, C.; Teng, J.; Liu, Q.; Lv, H.; Suñé, J.; Liu, M. Conductance quantization in resistive random access memory. *Nanoscale Res. Lett.* **2015**, *10*, 1-30.

(335) Ando, T.; Arakawa, Y.; Furuya, K.; Komiyama, S.; Nakashima, H. *Mesoscopic physics and electronics*; 1st ed.; Springer Berlin, **1998**.

(336) Krans, J.; Van Ruitenbeek, J.; Fisun, V.; Yanson, I.; De Jongh, L. The signature of conductance quantization in metallic point contacts. *Nature* **1995**, *375*, 767-769.

(337) Ohnishi, H.; Kondo, Y.; Takayanagi, K. Quantized conductance through individual rows of suspended gold atoms. *Nature* **1998**, *395*, 780-783.

(338) Terabe, K.; Hasegawa, T.; Nakayama, T.; Aono, M. Quantum point contact switch realized by solid electrochemical reaction. *Riken Rev.* **2001**, 7-8.

(339) Sune, J.; Miranda, E.; Nafria, M.; Aymerich, X. Point contact conduction at the oxide breakdown of MOS devices. In *International Electron Devices Meeting 1998. Technical Digest (Cat. No. 98CH36217)*, 1998; IEEE: pp 191-194.

(340) Wagenaar, J. J.; Morales-Masis, M.; Van Ruitenbeek, J. M. Observing "quantized" conductance steps in silver sulfide: Two parallel resistive switching mechanisms. *J. Appl. Phys.* **2012**, *111*, 014302.

(341) Jiang, L.; Xu, L.; Chen, J.; Yan, P.; Xue, K.; Sun, H.; Miao, X. Conductance quantization in an AgInSbTe-based memristor at nanosecond scale. *Appl. Phys. Lett.* **2016**, *109*, 153506.

(342) Gao, S.; Zeng, F.; Chen, C.; Tang, G.; Lin, Y.; Zheng, Z.; Song, C.; Pan, F. Conductance quantization in a Ag filament-based polymer resistive memory. *Nanotechnology* **2013**, *24*, 335201.

(343) Krishnan, K.; Muruganathan, M.; Tsuruoka, T.; Mizuta, H.; Aono, M. Highly reproducible and regulated conductance quantization in a polymer-based atomic switch. *Adv. Funct. Mater.* **2017**, *27*, 1605104.

(344) Nandakumar, S.; Minvielle, M.; Nagar, S.; Dubourdieu, C.; Rajendran, B. A 250 mV $Cu/SiO_2/W$ memristor with half-integer quantum conductance states. *Nano Lett.* **2016**, *16*, 1602-1608.

(345) Gao, S.; Chen, C.; Zhai, Z.; Liu, H.; Lin, Y.; Li, S.; Lu, S.; Wang, G.; Song, C.; Zeng, F. Resistive switching and conductance quantization in Ag/SiO_2/indium tin oxide resistive memories. *Appl. Phys. Lett.* **2014**, *105*, 063504.

(346) Tsuruoka, T.; Hasegawa, T.; Terabe, K.; Aono, M. Conductance quantization and synaptic behavior in a Ta_2O_5-based atomic switch. *Nanotechnology* **2012**, *23*, 435705.

(347) Banerjee, W.; Hwang, H. Quantized conduction device with 6-bit storage based on electrically controllable break junctions. *Adv. Electron. Mater.* **2019**, *5*, 1900744.

(348) Zhu, X.; Su, W.; Liu, Y.; Hu, B.; Pan, L.; Lu, W.; Zhang, J.; Li, R. W. Observation of conductance quantization in oxide-based resistive switching memory. *Adv. Mater.* **2012**, *24*, 3941-3946.

(349) Miranda, E.; Kano, S.; Dou, C.; Kakushima, K.; Suñé, J.; Iwai, H. Nonlinear conductance quantization effects in CeO_x/SiO_2-based resistive switching devices. *Appl. Phys. Lett.* **2012**, *101*, 012910.

(350) Mehonic, A.; Vrajitoarea, A.; Cueff, S.; Hudziak, S.; Howe, H.; Labbe, C.; Rizk, R.; Pepper, M.; Kenyon, A. Quantum conductance in silicon oxide resistive memory devices. *Sci. Rep.* **2013**, *3*, 1-8.

(351) Tappertzhofen, S.; Valov, I.; Waser, R. Quantum conductance and switching kinetics of AgI-based microcrossbar cells. *Nanotechnology* **2012**, *23*, 145703.

(352) Zhao, J.; Zhou, Z.; Zhang, Y.; Wang, J.; Zhang, L.; Li, X.; Zhao, M.; Wang, H.; Pei, Y.; Zhao, Q. An electronic synapse memristor device with conductance linearity using quantized conduction for neuroinspired computing. *J. Mater. Chem. C* **2019**, *7*, 1298-1306.

(353) Younis, A.; Chu, D.; Li, S. Voltage sweep modulated conductance quantization in oxide nanocomposites. *J. Mater. Chem. C* **2014**, *2*, 10291-10297.

(354) Chen, C.; Gao, S.; Zeng, F.; Wang, G.; Li, S.; Song, C.; Pan, F. Conductance quantization in oxygen-anion-migration-based resistive switching memory devices. *Appl. Phys. Lett.* **2013**, *103*, 043510.

(355) Milano, G.; Luebben, M.; Ma, Z.; Dunin-Borkowski, R.; Boarino, L.; Pirri, C. F.; Waser, R.; Ricciardi, C.; Valov, I. Self-limited single nanowire systems combining all-in-one memristive and neuromorphic functionalities. *Nat. Commun.* **2018**, *9*, 5151.

(356) Kuncic, Z.; Nakayama, T. Neuromorphic nanowire networks: principles, progress and future prospects for neuro-inspired information processing. *Adv. Phys.: X* **2021**, *6*, 1894234.

(357) Milano, G.; Pedretti, G.; Fretto, M.; Boarino, L.; Benfenati, F.; Ielmini, D.; Valov, I.; Ricciardi, C. Brain-inspired structural plasticity through reweighting and rewiring in multi-terminal self-organizing memristive nanowire networks. *Adv. Intell. Syst.* **2020**, *2*, 2000096.

(358) Milano, G.; Pedretti, G.; Montano, K.; Ricci, S.; Hashemkhani, S.; Boarino, L.; Ielmini, D.; Ricciardi, C. In materia reservoir computing with a fully memristive architecture based on self-organizing nanowire networks. *Nat. Mater.* **2022**, *21*, 195-202.

(359) Milano, G.; Montano, K.; Ricciardi, C. In materia implementation strategies of physical reservoir computing with memristive nanonetworks. *J. Phys. D: Appl. Phys* **2023**, *56*, 084005.

(360) Milano, G.; Agliuzza, M.; de Leo, N.; Ricciardi, C. Speech recognition through physical reservoir computing with neuromorphic nanowire networks. In *2022 International Joint Conference on Neural Networks (IJCNN)*, 2022; IEEE: pp 1-6.

(361) Milano, G.; Miranda, E.; Ricciardi, C. Connectome of memristive nanowire networks through graph theory. *Neural Netw.* **2022**, *150*, 137-148.

(362) Montano, K.; Milano, G.; Ricciardi, C. Grid-graph modeling of emergent neuromorphic dynamics and heterosynaptic plasticity in memristive nanonetworks. *Neuromorphic Comput. Eng.* **2022**, *2*, 014007.

(363) Stieg, A. Z.; Avizienis, A. V.; Sillin, H. O.; Martin-Olmos, C.; Aono, M.; Gimzewski, J. K. Emergent criticality in complex turing B-type atomic switch networks. *Adv. Mater.* **2012**, *24*, 286-293.

(364) Dunham, C. S.; Lilak, S.; Hochstetter, J.; Loeffler, A.; Zhu, R.; Chase, C.; Stieg, A. Z.; Kuncic, Z.; Gimzewski, J. K. Nanoscale neuromorphic networks and criticality: a perspective. *J. Phys. Complexity* **2021**, *2*, 042001.

(365) Manning, H. G.; Niosi, F.; da Rocha, C. G.; Bellew, A. T.; O'Callaghan, C.; Biswas, S.; Flowers, P. F.; Wiley, B. J.; Holmes, J. D.; Ferreira, M. S. Emergence of winner-takes-all connectivity paths in random nanowire networks. *Nat. Commun.* **2018**, *9*, 3219.

(366) Yuan, F.; Zhang, Z.; Pan, L.; Xu, J. A combined modulation of set current with reset voltage to achieve 2-bit/cell performance for filament-based RRAM. *IEEE J. Electron Devices Soc.* **2014**, *2*, 154-157.

(367) Emboras, A.; Niegemann, J.; Ma, P.; Haffner, C.; Pedersen, A.; Luisier, M.; Hafner, C.; Schimmel, T.; Leuthold, J. Atomic scale plasmonic switch. *Nano Lett.* **2016**, *16*, 709-714.

(368) Milano, G.; Ferrarese Lupi, F.; Fretto, M.; Ricciardi, C.; De Leo, N.; Boarino, L. Memristive devices for quantum metrology. *Adv. Quantum Technol.* **2020**, *3*, 2000009.

(369) Rolls, E. T.; Deco, G. *The noisy brain: stochastic dynamics as a principle of brain function*; Oxford University Press: Oxford, UK, **2010,** 34.

(370) Joksas, D.; Freitas, P.; Chai, Z.; Ng, W. H.; Buckwell, M.; Li, C.; Zhang, W.; Xia, Q.; Kenyon, A.; Mehonic, A. Committee machines—a universal method to deal with non-idealities in memristor-based neural networks. *Nat. Commun.* **2020**, *11*, 4273.

(371) Yi, S.-i.; Kendall, J. D.; Williams, R. S.; Kumar, S. Activity-difference training of deep neural networks using memristor crossbars. *Nat. Electron.* **2023**, *6*, 45-51.

(372) Misra, S.; Bland, L. C.; Cardwell, S. G.; Incorvia, J. A. C.; James, C. D.; Kent, A. D.; Schuman, C. D.; Smith, J. D.; Aimone, J. B. Probabilistic Neural Computing with Stochastic Devices. *Adv. Mater.* **2022**, 2204569.

(373) Cardwell, S. G.; Schuman, C. D.; Smith, J. D.; Patel, K.; Kwon, J.; Liu, S.; Allemang, C.; Misra, S.; Incorvia, J. A.; Aimone, J. B. Probabilistic Neural Circuits leveraging AI-Enhanced Codesign for Random Number Generation. **2022**, 2212.00625, ver.1, *arXiv (Emerging Technologies)*. DOI: 10.48550/arXiv.2212.00625 (accessed 05-27-2023).

(374) Kumar, S.; Wang, X.; Strachan, J. P.; Yang, Y.; Lu, W. D. Dynamical memristors for higher-complexity neuromorphic computing. *Nat. Rev. Mater.* **2022**, *7*, 575-591.

(375) Jo, S. H.; Chang, T.; Ebong, I.; Bhadviya, B. B.; Mazumder, P.; Lu, W. Nanoscale memristor device as synapse in neuromorphic systems. *Nano Lett.* **2010**, *10*, 1297-1301.

(376) Zidan, M. A.; Strachan, J. P.; Lu, W. D. The future of electronics based on memristive systems. *Nat. Electron.* **2018**, *1*, 22-29.

(377) Najem, J. S.; Taylor, G. J.; Weiss, R. J.; Hasan, M. S.; Rose, G.; Schuman, C. D.; Belianinov, A.; Collier, C. P.; Sarles, S. A. Memristive ion channel-doped biomembranes as synaptic mimics. *ACS Nano* **2018**, *12*, 4702-4711.

(378) Jo, S. H.; Lu, W. CMOS compatible nanoscale nonvolatile resistance switching memory. *Nano Lett.* **2008**, *8*, 392-397.

(379) Kim, S.; Choi, S.; Lu, W. Comprehensive physical model of dynamic resistive switching in an oxide memristor. *ACS Nano* **2014**, *8*, 2369-2376.

(380) Kim, S.; Choi, S.; Lee, J.; Lu, W. D. Tuning resistive switching characteristics of tantalum oxide memristors through Si doping. *ACS Nano* **2014**, *8*, 10262-10269.

(381) Zhu, X.; Li, D.; Liang, X.; Lu, W. D. Ionic modulation and ionic coupling effects in MoS_2 devices for neuromorphic computing. *Nat. Mater.* **2019**, *18*, 141-148.

(382) Li, H.; Wu, T. F.; Rahimi, A.; Li, K.-S.; Rusch, M.; Lin, C.-H.; Hsu, J.-L.; Sabry, M. M.; Eryilmaz, S. B.; Sohn, J. Hyperdimensional computing with 3D VRRAM in-memory kernels: Device-architecture co-design for energy-efficient, error-resilient language recognition. In *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016; IEEE: pp 16.11. 11-16.11. 14.

(383) Li, H.; Chen, W.-C.; Levy, A.; Wang, C.-H.; Wang, H.; Chen, P.-H.; Wan, W.; Wong, H.-S. P.; Raina, P. One-shot learning with memory-augmented neural networks using a 64-kbit, 118 GOPS/W RRAM-based non-volatile associative memory. In *2021 Symposium on VLSI Technology*, 2021; IEEE: pp 1-2.

(384) Du, C.; Ma, W.; Chang, T.; Sheridan, P.; Lu, W. D. Biorealistic implementation of synaptic functions with oxide memristors through internal ionic dynamics. *Adv. Funct. Mater.* **2015**, *25*, 4290-4299.

(385) Kim, S.; Du, C.; Sheridan, P.; Ma, W.; Choi, S.; Lu, W. D. Experimental demonstration of a second-order memristor and its ability to biorealistically implement synaptic plasticity. *Nano Lett.* **2015**, *15*, 2203-2211.

(386) Yoo, S.; Wu, Y.; Park, Y.; Lu, W. D. Tuning Resistive Switching Behavior by Controlling Internal Ionic Dynamics for Biorealistic Implementation of Synaptic Plasticity. *Adv. Electron. Mater.* **2022**, *8*, 2101025.

(387) Lee, S. H.; Moon, J.; Jeong, Y.; Lee, J.; Li, X.; Wu, H.; Lu, W. D. Quantitative, dynamic TaO_x memristor/resistive random access memory model. *ACS Appl. Electron. Mater.* **2020**, *2*, 701-709.

(388) Lee, J.; Lu, W. D. On-demand reconfiguration of nanomaterials: when electronics meets ionics. *Adv. Mater.* **2018**, *30*, 1702770.

(389) Wu, Y.; Moon, J.; Zhu, X.; Lu, W. D. Neural functional connectivity reconstruction with second-order memristor network. *Adv. Intell. Syst.* **2021**, *3*, 2000276.

(390) Du, C.; Cai, F.; Zidan, M. A.; Ma, W.; Lee, S. H.; Lu, W. D. Reservoir computing using dynamic memristors for temporal information processing. *Nat. Commun.* **2017**, *8*, 2204.

(391) Moon, J.; Ma, W.; Shin, J. H.; Cai, F.; Du, C.; Lee, S. H.; Lu, W. D. Temporal data classification and forecasting using a memristor-based reservoir computing system. *Nat. Electron.* **2019**, *2*, 480-487.

(392) Mountcastle, V. B. The columnar organization of the neocortex. *Brain* **1997**, *120*, 701-722.

(393) Spruston, N. Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.* **2008**, *9*, 206-221.

(394) Spruston, N.; Jaffe, D. B.; Johnston, D. Dendritic attenuation of synaptic potentials and currents: the role of passive membrane properties. *Trends Neurosci.* **1994**, *17*, 161-166.

(395) Kim, J.; Kim, S.-P.; Kim, J.; Hwang, H.; Kim, J.; Park, D.; Jeong, U. Object shape recognition using tactile sensor arrays by a spiking neural network with unsupervised learning. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020; IEEE: pp 178-183.

(396) Kang, P.; Banerjee, S.; Chopp, H.; Katsaggelos, A.; Cossairt, O. Event-Driven Tactile Learning with Location Spiking Neurons. In *2022 International Joint Conference on Neural Networks (IJCNN)*, 2022; IEEE: pp 1-9.

(397) Li, H.; Liu, H.; Ji, X.; Li, G.; Shi, L. CIFAR10-DVS: an event-stream dataset for object classification. *Front. Neurosci.* **2017**, *11*, 309.

(398) Amir, A.; Taba, B.; Berg, D.; Melano, T.; McKinstry, J.; Di Nolfo, C.; Nayak, T.; Andreopoulos, A.; Garreau, G.; Mendoza, M. A low power, fully event-based gesture recognition system. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017; pp 7243-7252.

(399) Shadmehr, R.; Smith, M. A.; Krakauer, J. W. Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* **2010**, *33*, 89-108.

(400) Yoo, S.; Park, Y.; Wang, Z.; Wu, Y.; Medepalli, S.; Thio, W.; Lu, W. D. Columnar Learning Networks for Multisensory Spatiotemporal Learning. *Adv. Intell. Syst.* **2022**, *4*, 2200179.

(401) Hawkins, J.; Ahmad, S.; Cui, Y. A theory of how columns in the neocortex enable learning the structure of the world. *Front. Neural Circuits* **2017**, 81.

(402) Golding, N. L.; Mickus, T. J.; Katz, Y.; Kath, W. L.; Spruston, N. Factors mediating powerful voltage attenuation along CA1 pyramidal neuron dendrites. *J. Physiol.* **2005**, *568*, 69-82.

(403) Rall, W. Theoretical significance of dendritic trees for neuronal input-output relations. In *Neural theory and modeling*, 1964; pp 73-97.

(404) Rall, W.; Burke, R.; Smith, T.; Nelson, P. G.; Frank, K. Dendritic location of synapses and possible mechanisms for the monosynaptic EPSP in motoneurons. *J. Neurophysiol.* **1967**, *30*, 1169-1193.

(405) Agmon-Snir, H.; Carr, C. E.; Rinzel, J. The role of dendrites in auditory coincidence detection. *Nature* **1998**, *393*, 268-272.

(406) Single, S.; Borst, A. Dendritic integration and its role in computing image velocity. *Science* **1998**, *281*, 1848-1850.

(407) Wang, W.; Gao, S.; Wang, Y.; Li, Y.; Yue, W.; Niu, H.; Yin, F.; Guo, Y.; Shen, G. Advances in Emerging Photonic Memristive and Memristive-Like Devices. *Adv. Sci.* **2022**, *9*, 2105577.

(408) Chua, L. If it's pinched it's a memristor. *Semicond. Sci. Technol.* **2014**, *29*, 104001.

(409) Biolek, D.; Biolek, Z.; Biolkova, V. Pinched hysteretic loops of ideal memristors, memcapacitors and meminductors must be 'self-crossing'. *Electron. Lett.* **2011**, *47*, 1385-1387.

(410) Hamdioui, S.; Xie, L.; Du Nguyen, H. A.; Taouil, M.; Bertels, K.; Corporaal, H.; Jiao, H.; Catthoor, F.; Wouters, D.; Eike, L. Memristor based computation-in-memory architecture for data-intensive applications. In *2015 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, Grenoble, France, 2015; IEEE: pp 1718-1725.

(411) Mead, C. Neuromorphic electronic systems. *Proc. IEEE* **1990**, *78*, 1629-1636.

(412) Mead, C. How we created neuromorphic engineering. *Nat. Electron.* **2020**, *3*, 434-435.

(413) Ielmini, D. Brain-inspired computing with resistive switching memory (RRAM): Devices, synapses and neural networks. *Microelectron. Eng.* **2018**, *190*, 44-53.

Table of Contents