Resistive switching materials for information processing

*Zhongrui Wang[1], Huaqiang Wu ⓘ[2], Geoffrey W. Burr ⓘ[3], Cheol Seong Hwang ⓘ[4], Kang L. Wang[5], Qiangfei Xia ⓘ[1] * and J. Joshua Yang ⓘ[1] *

Abstract | The rapid increase in information in the big-data era calls for changes to information-processing paradigms, which, in turn, demand new circuit-building blocks to overcome the decreasing cost-effectiveness of transistor scaling and the intrinsic inefficiency of using transistors in non-von Neumann computing architectures. Accordingly, resistive switching materials (RSMs) based on different physical principles have emerged for memories that could enable energy-efficient and area-efficient in-memory computing. In this Review, we survey the four physical mechanisms that lead to such resistive switching: redox reactions, phase transitions, spin-polarized tunnelling and ferroelectric polarization. We discuss how these mechanisms equip RSMs with desirable properties for representation capability, switching speed and energy, reliability and device density. These properties are the key enablers of processing-in-memory platforms, with applications ranging from neuromorphic computing and general-purpose memcomputing to cybersecurity. Finally, we examine the device requirements for such systems based on RSMs and provide suggestions to address challenges in materials engineering, device optimization, system integration and algorithm design.

[1]Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA, USA.

[2]Institute of Microelectronics, Tsinghua University, Beijing, China.

[3]IBM Research–Almaden, San Jose, CA, USA.

[4]Department of Materials Science and Engineering, and Inter-University Semiconductor Research Center, Seoul National University, Seoul, Republic of Korea.

[5]Department of Electrical and Computer Engineering, University of California, Los Angeles, CA, USA.

*e-mail: qxia@umass.edu; jjyang@umass.edu

The rapid development of information processing has relied on advances in the underlying electronic devices, particularly metal–oxide–semiconductor field-effect transistors. However, the main driving forces — Moore's law[1] and Dennard scaling[2] — have become difficult to sustain at their past pace, as shrinking technology nodes encounter physical limitations that make further scaling less cost-effective than before.

Meanwhile, the increase in big data, in part due to the Internet of Things, and big-data workloads, such as machine learning, are driving the development of increasingly data-centric computing, which poses a challenge to traditional digital computers based on the von Neumann architecture. The frequent data shuttling between the physically separated processing and memory units in these computers incurs considerable penalties on the energy efficiency and data bandwidth, termed the von Neumann bottleneck. This bottleneck is further intensified by the increasing disparity between the speed of the memory unit and the processor, termed the memory wall. Therefore, radically different computing paradigms are needed to boost the area and energy efficiencies to meet the ever-increasing demands of computing capacity in the big-data era.

Processing-in-memory (or memcomputing) is one of the most promising solutions to address these challenges, owing to the collocation of information storage and processing[3]. Recent efforts in stacking logic and memory dies in a single package have already shown remarkable advantages in bandwidth and energy[4]. Nevertheless, transistors were created and optimized for a different purpose, namely digital logic, and are thus inefficient to unleash the full potential of processing-in-memory. Therefore, new computing approaches require new memory devices that are fundamentally different in terms of device physics and materials to enable high-throughput, energy-efficient and area-efficient information processing.

Resistive switching materials (RSMs) exhibit tunable resistance states, induced by an external electrical bias. The corresponding resistive switching devices are normally in the form of a metal–dielectric–metal two-terminal junction, in which the resistive switching is typically driven by one of four mechanisms: electrochemical reactions[5,6] (namely, redox and ion migration), phase changes[7–9] (such as thermally activated amorphous–crystalline transitions), tunnel magneto-resistance[10,11] (such as spin-dependent tunnel resistance) or ferroelectricity[12,13] (for example, lattice-polarization-dependent resistance due to tunnelling or domain-wall transport). Owing to the dependence of their resistance states on the history of the electrical bias, RSMs can store information in the form of electrical resistance, in devices that can sometimes be termed memristors[14–16].

This ability, in addition to their high information-representation capability, reliability, low operating energy, high operating speed and small footprint, renders RSMs competitive with conventional charge-based memories. In addition, such resistance-based memory could process information intrinsically through 'compute-in-physics' (that is, the use of physical phenomena to implement complicated signal transformations), benefiting the area and energy efficiencies of various applications, including neuromorphic computing[17–23], general-purpose computing[24–28] and hardware security[29,30].

In this Review, we first discuss the switching mechanisms of different types of RSMs and highlight the microscopic origins. We then examine how these mechanisms lead to the unique electrical properties that are the basis of the recent processing-in-memory implementation of neuromorphic computing, general-purpose memcomputing and cybersecurity. Finally, we discuss the major challenges in achieving high-performance information processing and provide our perspective on the future directions for materials engineering, device optimization, system integration and algorithm design.

RSMs and mechanisms

Electrical-bias-induced resistance changes in metal–dielectric–metal junctions can be driven by one of four chemical and physical effects, as discussed in the following sections.

Redox RSMs

Electrochemical redox reactions change the chemical oxidation states of atoms, resulting in changes in electronic structure and charge transport. Since the 1960s, electrically induced redox switching has been observed in various non-crystalline dielectrics sandwiched between electrodes[31].

Switching mechanism. Within the general theory of resistive switching, the switching is attributed to redox reactions and ion migration, driven by electric potential, chemical potential and temperature gradients over the reaction coordinates[5,6,32–37]. In many dielectrics, especially transition-metal oxides and perovskites, anions such as oxygen ions (or, equivalently, oxygen vacancies) are more mobile than the cations when they are thermally and/or electrically activated. An applied electric field can then drive oxygen anions towards the anode, which modifies the electrical conductivity by inducing changes in the valence of the metal cations in the dielectric[38] (FIG. 1a). Electrochemical or electrothermal dissolution of the conduction channel occurs upon application of a reverse electric field or sufficient temperature, resetting the device back to a low-conductance (OFF) state.

The structure of a localized conduction channel of an anion-type device was first indicated using conductive-tip atomic force microscopy, which revealed the conductivity enhancement of a nanospot upon application of an electric field in an undoped $SrTiO_3$ single crystal[39]. A conical-pillar-shaped nanocrystalline filament of Ti_4O_7 (Magnéli phase) was visualized using transmission electron microscopy (TEM) in a unipolar switching $Pt/TiO_2/Pt$ device[40] (FIG. 1b). The evolution of O_2 gas at the anode, leaving oxygen vacancies in the dielectric, can result in deformation of the junction; for example, discontinuous and even delaminated top electrodes with dielectric voids have been observed in a $Pt/TiO_2/Pt$ device[41]. The composition of the nanofilaments has also been investigated by X-ray absorption spectroscopy in TiO_x (REF.[42]), $SrTiO_x$ (REF.[43]), $SrFeO_x$ (REF.[44]), HfO_x (REF.[45]) and TaO_x (REF.[46]) and also by electron-energy-loss spectroscopy in HfO_x (REFS[47,48]), $SrTiO_3$ (REFS[49,50]) and TaO_x (REFS[51,52]).

The conduction channels of redox RSMs can also be created by the redox and migration of cations, which involves the oxidation of an electrochemically active metal, such as Ag or Cu, followed by the drift of mobile cations in the solid electrolyte and the nucleation of cations to form a conducting filament (or filaments) upon reduction. Differences in ion mobilities and redox rates lead to different filament-growth modes. For instance, a low redox rate and ion mobility (as observed in poor ion-conductor RSMs, such as oxides and nitrides) typically result in filament growth from the active metal electrode (FIG. 1c), whereas a high redox rate and ion mobility (as found in good ion-conductor RSMs, such as sulfides, iodides, selenides, tellurides and ternary chalcogenides) could result in filament growth in the opposite direction[53].

Cation-based filaments were first visualized in a planar Ag/As_2S_3:Ag/Ag device[54], and similar Ag dendrites have been observed in aqueous electrolytes[55]. The dynamic switching process of a planar Au/SiO_x:Ag/Au diffusive RSM device (which differs from typical cation-type redox-RSM devices owing to spontaneous filament rupture upon the removal of the electrical bias) was captured by in situ TEM, revealing that a Ag bridge forms between the two electrodes (FIG. 1d), owing to the bipolar-electrode effect, whereby both oxidation and reduction occur within the same Ag nanoparticle[56]. This effect has been observed not only with Ag or Cu cation-type devices[57–62] but also in devices with more inert metals[53]. Following removal of the bias, the elongated cluster of nanoparticles quickly retracts through diffusion by minimization of the interfacial energy between the metal and the dielectric[56,63,64] (FIG. 1d). The dynamic evolution of a Cu filament in SiO_2, a challenging process to investigate computationally, has been modelled using reactive molecular dynamics[65]. The composition of Ag or Cu filaments in various solid electrolytes has been verified by energy-dispersive X-ray spectroscopy[57–59,62,66,67]. Moreover, it has been revealed that cation transport is possible in a typical valence-change redox device (such as TaO_x) under certain conditions[68].

In addition to filamentary switching, homogenous interfacial switching, in which the conductance scales with the junction area of the device, is also observed in some perovskites (for example, $Pr_{0.7}Ca_{0.3}MnO_3$). The mechanism of such interfacial resistive switching could be either charge trapping[69,70] or an electrochemical reaction[71–74], subject to different pairs of electrodes and RSMs.

Transport mechanism. Redox RSMs are often non-crystalline insulators in their OFF states with localized electronic states bounded by the mobility edges and, thus, favour thermally facilitated transport mechanisms, such as electron hopping[75] (FIG. 1e).
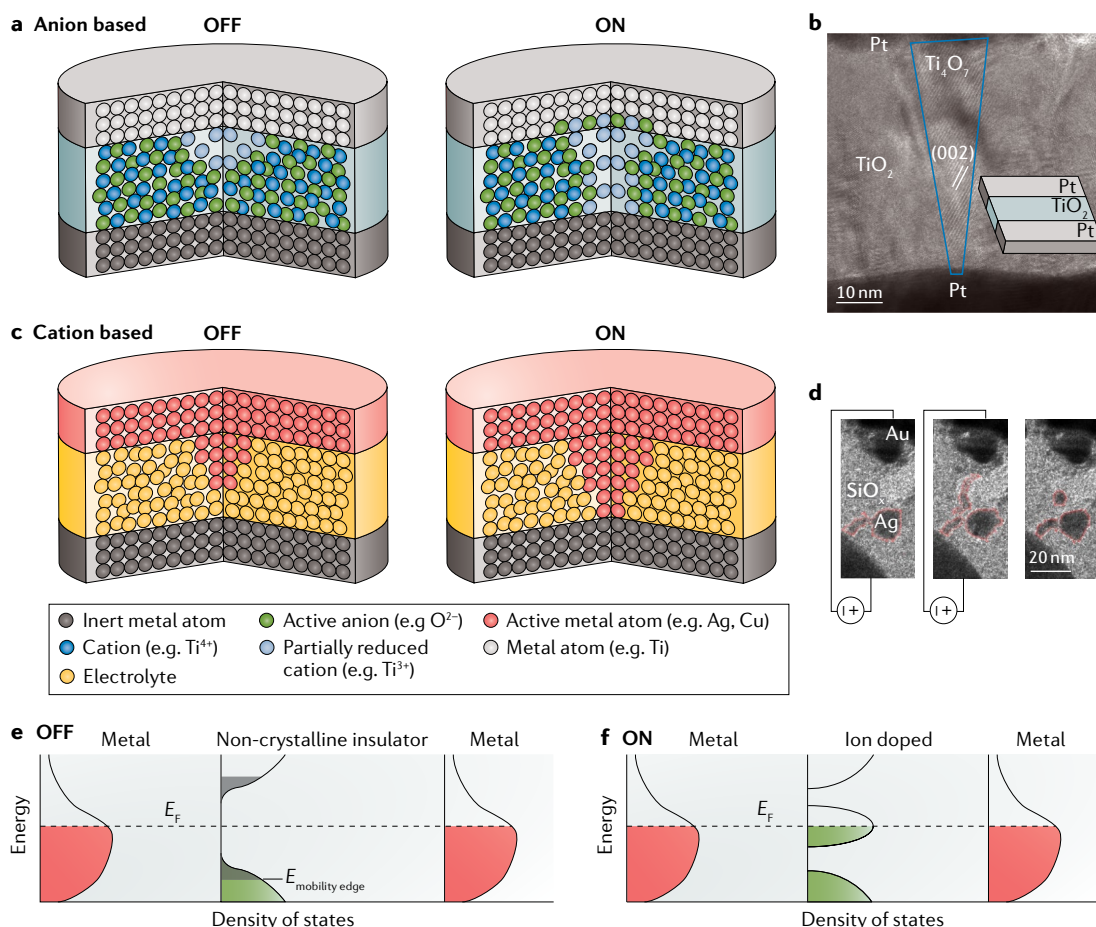
Fig. 1 | **Redox RSMs. a** | Illustration of an oxygen-vacancy filament in the OFF state (left) and ON state (right) in an anion-type resistive switching material (RSM) device. **b** | High-resolution transmission electron microscopy image of a nanocrystalline Ti_4O_7 (Magnéli phase) filament with a conical shape (outlined in blue) in a $Pt/TiO_2/Pt$ anion-type RSM device. **c** | Illustration of a metal-atom filament in the OFF state (left) and ON state (right) in a cation-type RSM device. **d** | Observation of a Ag filament in a planar $Au/SiO_x:Ag/Au$ cation-type RSM device with in situ high-resolution transmission electron microscopy. The images show the formation of the conducting Ag bridge upon application of an external bias (left and centre), followed by spontaneous retraction upon removal of the bias (right). **e** | Density of states of the metal electrodes and redox RSM in the OFF state. Localized states (grey area) within the mobility edges ($E_{mobility\ edge}$) lead to various thermally facilitated transport mechanisms, such as electron hopping. **f** | Density of states of the metal electrodes and the redox RSM in the ON state. The excessive doping of cations or anions leads to the overlap of orbitals of the localized impurity states and the formation of an impurity band. E_F, Fermi level. Panel **b** is adapted from REF.[40], Springer Nature Limited. Panel **d** is adapted from REF.[56], Springer Nature Limited.

The redox reactions change the stoichiometry of the material[38], resulting in doping with electronic impurities. The orbitals of the localized impurity states could overlap significantly at a large non-stoichiometry, leading to the formation of a delocalized impurity band(s) in the ON state, similar to that in a degenerately doped amorphous semiconductor[75] (FIG. 1f). Furthermore, a large non-stoichiometry is usually accompanied by structural transformations such as defect ordering, assisted by Joule heating, which further affects the electronic structure[40,76].

Phase-change RSMs

Phase-change RSMs are based on reversible phase transitions in a family of chalcogenide glass materials that frequently, but not always, involve Ge, Te and/or Sb (REFS[77–79]), as illustrated by the Ge–Sb–Te ternary phase diagram (FIG. 2a). In their metastable amorphous phase, which exhibits long-range disorder (but usually with some short-range order), the electrical resistivity of these materials is typically high. In the crystalline phase, which exhibits long-range order, the electrical resistivity is much lower[7–9]. The amorphous and crystalline phases also exhibit notable optical contrast, which led to the invention of optical rewritable disks in the late 1980s[7]. One question that continues to fascinate material scientists is why the macroscopic properties of the phases can change so dramatically[9], given the relatively subtle changes in atomic positions.

Switching mechanism. Unipolar electrical switching of phase-change RSMs is due to temperature changes, driven by Joule heating and passive thermal dissipation (FIG. 2b). Switching the crystalline phase (FIG. 2c) into the amorphous phase (FIG. 2d) involves heating
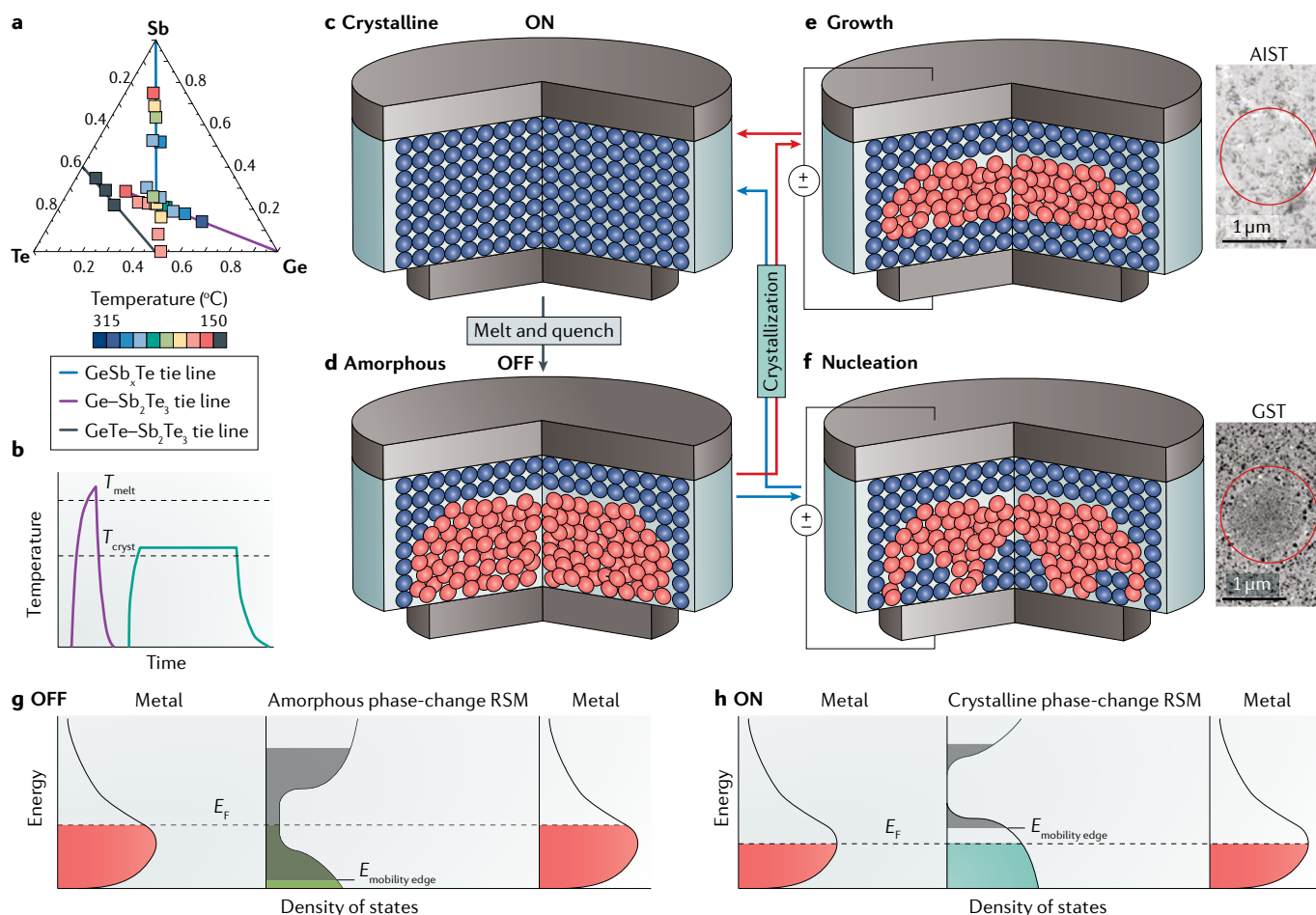
Fig. 2 | **Phase-change RSMs. a** | Ge–Sb–Te ternary phase diagram showing the composition dependence of the crystallization temperatures of phase-change resistive switching materials (RSMs). **b** | In phase-change RSMs, reversible phase transitions between the crystalline and amorphous phases are induced by temperature changes (where T_{melt} is the melting point and T_{cryst} is the crystallization temperature). **c,d** | Illustration of the crystalline (panel **c**) and amorphous (panel **d**) states of a phase-change RSM. Switching the crystalline phase into the amorphous phase involves heating above T_{melt} to melt and then rapidly quench a volume of the phase-change RSM (purple line in panel **b**). Switching the amorphous phase into a crystalline phase involves heating above T_{cryst} for sufficient time to crystallize the phase-change RSM (green line in panel **b**). **e,f** | Illustrations of the growth-dominated (panel **e**) and nucleation-dominated (panel **f**) crystallization processes, together with transmission electron microscopy images of exemplar materials, namely crystalline $Ag_4In_3Sb_{67}Te_{26}$ (AIST) and polycrystalline $Ge_2Sb_2Te_5$ (GST), respectively. **g** | Density of states of the metal electrodes and the phase-change RSM in the OFF state. The pinned Fermi level (E_F) and Anderson localization lead to thermally facilitated charge transport. **h** | Density of states of the metal electrodes and the phase-change RSM in the ON state. The increase in long-range order raises the mobility edge ($E_{mobility\,edge}$), whereas self-doping lowers E_F. Panel **a** is adapted with permission from REF.[299], IEEE. Panel **b** is adapted with permission from REF.[79], IEEE. The right parts of panels **e** and **f** are adapted from REF.[83], CC-BY-4.0 (https://creativecommons.org/licenses/by/4.0/).

the phase-change RSM above the melting point (typically, $T_{melt} \approx 600\,°C$), followed by a rapid quench of the undercooled liquid. Usually, the electrical heating power peaks in the melting process, whereas passive cooling is sufficient for amorphization on the sub-nanosecond timescale, owing to the nanosize of the melted region and the presence of high-thermal-conductivity wiring[80]. In addition, the amorphous state, effectively an undercooled liquid, continues to relax over time, long after the quench to low temperatures is completed[81]. This phenomenon, known as resistance drift, leads to a slow but steady increase in resistance over time. Resistance drift can be relatively substantial in amorphous phase-change RSMs[79] but is smaller in crystalline phase-change RSMs.

Switching into the crystalline phase involves holding the temperature of the amorphous phase-change RSM above the crystallization temperature (typically, $T_{cryst} \approx 400\,°C$) for long enough for crystallization to occur[82,83]. As this process usually takes 50–1,000 ns, it dominates the operational speed of phase-change RSMs and has inevitably led to searches for faster-crystallizing materials. Depending on the thermodynamic factors (for example, free energy as a function of both the specific nuclei radius and interfacial area) during crystallization, phase-change RSMs are categorized as either growth-dominated or nucleation-dominated materials[9,83]. Growth-dominated phase-change RSMs recrystallize at the interface between the crystalline and amorphous

matrices, which usually yields a single crystalline phase. An example of a growth-dominated phase-change RSM is Ag-doped and In-doped SbTe ($Ag_4In_3Sb_{67}Te_{26}$, AIST) (FIG. 2e). However, most phase-change RSMs are nucleation dominated, featuring stochastic formation of critical nuclei. Note that, in nucleation-dominated phase-change RSMs such as $Ge_2Sb_2Te_5$, both growth and nucleation are fast. Therefore, the formation of nuclei is accompanied by rapid grain growth, leading to a polycrystalline phase of different grain orientations[83] (FIG. 2f). The speed of crystal growth varies by more than ten orders of magnitude, primarily owing to temperature[82]. This tunability of the crystallization speed enables analogue programming of the conductance between the two extremes, with applications in both storage-class memory[84,85] and in-memory computing[23,27,86–88].

Charge transport. The OFF state of phase-change RSMs is typically amorphous with substantial structural randomness, which yields many localized electronic states with wavefunctions that decay exponentially. These localized states, bounded by the mobility edges, pin the Fermi level to the bandgap[89] (FIG. 2g). Such Anderson-type localization of the electronic states facilitates thermally activated transport[90,91], which is further complicated by the strong electron–lattice interactions and the negative effective correlation energy[92].

In the crystalline ON state of phase-change RSMs, which has been referred to as a metavalent solid or incipient metal[93], annealing leads to increased long-range order with fewer localized states and raises the mobility edge[89]. Self-doping, such as that due to excessive Ge and Sb vacancies in rock-salt $GeTe–Sb_2Te_3$ (REF.[91]), lowers the Fermi level to the valence band, which increases the concentration of delocalized carriers[89] and gives rise to a larger conductance (FIG. 2h).

Magnetic-tunnelling RSMs

Magnets have been used for memory since the 1930s and have been the backbone of storage devices since the development of hard disk drives in the 1950s[94]. More recently, there has been a major effort to integrate magnetoresistive-RSM-based memory, such as magnetic tunnelling junctions (MTJs)[95,96], with complementary metal–oxide–semiconductor (CMOS) logic chips[94]. A MTJ consists of two ferromagnetic layers that sandwich a thin (1–2-nm) insulator as a tunnelling barrier (FIG. 3). The reference, or fixed, magnetic layer has a pinned magnetization orientation, whereas the other magnetic layer, the magnetic free layer, can be switched between two opposite orientations.

Switching mechanism. Over the past 30 years, considerable efforts have been made to optimize MTJs for CMOS integration. There have been three key milestones. The first milestone was the evolution from giant magnetoresistance spin valves comprising Cu between ferromagnetic layers[97,98] to MTJs using an Al_2O_3 tunnelling barrier[95]. The second breakthrough was the improvement in the tunnelling magnetoresistance ratio by changing the tunnel barrier from amorphous Al_2O_3 to crystalline MgO (REFS[99,100]). Last, but not least, the

scalability has been improved by changing the magnetic anisotropy of the ferromagnetic layer from in-plane to out-of-plane[101]. The switching of the ferromagnetic layer has also gone through several phases of development, moving from the magnetic-field switching generated by a running current to using quantum phenomena, such as spin-transfer torque (STT), voltage-controlled magnetic anisotropy (VCMA) and spin–orbit torque (SOT).

STT RSMs use the STT effect[102,103] to switch the magnetic state of the free layer. When electrons first flow through the magnetically fixed layer, the majority of them become spin polarized by the magnetization orientation of the fixed layer and, subsequently, transfer spin angular momentum to the free magnetic layer, thereby switching the magnetization orientation of the free layer to that of the fixed layer (resulting in a parallel configuration) (FIG. 3a). If the current direction is reversed under the parallel configuration, the minority electrons with spin polarization opposite to that of the free layer are reflected by the fixed layer back to the free layer, thereby switching the free layer to be antiparallel to the fixed layer[10]. STT RSMs, as the most technically mature magnetoresistive RSM, are currently being developed by several major foundries for memory applications[104–106].

VCMA RSMs use the capacitive nature of the MgO barrier for switching. When a voltage is applied across a VCMA MTJ, charge accumulation or depletion at the magnetic free layer/MgO interface modifies the electronic occupation of atomic orbitals[107], causing a change in the total magnetic anisotropy of the free layer (FIG. 3b). When sufficient voltage is applied to the MTJ, the free layer will precess around the in-plane direction, as described by the Landau–Lifshitz–Gilbert equation[11]. A voltage pulse with a duration half that of the precession period can, therefore, switch the magnetization direction of the free layer. A thicker MgO barrier with a resistance–area product at least one order of magnitude larger than that in STT MTJs is typically used[108,109] (on the order of hundreds of ohm μm^2) to decrease the current and suppress the STT effect. To enhance the VCMA effect, MgO seed and Cr buffer layers have been optimized for epitaxial Fe/MgO/Fe growth[110]. In addition, Mg insertion at a CoFeB/MgO interface increased the response of the electronic orbitals to the applied voltage[111]. To make voltage-controlled MTJs compatible with the thermal budget of the industrial CMOS back-end-of-line process, thermal-barrier materials, such as Mo and W, are being explored to prevent seed or capping layers from diffusing into the CoFeB/MgO layers[112,113].

SOT RSMs are based on the surface spin accumulation from materials with high spin–orbit coupling (SOC)[114], which efficiently manipulates the magnetization adjacent to the SOC materials[115,116]. Unlike STT magnetoresistive random-access memory (MRAM), the writing current goes through a separate high-SOC metal layer that is underneath the MTJ stack. As the current does not pass through the MTJ stack, the MTJ write endurance improves and there is a wider range of material choices that could be used to increase the writing-energy efficiency (FIG. 3c). The main high-SOC materials under study today are bulk heavy metals and Rashba interfaces[115,116].

Transport mechanism. The conductivity of MTJs is based on the tunnelling magnetoresistance effect (FIG. 3d,e). Specifically, the spin-dependent electrons tunnelling across a thin insulating layer between two ferromagnetic layers yield different tunnelling magnetoresistance, because the tunnelling probability is approximately proportional to the product of the density of states of both ferromagnetic layers[117,118]. When the magnetization orientations of the ferromagnetic layers are antiparallel (parallel), the device is in its OFF state (ON state) and displays a low (high) conductance because the majority and minority electrons tunnel
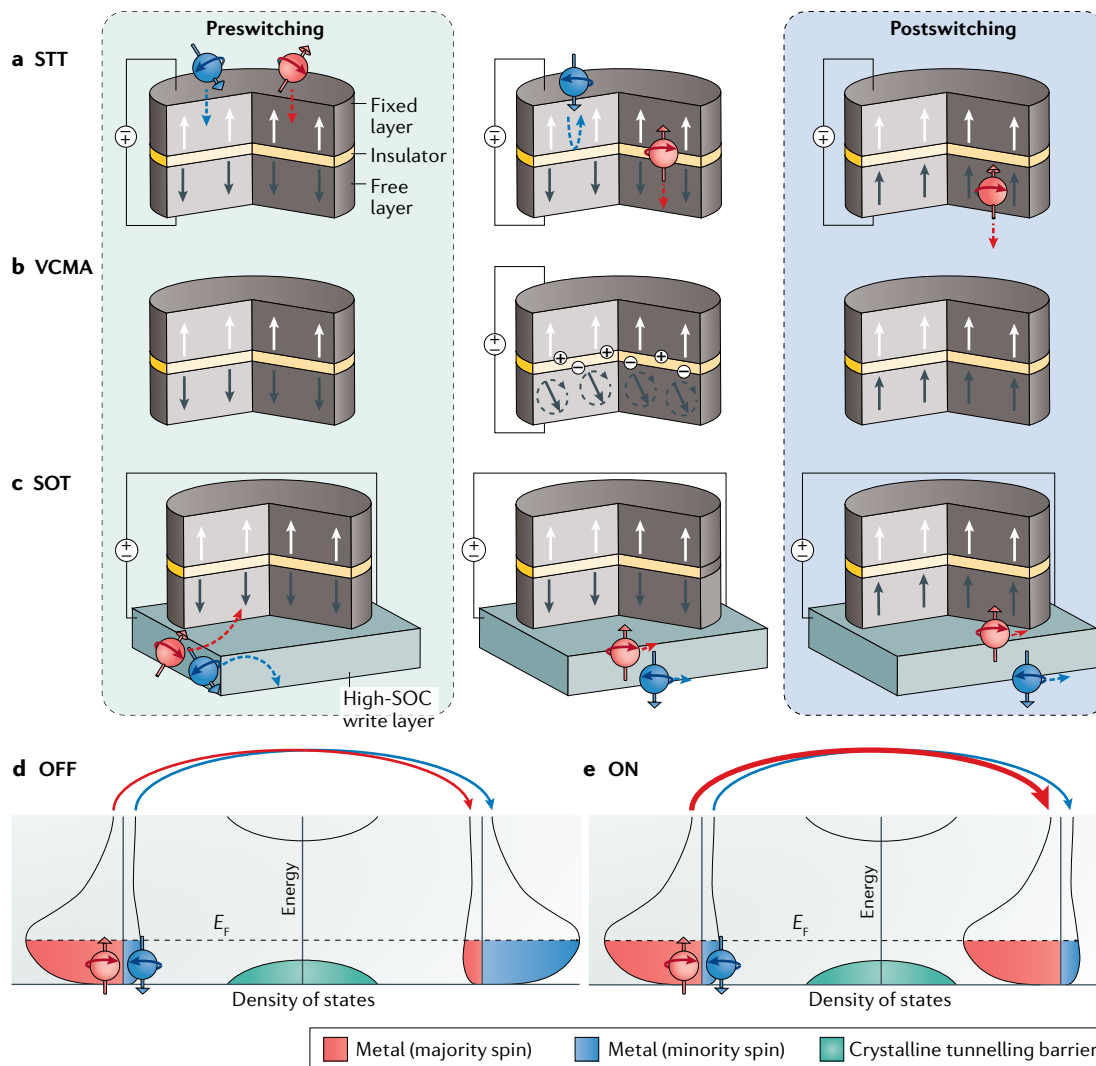

Fig. 3 | **Magnetic-tunnelling RSMs. a** | Spin-transfer torque (STT) switching. When electrons flow through the magnetically fixed layer, they become spin polarized by the magnetic moments in the fixed layer (white arrows). Electrons with the same orientation (red balls) as the fixed layer pass through the tunnelling barrier, whereas those with the opposite orientation (blue balls) are reflected. The electrons that pass through thus transfer their spin angular momentum to the moments of the magnetic free layer (grey arrows), thereby switching its state. **b** | Voltage-controlled magnetic anisotropy (VCMA) switching. When a voltage is applied across a magnetic tunnelling junction, charge accumulation or depletion at the magnetic free layer/MgO interface modifies the anisotropy of the magnetic free layer, leading to a precessional (oscillating) motion of the free-layer magnetization. A pulse timed to be half that of the precession period inverts the free-layer magnetization and, thus, the device state. **c** | Spin–orbit torque (SOT) switching. To facilitate this mechanism, the magnetic tunnelling junction is placed on top of a material with a high spin–orbit coupling (SOC). As current passes through this material, electrons with one spin orientation accumulate on top of the material, while those of the opposite spin accumulate at the bottom of the material. The electrons accumulated on top flow into and transfer their spin momenta to the free layer, thereby changing its magnetization. **d,e** | Density of states of the ferromagnetic layer and the tunnelling barrier in the OFF state (panel **d**) and the ON state (panel **e**). As the spin-polarized tunnelling probability is approximately proportional to the density of states of both ferromagnetic layers, the parallel magnetization orientation yields a larger tunnelling conductance than that of the antiparallel orientation because the majority-spin and minority-spin electrons tunnel into the majority and minority states, respectively, of the other ferromagnetic layer. By contrast, when the ferromagnetic layers have antiparallel magnetization orientations, the majority-spin and minority-spin electrons tunnel into the minority and majority states, respectively. E_F, Fermi level.
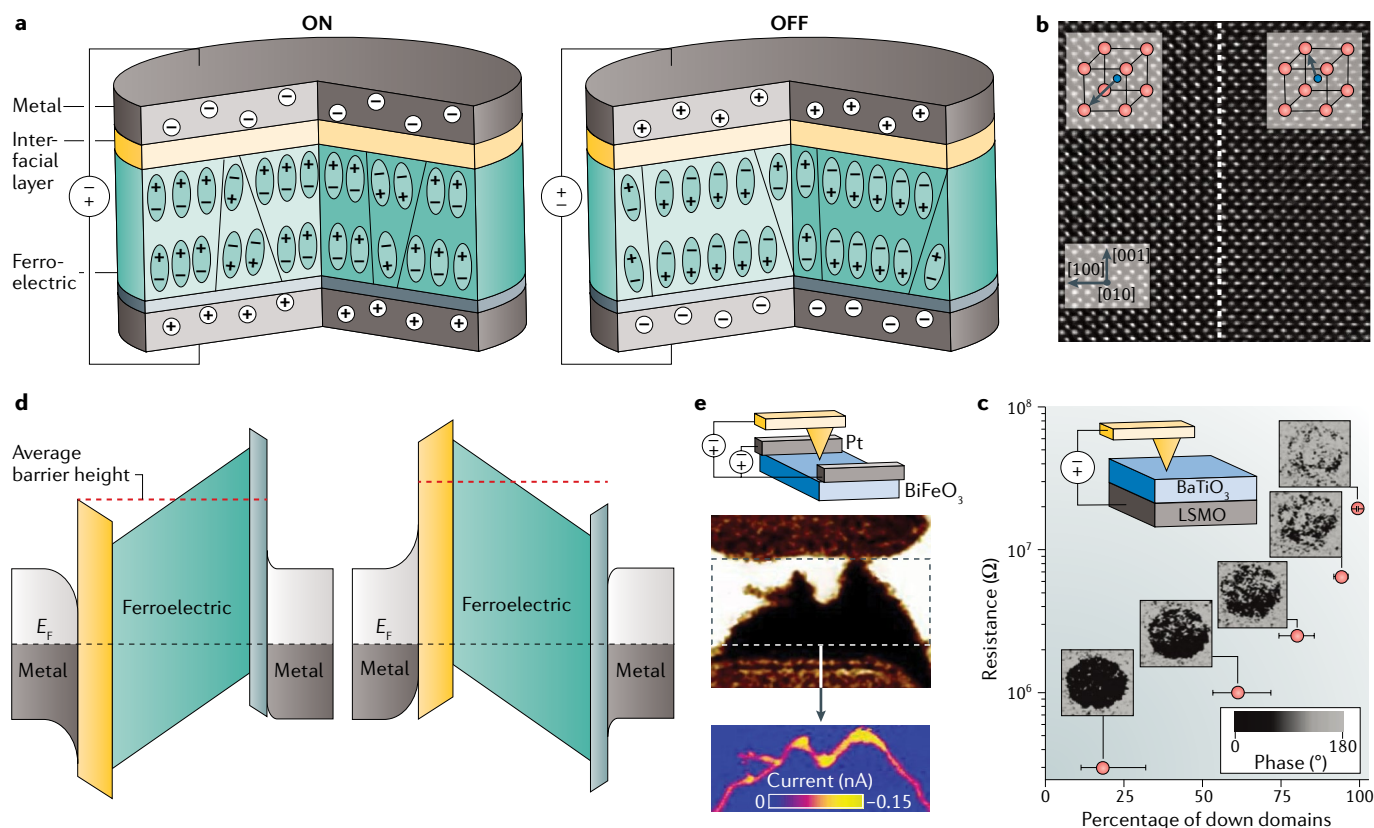
Fig. 4 | **Ferroelectric RSMs. a** | Illustration of two different polarized states (ON (left) and OFF (right)) of a ferroelectric tunnelling junction (FTJ), in which a ferroelectric is sandwiched by non-ferroelectric interfacial layers and electrodes of different screening capabilities. **b** | Experimental observation of a 109° domain wall (white dashed line) along the [010] zone axis of a cross-sectioned $BiFeO_3$ ferroelectric thin film grown on a (110) surface, obtained using exit-wave-reconstructed high-resolution transmission electron microscopy. The insets show the unit cells, with the Fe ions shown in blue. **c** | Resistance of an epitaxial $BaTiO_3$-based FTJ (where LSMO is $La_{0.67}Sr_{0.33}MnO_3$) versus the relative fraction of downwards polarization domains, as extracted from the piezoresponse force microscopy phase images. The points were acquired during the RESET process. The resistance gradually increases with the size of the down domains. **d** | Energy-band diagrams of the FTJ in the two different polarized states shown in panel **a**. The tunnelling resistance of the FTJ is dependent on the polarization. The average barrier height of the ON state (left part) is lower than that of the OFF state (right part). **e** | In-plane piezoresponse force microscopy phase image (middle part) of volatile 71° domain walls in a mesa-structured, $BiFeO_3$-based junction (top part). The lower part shows the concurrent conductive atomic force microscopy current map, indicating the large conductivity at domain walls. E_F, Fermi level. Panel **b** is adapted from REF.[127], Springer Nature Limited. Panel **c** is adapted from REF.[144], Springer Nature Limited. Panel **e** is adapted from REF.[132], Springer Nature Limited.

into the minority and majority (majority and minority) states, respectively, of the other ferromagnetic layer.

Ferroelectric-polarization-induced RSMs

Ferroelectric materials feature reversible remnant polarization, which can be changed by applying an external electric field[13]. In such materials, the symmetry-breaking polarized states correspond to the thermodynamic energy minimum according to the Landau–Devonshire theory[119]. These polarized states can encode bits, which led to the commercialization of ferroelectric random-access memory in the late 1990s[120]. Ferroelectric random-access memory has a similar construction to that of dynamic random-access memory but suffers from poor scaling, destructive capacitive readout and high processing costs on materials such as $Pb(Zr,Ti)O_3$ (REF.[121]). The emergence of thin films based on doped-HfO_2 or $(Hf,Zr)O_2$ (REFS[122–125]), which can be processed using atomic layer deposition, revived interest

in ferroelectric-based memories. Remarkable progress has been made in the development of ferroelectric tunnel junctions (FTJs)[12], which were conceptualized in the 1970s[126], and emerging domain-wall RSMs[127–132].

Switching mechanism. Both FTJs and domain-wall RSMs operate on the compensation of polarization charges. The two polarization-terminating surfaces of the ferroelectric layer possess uncompensated polarization charges (FIG. 4a), which are partially screened by the free carriers of the contacting electrodes[133]. According to Thomas–Fermi theory, the screening length is determined by the density of states at the Fermi level, although the effective screening length depends on the microscopic properties of the ferroelectric/electrode interface[12,134]. Incomplete screening yields depolarization fields that can reverse the polarization of an extremely thin film (<3 nm) and, thus, degrade the retention of the polarization state. The compensation of polarization

charges also occurs inside the ferroelectric thin film. Each domain, with a small volume of identical polarization, can collect compensating charge, possibly owing to the accumulation of defects[129,135], voltage-tunable distortion states[136], a high-symmetry phase[137] or a reduced local bandgap at the domain walls[127]. It has been suggested that these compensating charge injections to the domain walls between the reversed domains are the clue to solving the well-known Landauer's paradox[138].

Classical thermodynamics portrays polarization switching as a state transition in a ferroelectric double-well potential[119], which could show notable differences between the predicted and experimental coercive fields[139]. The localized nucleation of domains and subsequent coalescence achieved by domain-wall motion has been the source of this complication[133]. The microscopic structure of domain walls has been examined in various ferroelectric materials, such as epitaxial $PbZr_{0.2}Ti_{0.8}O_3$ (REFS[140,141]) and $BiFeO_3$ (REFS[127,129,131,142]) using atomic-resolution TEM. For example, this technique was used to image a 109° domain wall along the [010] zone axis of a cross-sectioned $BiFeO_3$ film grown on a (110) surface (FIG. 4b). The polarization direction was determined by analysing the relative displacement of an Fe ion from the centre of the unit cell; the Fe-ion displacement changed at the domain wall[127] (indicated by the dashed line in FIG. 4b). Further microscopy studies suggest the existence of lattice defects, such as Bi vacancies, at the domain walls in $BiFeO_3$ thin films[129]. The switching of multidomain FTJs could show analogue and memristive conductance changes[143,144]. For example, there is a correlation between the multilevel resistance and the ferroelectric domain configuration of a FTJ with an epitaxial $BaTiO_3$ RSM (FIG. 4c); this behaviour could be a useful asset for analogue computing[144]. Such analogue switching has recently been reported for $(Hf,Zr)O_2$-based devices[145,146]. Domain-wall motion has also been investigated theoretically using large-scale molecular dynamics simulations, which revealed that the diffuseness of the domain wall is dependent on both the temperature and the electric field[147].

Transport mechanism. The FTJ is a metal/ferroelectric/metal junction with two polarization directions of the ultrathin ferroelectric film. Switching the polarization changes the tunnelling barrier and the electron-tunnelling resistance at a small bias. Such a polarization-dependent difference in the tunnelling barrier (or the low-temperature tunnelling probability, which is exponentially dependent on the square root of the average barrier height (red dashed lines in FIG. 4d)) is dependent on the asymmetry of the two ferroelectric/electrode interfaces. Asymmetry arises from a difference in the thickness of the non-ferroelectric interfacial layers[148] and/or screening capabilities of the electrodes[149] (FIG. 4a,d), and can be further enhanced through the introduction of a polarization-induced Schottky barrier[150] or metal–insulator transition[151]. The converse piezoelectric effect has been suggested to contribute to the polarization-induced transport changes, owing to the modulation of the barrier thickness, electron effective mass and position of the conduction-band edge[152].

Although tunnelling and thermionic emission are primarily responsible for the conductance switching of FTJs, the structural deviation of a domain wall from its associated domains and the polarization discontinuities of adjacent domains can also affect its conductance; this effect can be used for both memory and computing applications[128,130]. Density functional theory simulations on the observed domain walls in $BiFeO_3$ (FIG. 4b) suggest that there is a reduced local bandgap at the domain walls[127] of specific angles. Electric-field-induced distortion of the polarization structure might lead to voltage-tunable electronic states that aid charge transport[136]. Local structures of the domain wall of the $BiFeO_3$ thin film could have a centrosymmetric phase rather than the ferroelectric phase of the associated domain[137]. Lattice defects, including oxygen vacancies[135] and bismuth vacancies[129], might exist in domain walls and could also result in the observed tunable domain-wall conductance.

The electrical current through domain walls, however, is usually too small for electronic applications. Recent improvements in nanofabrication and domain control have increased the domain-wall current to several tens of nanoamperes[132]. An even more intriguing finding is the volatile formation and spontaneous retraction of conducting domain walls in a mesa-structured $BiFeO_3$ junction[132] (FIG. 4e). Application of a voltage across the junction temporarily induces partially switched domains and highly conductive domain walls, which could be responsible for the large cross-junction current[132]. The cross-junction conductance drops when the electrical bias ceases, suggesting the disappearance of the conductive domain walls[132].

RSM device properties

==Although the resistance changes of the four RSMs are manifestations of different underlying physics, RSMs share common device properties (FIG. 5), which enable various computing applications.==

Stochasticity and number of states

Different computing tasks have different requirements of the information-representation capabilities. For both neural networks and analogue memcomputing, performance is largely dictated by the number of bits that RSMs can represent. For instance, a large number of distinguishable physical states or a large ratio between the conductance range and the variability of programming and reading means a greater number of effective bits per device. However, applications such as cybersecurity could benefit from a certain level of randomness, whereas digital memcomputing operates on binary states (FIG. 5).

Both redox and phase-change RSMs exhibit relatively large inherent programming variability because of the significant rearrangements of atoms involved in the stochastic switching process. The programming variability can be suppressed by 1D channels in epitaxial dielectrics[153], confinement nanolayers[154] and iterative programming algorithms[85], which have led to demonstrations of 64-level and 16-level usable conductance in a redox RSM[26] and a doped-$Ge_2Sb_2Te_5$ phase-change RSM[85], respectively.

Magnetoresistive and ferroelectric RSMs operate on domain effects. Typical MTJs feature a single domain and the lowest programming variability of the different RSMs, at the expense of encoding only a single bit per device. STT MTJs feature a low programming error rate ($<10^{-6}$) across a wide operating temperature range
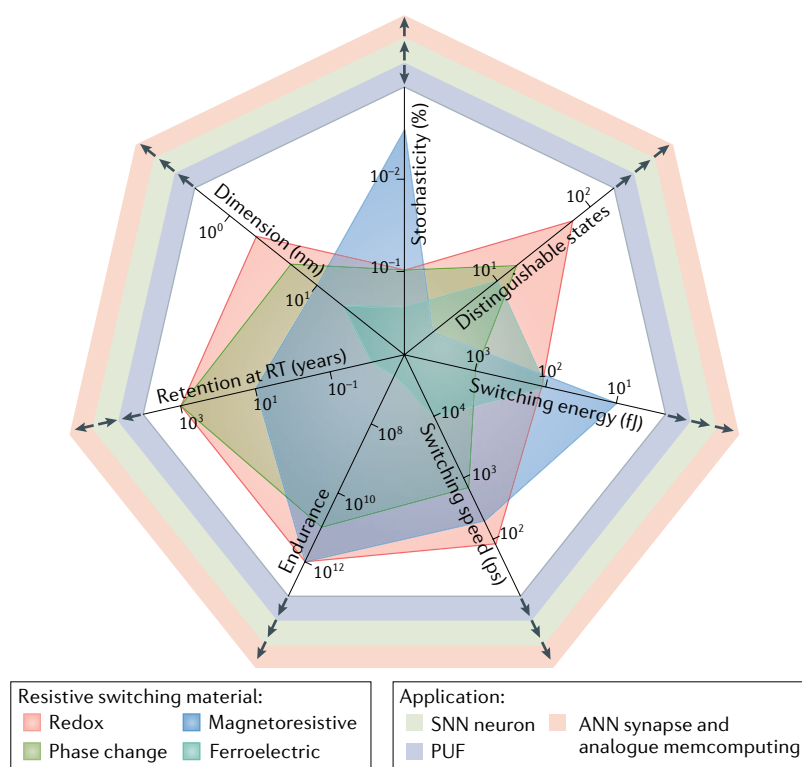
(approximately -40–$150\,°C$)[105], whereas VCMA MTJs have a moderately high programming error rate ($\geq10^{-5}$), owing to the precessional nature of the switching[155]. A multidomain $BiFeO_3$ FTJ yields ~10 usable conductance levels by varying the amplitude or the number of applied programming pulses[144]. However, scaling such devices down to their domain size decreases the number of analogue states. A Schottky barrier was introduced in addition to the tunnel barrier in a $BaTiO_3/Nb:SrTiO_3$ FTJ to enlarge the dynamic range, leading to an ON/OFF ratio of ~10^4 (REF.[150]). Furthermore, topological protection has been explored as an approach to generate highly controllable and repeatable domains and domain walls, such as in $BiFeO_3$ nanoislands, and, thus, reduced stochasticity[131].

Switching energy and speed

A small switching energy and a fast switching speed benefit almost all types of computing (FIG. 5), particularly applications that demand frequent device programming, such as digital memcomputing.

For redox and phase-change RSMs, the switching energy is mainly dominated by the RESET switching process. Although a RESET energy of 6 fJ has been reported[156], typical anion-type redox RSMs show a minimal switching energy of ~100 fJ (REF.[157]). The RESET energy of a phase-change RSM is roughly proportional to its switching volume; thus, devices with carbon-nanotube electrodes can achieve RESET energies of ~100 fJ (REF.[158]), although a more typical minimal value is ~1,000 fJ (REF.[159]). The switching speed is normally limited by the redox reaction rate and ion migration speed in redox RSMs or the growth and nucleation speeds in phase-change RSMs. Both types of RSM show a dependence on the applied electric field, temperature and choice of the active species. For instance, the fastest switching speed of 85 ps was observed in a redox RSM[160] operating on the migration of nitrogen vacancies, and the fastest crystallization speed of 800 ps was reported in a scandium-doped antimony telluride phase-change RSM[161]. Prestructural ordering has been investigated as an approach to accelerate the crystallization of $Ge_2Sb_2Te_5$ phase-change RSMs, albeit at the expense of the energy and time of preprogramming voltages[162].

For magnetoresistive and ferroelectric RSMs, the changes in magnetic or electric polarization involve minimal atom displacement and, thus, typically require a smaller switching energy and a shorter switching time than those of redox and phase-change RSMs. For STT MTJs, a switching time of $<10\,ns$ and a low switching current ($<50\,\mu A$) have been realized with an engineered free layer[163]. As no static current is needed for programming precessional VCMA MTJs, the switching energy is mostly reduced to the charging and discharging energy of the MTJ capacitor; thus, the lowest switching energy ($<10\,fJ$) among all RSMs has been reported for a $Co_{20}Fe_{60}B_{20}/MgO/Co_{20}Fe_{60}B_{20}/Ta/[Co/Pd]_{10}$ MTJ with a diameter of 50 nm (REF.[108]). In addition, a related STT MTJ switched at ~200 ps with low-voltage pulses[164], approaching the performance of static random-access memory. For ferroelectric RSMs, the switching of a $Co/BaTiO_3/La_{0.67}Sr_{0.33}MnO_3$ FTJ was achieved by

Resistive switching material:
- Redox
- Phase change
- Magnetoresistive
- Ferroelectric

Application:
- SNN neuron
- PUF
- ANN synapse and analogue memcomputing

Comparison of selected properties

RSM	Min. stochasticity (%)	Max. no. of distinguishable states	Min. switching energy (fJ)	Max. switching speed (ps)	Max. endurance (cycles)	Max. retention at RT (years)	Min. dimensions (nm)
Redox	~9.65 (ref.[300])	64 levels (ref.[26])	115 (ref.[157])	85 (ref.[160])	10^{12} (ref.[170])	>1,000 (ref.[165])	~2 (ref.[179])
Phase change	~9.62 (ref.[85])	16 levels (refs[84,85])	1,000 (ref.[159])	700 (ref.[161])	10^{11} (ref.[174])	>1,000 (ref.[166])	~5 (refs[158,159])
Magneto-resistive	~0.29 (ref.[176])	2 levels (refs[113,276])	10 (ref.[108])	200 (ref.[164])	10^{12} (ref.[176])	10 (ref.[101])	~10 (ref.[185])
Ferro-electric	~24.5 (ref.[131])	~10 levels (refs[143,144,177])	100 (ref.[143])	10,000 (ref.[143])	4×10^6 (ref.[177])	0.0078 (ref.[168])	~20 (ref.[186])

Fig. 5 | **Properties of RSMs and application requirements.** Selected properties of four types of resistive switching material (RSM) are compared in the plot. The maximum or minimum values of the properties are listed in the table. The outer coloured rings indicate the impact of the RSM properties on applications in artificial neural network (ANN) synapses and analogue memcomputing, spiking neural network (SNN) neurons and physical unclonable functions (PUFs). The example applications (illustrated in FIGS 6, 7 and 8) typically benefit from a low switching energy, rapid switching speed, high endurance and small dimensions. ANN synapses and SNN neurons favour low stochasticity, whereas PUFs could be aided by a certain amount of randomness. Moreover, ANN synapses and PUFs favour longer retention, whereas SNN neurons can benefit from fast, spontaneous conductance relaxation. Note that the minimal stochasticity is estimated as $2\{\sigma[\log(R_{HRS})]+\sigma[\log(R_{LRS})]\}\,/\,[\overline{\log(R_{HRS})}-\overline{\log(R_{LRS})}]$, where R_{HRS} and R_{LRS} are the distributions of the resistance in the high-resistance state and the low-resistance state, respectively. $\sigma[\log(R_{HRS})]$ and $\overline{\log(R_{HRS})}$ are the standard deviation and the mean of $\log(R_{HRS})$, with equivalent definitions for $\log(R_{LRS})$. RT, room temperature. Data from REFS[26,84,85,101,108,113,131,143,144,157–161,164–166,168,170,174,176,177,179,185,186,276,300].

applying 10-ns voltage pulses with an associated switching energy of ~100 fJ (~10 fJ) for devices with a 200-nm (50-nm) diameter[143].

Retention and endurance

Data stability is crucial for synapses of artificial neural networks (ANNs), memcomputing and hardware identities (FIG. 5). However, fast conductance decay can be exploited to emulate certain synaptic and neural dynamics and to implement random number generation.

An increase in the potential barrier between switchable states increases the switching energy but also increases the retention of the conductance state, with the retention largely conforming to the Arrhenius dependence on temperature. Both redox and phase-change RSMs can achieve a retention of >10 years at room temperature using, for example, Ta–O conduction channels[165] or a high-crystallization-temperature phase-change RSM with germanium enrichment and nitrogen doping[166] (FIG. 5). The thermal stability factor of MTJs is linearly proportional to the energy barrier, which scales with the device volume. STT MTJ arrays with large thermal stability factors (retentions of >10 years) have been reported by Samsung and Intel[104,167]. By contrast, the longest measured retentions for ferroelectric RSMs are 68 h and 278 h at room temperature for a $BiFeO_3$ FTJ[168] and a three-terminal cell[132] (the bulk polarization of which can be non-destructively read by domain-wall currents between the middle electrode and an edge electrode), respectively.

A high cycling switching endurance (that is, the maximum number of switching cycles) benefits all types of computing, particularly those involving frequent programming, such as digital memcomputing (FIG. 5). Endurance failure of devices may originate from structural fatigue, particularly in redox and phase-change RSMs. For redox RSMs, these fatigues include undesired redox reactions with electrodes, overgrowth of filaments or unwanted diffusion (or loss) of filament atoms[169]. A thermodynamically stable redox RSM can achieve an endurance of up to 10^{12} cycles[170] (FIG. 5). For phase-change RSMs, the endurance is closely correlated with the electromigration[171] and, thus, phase segregation within the melt[172], together with the mechanical stress due to a change in density[173]. The endurance of phase-change RSMs can be increased by engineering the film-deposition technique, composition and device geometry. Using this approach, the endurance of a 7.5-nm × 17-nm Sb-rich Ge–Sb–Te phase-change RSM device was measured up to 10^{11} cycles[174] (FIG. 5).

Magnetoresistive and ferroelectric RSMs are predicted to have high endurances, as they exhibit less structural fatigue than other types of RSM, owing to the minimal atomic displacement during switching. STT MTJs have been reported with an endurance of 10^{10} and moderate thermal stability after mitigation of damage caused by reactive ion etching[175]. SOT MTJs based on CoFeB/MgO/CoFeB stacks have been switched 10^{12} times, as no direct programming current flows through the junction[176]. Although a large endurance has been predicted for ferroelectric RSMs, so far, endurances of only 4×10^6 (REF.[177]) and 10^7 (REF.[132]) cycles have

been demonstrated for FTJ and volatile domain-wall devices, respectively.

Small footprint and stackability

A high device density increases the portability and decreases the cost of computing systems, and is, thus, a universal advantage to all applications (FIG. 5), especially to those such as neural networks and memcomputing, which demand a large number of RSM devices.

In most cases, redox-RSM devices are scalable, owing to their localized conduction channels. Redox HfO_x devices with a junction size of 10×10 nm^2 exhibited fast and reliable operation[178], and low-current HfO_2/TiO_x RSM crossbars with a half-pitch of 6 nm and a critical dimension of 2 nm have been demonstrated[179]. TaO_x-based redox-RSM devices have also been integrated on an Intel 22FFL (22-nm FinFET low-power) node in a 7.2-Mbit array[180]. The scalability of phase-change RSMs was verified with carbon-nanotube electrodes with a contact area of ~5 nm^2 (REFS[158,159]). The scaling of phase-change RSM devices also decreased the switching energy and enabled the integration of an 8-Gb phase-change random-access memory on a 20-nm node chip using diode-based accessing devices[181]. The stackability of redox and phase-change RSMs has been experimentally demonstrated with the fabrication of an eight-layer 3D vertical $TiN/HfO_2/TaO_x/Ti/TiN/W$ array[182] and the 'Optane memory' of Micron and Intel[183], respectively.

For magnetoresistive and ferroelectric RSMs based on domain switching, a moderately large footprint increases retention and the number of states; however, this requires a trade-off between the density and other properties according to the application. STT MTJs have demonstrated megabit integration on 22-nm or 28-nm technology nodes[104–106], with a scaling capability down to nodes of <10 nm (REFS[184,185]). Ferroelectric RSMs based on $BaTiO_3$ have been scaled down to 20 nm with Ag nanoisland electrodes[186].

RSM-based neuromorphic computing

Inspired by and abstracted from how brains function, neuromorphic computing is a key technology to empower artificial intelligence. RSM-based synapses, neurons and their networks use the physical properties of RSMs to process information in a brain-inspired manner.

Synapses and neurons

The variable conductance of an RSM mimics the synaptic weight by modulating the charge transmission at a given voltage; thus, RSMs are a compact hardware implementation of synapses for ANNs and spiking neural networks (SNNs). In addition, spikes encoded with spatiotemporal information and generated by SNN neurons can induce synapses to change their weight through both short-term and long-term plasticity, which has been efficiently emulated by RSM switching dynamics[56,153,187–192].

Redox RSMs with highly mobile cations[188] or anions[193] may experience spontaneous conductance decay, which can emulate short-term plasticity. The synaptic weight (or conductance; FIG. 6a, bottom left)

of a redox-RSM synapse evolves upon application of a train of voltage pulses (FIG. 6a, top left): the conductance increases upon excitation with a voltage pulse and gradually relaxes between pulses. For example, this behaviour was observed in a Ag_2S redox RSM and was attributed to the voltage-induced formation and spontaneous decay of a Ag filament[188]. Long-term plasticity has been achieved by controllably driving Ag cations in and out of dislocations within an epitaxial SiGe film[153], resulting in non-volatile linear synaptic potentiation
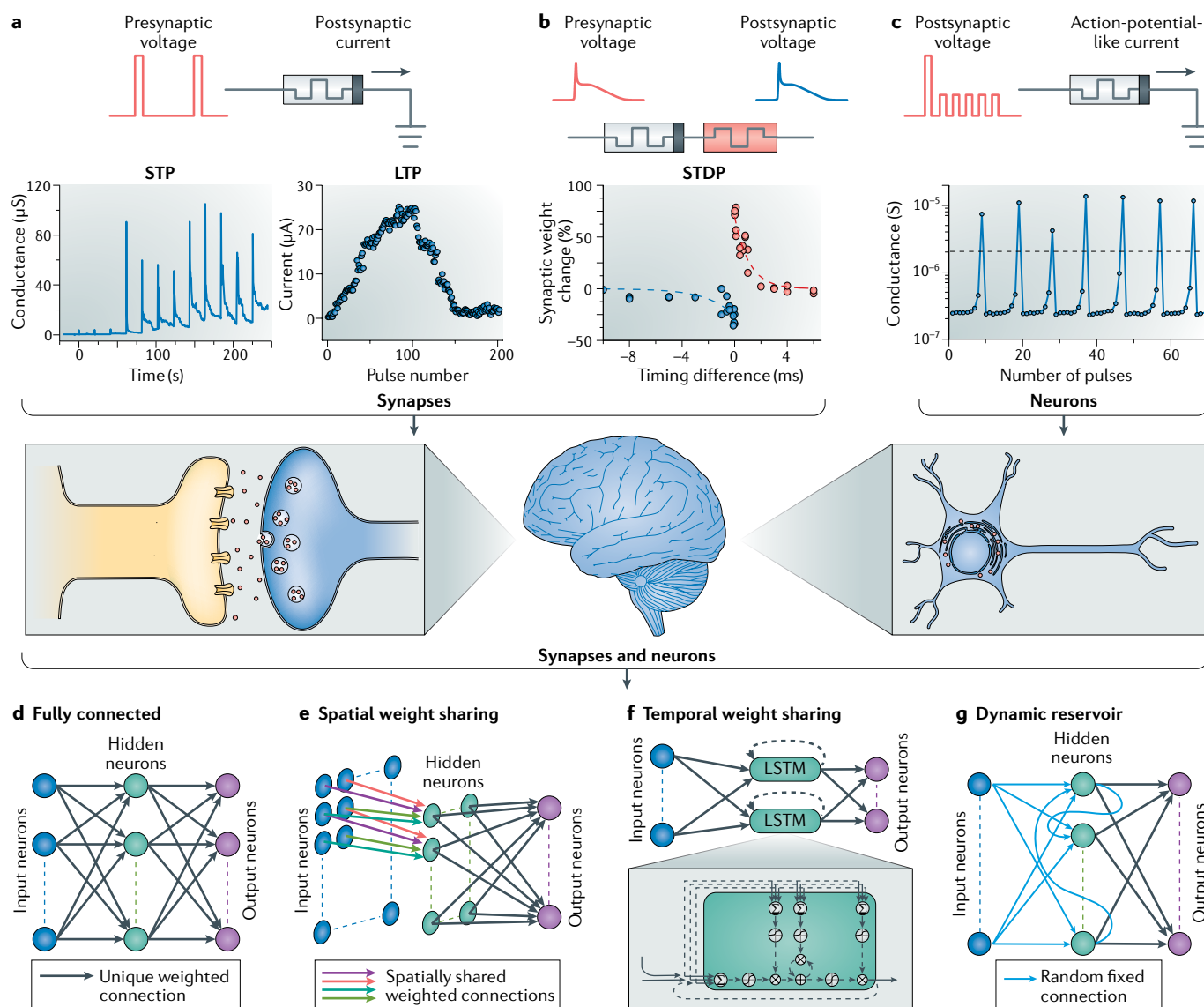

Fig. 6 | **RSM neuromorphic computing applications. a** | Resistive switching material (RSM) synapses for short-term and long-term potentiation and depression. The synaptic weight (or conductance; bottom) evolves with the application of voltage pulses (top). Short-term plasticity (STP) can be implemented by application of electric pulses, which cause the formation and spontaneous decay of a Ag filament in a Ag_2S redox RSM[188]; the conductance relaxes between pulses (bottom left). Long-term plasticity (LTP) can be achieved by using a voltage pulse to drive Ag filament formation and rupture in an epitaxial SiGe film[153], resulting in non-volatile analogue conductance (current) potentiation and depression (bottom right). **b** | An RSM synapse, consisting of serially connected non-volatile (middle part, grey) and volatile (middle part, red) RSM devices, for spike-timing dependent plasticity (STDP). The synaptic weight change of STDP (bottom) is determined by the relative timing between the presynaptic and postsynaptic voltage pulses[56] (top). **c** | An RSM neuron for integrate-and-fire. The doped $Ge_2Sb_2Te_5$ phase-change RSM neuron sums up postsynaptic voltage spikes over time and produces action-potential-like conductance spikes (bottom) once the integration threshold (dashed line) is reached[196]. The tall orange spike (top left) was used to RESET the device after the fire. **d** | Illustration of a two-layer, fully connected network. The neurons are represented by circles and the black arrows represent synapses (for example, RSM devices with their conductance as weights)[213]. **e** | Illustration of a spatial weight-sharing network in which a convolutional layer with a 2×2 kernel is followed by a fully connected layer. The neurons are represented by circles and the coloured arrows are synapses (for instance, RSM devices) of the localized neural connections shared among all hidden neurons (green circles)[226]. **f** | Illustration of a temporal weight-sharing network in which a long short-term memory (LSTM) layer is followed by a fully connected layer. The inset shows the internal structure of an LSTM unit[229]. **g** | Illustration of a reservoir computing network in which a dynamic reservoir, of random but fixed connections, is followed by a trainable, fully connected readout layer[234]. The bottom sections of panel **a** are adapted from REFS[153,188], Springer Nature Limited. The bottom section of panel **b** is adapted from REF.[56], Springer Nature Limited. The bottom section of panel **c** is adapted from REF.[196], Springer Nature Limited.

and depression (or current; FIG. 6a, bottom right) upon application of trains of voltage pulses of different polarities. Spike-timing-dependent plasticity (STDP) is a widely used long-term learning rule in SNNs that has been implemented with redox[187,191], phase-change[190,194] and ferroelectric[128,145,146] RSMs and envisioned for magnetoresistive[195] RSM synapses. The synaptic weight change of STDP (FIG. 6b, bottom) is determined by the relative timing between the presynaptic and postsynaptic voltage pulses (FIG. 6b, top), which originate from the ionic (for example, Ca^{2+}) diffusive dynamics of biological synapses. Conventional non-volatile RSMs lack such dynamics and, thus, waveform engineering is used to simulate STDP. However, because the ionic dynamics in some volatile RSMs are similar to those in chemical synapses, STDP was naturally realized in a combined device consisting of a redox volatile $Pt/SiO_xN_y{:}Ag/Pt$ RSM device and a non-volatile $Pt/TaO_x/Ta$ RSM device[56] (FIG. 6b, middle).

In addition to RSM synapses, phase-change (including Mott-type)[196,197], magnetoresistive[198,199] and redox-diffusive[200] RSMs are capable of simulating the integrate-and-fire neuron model (FIG. 6c). For example, the doped $Ge_2Sb_2Te_5$ phase-change-RSM neuron integrated post-synaptic voltage spikes (FIG. 6c, top left) over time and produced an action-potential-like conductance or current spikes (FIG. 6c, bottom) when the integration threshold was reached[196]. Moreover, Mott-type phase-change[201,202], Ovonic electronic[203] and magnetoresistive[204] RSMs have been explored to replicate oscillatory neural behaviour, leading to chaotic dynamics for accelerating Hopfield network convergence[202] and reservoir computing for speech classification[204].

Neural networks

Approximately 10^{14-15} synapses and 10^{11} neurons constitute the core of the human brain. Each neuron communicates through synapses to several thousands of other neurons. The topology of the connectivity endows the networks with information-processing capabilities. Two types of topologies are widely used in RSM networks, namely, fully connected and weight shared.

Fully connected networks.

Fully connected networks are layered structures wherein each neuron (circles in FIG. 6d) is connected to all neurons in the previous layer. The neurons receive stimuli through synapses (black arrows) and generate outputs to synapses of the next layer. Depending on how the neurons encode information, two types of networks have generally been adopted: the ANN and the SNN.

For ANNs, neurons feature static and continuous-valued outputs, without an intrinsic dependence on the history of neural activity. The adjustment of synaptic weight, or the 'learning process', is determined by mathematical optimization methods, such as stochastic gradient descent with error backpropagation[205], which can be partially implemented in RSM devices. The learning process can also be ex situ by transferring pre-optimized synaptic weights learned on other platforms (such as digital computers) to the RSM synapses; this approach is more suitable for inference-intensive tasks[206].

ANN learning can also be classified as supervised, unsupervised or reinforcement, depending on the kind of information available to the network.

Supervised learning demands a collection of input samples (for example, images) that each have a labelled output (for example, image content). The fully connected network learns by adjusting the synaptic weights to map an input to the correct output using the training input–output pairs as examples. The learned networks can then predict the corresponding labels of new input samples. Using this approach, redox (for example, TiO_{2-x} (REF.[207]), $Pr_{0.7}Ca_{0.3}MnO_3$ (REF.[208]), Al_2O_3/TiO_{2-x} (REFS[17,209]), $HfAl_yO_x$ (REF.[210]), TaO_x (REF.[211]), WO_x (REF.[212]) and HfO_x (REFS[206,213])) and phase-change (for example, doped $Ge_2Sb_2Te_2$ (REFS[23,86])) RSM synapses have been used to classify patterns, including simple letters and images[17,207,209,212], electroencephalography signals[208], images in the Yale Face Database[210] and handwritten digits in the MNIST database[23,86,206,211,213,214]. A maximum energy efficiency of 66.5 tera operations per second per watt (TOPS W^{-1}) was achieved using a redox TaO_x RSM integrated on 40-nm technology nodes[211].

In contrast to supervised learning, unsupervised learning deals with unlabelled input samples. The fully connected network learns to cluster inputs that look similar to each other based on different optimization methods, such as the competitive learning rule of self-organizing maps. For instance, the generalized Hebbian algorithm[215] (which changes the synaptic weight according to the activation of the presynaptic and postsynaptic neurons) has been implemented on redox TaO_x (REF.[216]) or WO_x (REF.[212]) RSM synapses for principal component analysis of a breast cancer dataset. Similar Hebbian rules and winner-take-all neurons were used for offline learning of the dictionary elements encoded on a redox WO_x RSM crossbar for sparse coding[25,212] (discussed further in the section on data reduction).

Moreover, the fully connected network could learn how to make decisions in an unknown environment to maximize certain 'rewards' using reinforcement learning, by approximating either the value (or Q, which is the sum of future rewards weighted over time) or policy (probabilities of different actions) of a given state. In terms of Q-learning, redox-RSM synapses of a three-layer network acted as a Q-value approximator to balance a hinged pole on a moving cart and to drive an underpowered car out of a steep valley[217]. In addition, policy function was implemented in a FTJ crossbar for path finding, wherein the programming stochasticity helped to explore actions[218]. Fully connected networks also constituted the generator and discriminator sub-networks of a generative adversarial network, which generated outputs that looked like its inputs, using redox $HfAl_yO_x$ RSM synapses and the MNIST dataset[219].

Unlike ANNs, SNNs explicitly incorporate time as a computational dependence, similar to that of the brain. SNNs could be trained by biologically plausible local learning rules, whereby synapses autonomously change their weights under certain plasticity (such as STDP), driven by signals from the linked neurons. Such a learning mechanism could be facilitated by the intrinsic switching dynamics of RSMs. Alternatively, SNNs

could be trained by spike variants of backpropagation and ANN-to-SNN conversion[220].

Unsupervised learning with STDP in fully connected networks has been demonstrated with redox (for example, TiO_x (REF.[221]), HfO_x (REF.[200]), $AlTe_xO_y$ (REF.[222])) and phase-change (such as doped $Ge_2Sb_2Te_2$ (REFS[223,224])) RSM synapses and used to, for example, cluster patterns[200,221,222], implement auto-associative memory[223] and detect temporal correlations[224]. The synaptic weight changes were driven by integrate-and-fire spiking neurons that were built on transistors with lateral inhibition[221], or phase-change[224] and redox-diffusive[200] RSM neurons that could integrate-and-fire. A diffusive RSM neuron was able to physically program RSM synapses by increasing voltage drops onto the synapses when the neuron fires (that is, switched ON and, thus, has a smaller resistance)[200]. With spiking neurons that could receive teacher signals and synapses with transistor-based timed switches, supervised learning using STDP was demonstrated with redox HfO_x:Si RSM synapses, which could detect spatiotemporal patterns[225].

Weight-shared networks. State-of-the-art neural networks are mostly weight shared. Convolutional neural networks, which feature spatially shared weights, differ from fully connected networks because each neuron links to only a portion of the neurons in the previous layer (FIG. 6e), or the so-called receptive field, inspired by the visual cortex in the brain. More importantly, all neurons of a convolutional layer use the same set of synapses (coloured arrows in FIG. 6e, where each colour represents a unique, spatially shared synapse) to connect to neurons of the previous layer. Such spatial synaptic weight sharing could use the spatial correlation of the inputs to extract and combine local features. In situ supervised learning with a convolutional neural network based on a redox TaO_x RSM achieved similar accuracy in classifying the MNIST dataset to that of a fully connected RSM network, but using just ~25% of the number of synapses[226]. Ex situ supervised learning using convolutional neural networks with redox-RSM synapses (for example, TiO_xN_y) showed a maximum energy efficiency of 53.17 TOPS W^{-1} in classifying the MNIST and CIFAR-10 datasets, which is one of the highest efficiency values reported to date[227,228]. In addition, a hybrid software–hardware convolutional neural network that incorporated a phase-change RSM has been used to classify the CIFAR-10 and CIFAR-100 datasets; this was achieved by 'transfer learning' for the last fully connected layer using phase-change RSM synapses[23].

RSM synapses can also be temporally shared. Recurrent networks are those with loops in their computational graphs, such as ANNs with a long short-term memory (LSTM) layer. At each time step, an LSTM cell (green boxes in FIG. 6f) receives both external inputs from neurons of the previous layer and its own output from the previous time step. Its internal structures add and remove information of the cell based on weighted-summation performed by the same set of RSM synapses in all time steps, which could exploit the temporal correlation of the inputs for feature extraction. An LSTM-based recurrent network of redox HfO_x RSM

synapses was trained in situ using supervised learning for the classification of human walking gaits[229]. In addition, ex situ supervised training of an LSTM network of phase-change RSM synapses was applied for language modelling, such as word-level prediction using the Penn Treebank dataset[230]. Moreover, simultaneous spatial and temporal weight sharing was realized in a network comprising a convolutional LSTM layer for classifying short videos using redox TaO_x RSM synapses[226]. Other recurrent networks, such as the Hopfield networks of redox HfO_x (REF.[231]) and $HfAl_yO_x$ (REF.[232]) RSM synapses, have been demonstrated to implement associative memory.

For weight-shared SNNs, a convolutional SNN trained ex situ using supervised learning was implemented based on HfO_x RSM synapses with on-chip compact integrate-and-fire neurons for classifying the MNIST and CIFAR-10 datasets, showing an energy efficiency of up to 16.9 TOPS W^{-1} (REF.[233]). A special variant of the SNN recurrent network is the so-called dynamic reservoir (FIG. 6g), which is a layer of spiking neurons with recurrent, random and fixed synaptic weights that mimic the sparse and irregular neural connections of the brain. The reservoir maps the temporal input to the neural states, a vector in a high-dimensional space that serves as an input to a fully connected readout layer, which can be trained in a supervised way (FIG. 6g). As it is the readout layer that is trained, the reservoir can substantially reduce the training complexity compared with ordinary recurrent SNNs. Reservoirs of redox WO_x (REFS[234,235]) and Ag-diffusive[236] RSM neurons were used to classify the MNIST dataset; the former also solved second-order nonlinear tasks[234], classified spoken digits and forecasted the Mackey–Glass time series[235].

RSM general-purpose memcomputing

Information processing with neural networks, as described above, is usually termed a computationally 'soft' task, as it can withstand low computing precision[28]. RSMs are also used to accelerate computationally intensive tasks, such as analogue vector–matrix multiplications and digital Boolean logic.

Analogue memcomputing

Analogue memcomputing uses RSM crossbars as multiplier–accumulators (MACs) to compute a vector–matrix multiplication in a single computing cycle (FIG. 7a) using Ohm's law for multiplications and Kirchhoff's current law for accumulations in parallel. Alternatively, forcing analogue currents and measuring voltages could yield matrix inversion[237]. High-throughput and energy-efficient analogue memcomputing benefits applications such as data reduction, linear system and eigenvector solvers, and combinatorial optimization, as demonstrated in the following examples.

Data reduction.
Compressed images, such as those in JPEG formats, are commonly used. The core of the compression is the vector–matrix multiplication, which could be greatly accelerated using RSM crossbars. For example, a greyscale image represented by a matrix X in the spatial domain was converted to $[(X \cdot G)^T \cdot G]^T$ in

the frequency domain using a redox HfO_x RSM crossbar, with the discrete cosine transform matrix G represented by the RSM crossbar conductance matrix (FIG. 7b). $X \cdot G$ or $(X \cdot G)^T \cdot G$ was calculated by multiplying each row of X or $(X \cdot G)^T$ with G, respectively[26].

Compared with standard compression methods such as JPEG compression, neural networks could be better at compressing images that are similar to those that the neural networks were trained on. Sparse dictionary learning aims to find a sparse row vector **a** to
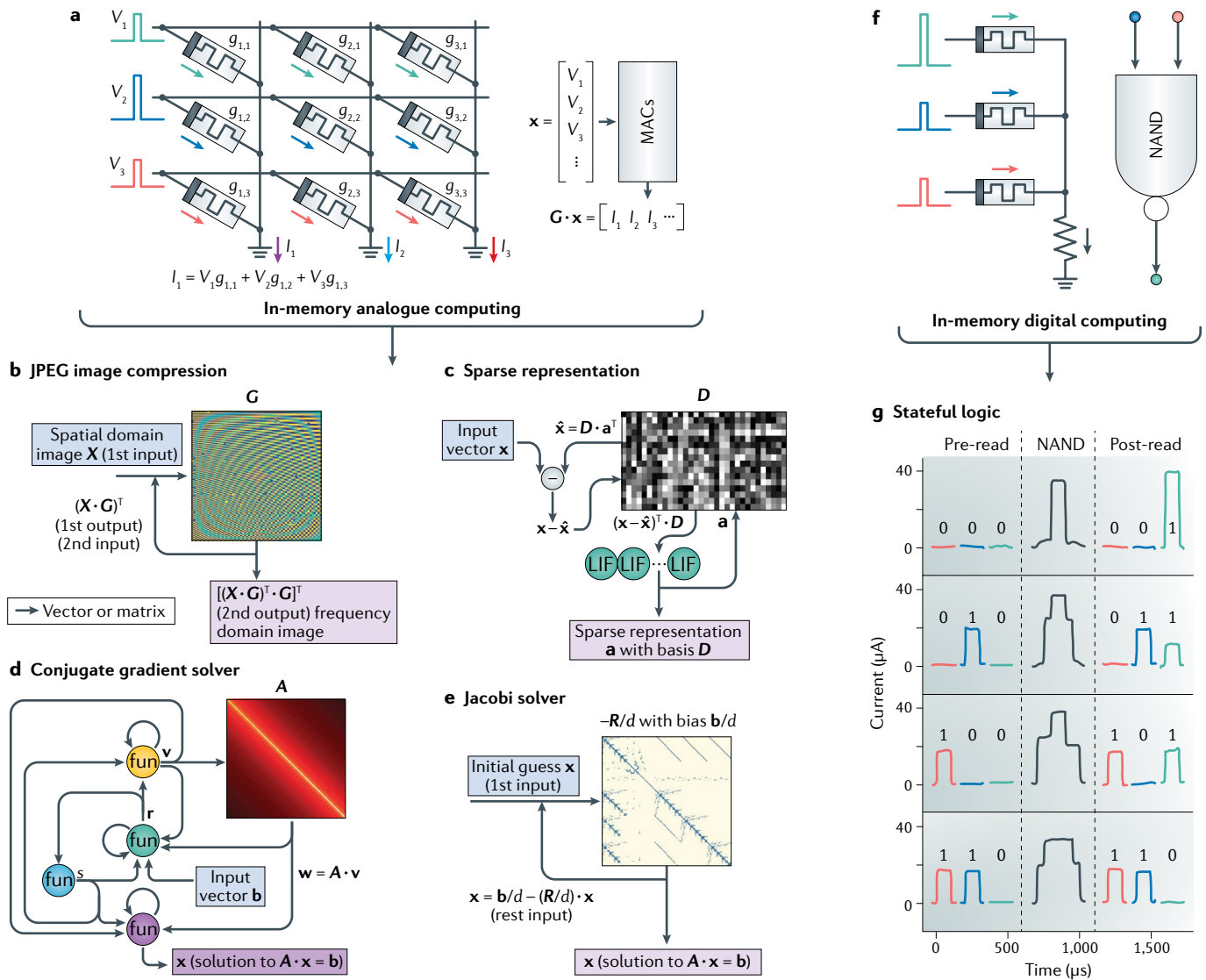

Fig. 7 | **RSM memcomputing applications. a** | Analogue memcomputing uses resistive switching material (RSM) crossbars as multiplier–accumulators (MACs) to compute the vector–matrix multiplication in a single step. The vector and matrix are represented by the row voltages (**x**, with elements V_i, where i is the row number) and the RSM-crossbar conductance matrix (G, with elements $g_{j,i}$, where j is the column number), respectively. The output vector is in the form of the column currents ($G \cdot \mathbf{x}$, with elements I_j). **b** | An RSM crossbar for JPEG image compression. An image in the spatial domain (matrix X) can be transformed into $[(X \cdot G)^T \cdot G]^T$ in the frequency domain, where the discrete cosine transform matrix G is represented by the conductance of a redox-RSM crossbar. $X \cdot G$ and $(X \cdot G)^T \cdot G$ are evaluated by multiplying each row of X and $(X \cdot G)^T$ with G, respectively[26]. **c** | A sparse vector **a** approximates the input column vector **x** using the dictionary D (for example, the conductance of an RSM crossbar), $\mathbf{x} \cong D \cdot \mathbf{a}^T$. An iterative process is used to compute **a**. The row vector $(\mathbf{x} - \hat{\mathbf{x}})^T$ is first physically multiplied with D. The result (row vector $(\mathbf{x} - \hat{\mathbf{x}})^T \cdot D$) is the input to the leaky integrate-and-fire neurons (LIFs, green circles), which output the row vector **a**. The reconstructed input $\hat{\mathbf{x}} = D \cdot \mathbf{a}^T$ is calculated physically in a backward pass. This process is repeated until convergence[25]. **d** | An RSM crossbar for

the conjugate gradient method. Finite iterations are used to solve $A \cdot \mathbf{x} = \mathbf{b}$ by updating the intermediate column vectors or scalar (**r**, **v**, s, **w**) and the solution **x** sequentially in each iteration, where A is mapped to the conductance of a computational memory based on a phase-change RSM to physically compute $\mathbf{w} = A \cdot \mathbf{v}$ (REF.[27]). **e** | The solution **x** to $A \cdot \mathbf{x} = \mathbf{b}$ was iteratively improved by multiplying **x** with the conductance matrix of redox Ta_2O_{5-x} RSM crossbars consisting of $-R/d$ (R is the non-diagonal part of A and d is the constant diagonal element of A) and bias $\mathbf{b}/d$ to implement $\mathbf{x} = \mathbf{b}/d - (R/d) \cdot \mathbf{x}$ (REF.[28]). **f** | Digital memcomputing with RSMs. The Boolean logic states ('0' and '1') are mapped onto the binary resistance states (ON state and OFF state). Stateful logic operations are based on voltage divisions between the RSM devices representing the input bits (blue and orange lines) and the device representing the output bit (green line), which implemented the NAND gate. **g** | In-memory NAND logic-gate operation using a redox HfO_2 RSM with the circuit shown in panel **f**[242]. fun, function. Panel **b** is adapted from REF.[26], Springer Nature Limited. Panel **c** is adapted from REF.[25], Springer Nature Limited. Panel **d** is adapted from REF.[27], Springer Nature Limited. Panel **e** is adapted from REF.[28], Springer Nature Limited. Panel **g** is adapted with permission from REF.[242], Wiley-VCH.

approximate the input column vector $\mathbf{x}$ (for example, a patch of an image) using the dictionary $\mathbf{D}$ (for example, the conductance of a WO_x RSM crossbar) as the basis, or $\mathbf{x} \cong \mathbf{D} \cdot \mathbf{a}^T$ (FIG. 7c). An iterative procedure, using the locally competitive algorithm to minimize an energy function, was used to produce $\mathbf{a}$ through two processes. The first is the vector–matrix multiplication $(\mathbf{x} - \hat{\mathbf{x}})^T \cdot \mathbf{D}$, which is the input to the leaky integrate-and-fire neurons (green circles in FIG. 7c). The neurons then output $\mathbf{a}$. The second process is to reconstruct the image with another vector–matrix multiplication, $\hat{\mathbf{x}} = \mathbf{D} \cdot \mathbf{a}^T$ (a backward pass that biases columns instead of rows of the RSM crossbar), for constructing the residue by removing the reconstructed input from the original input to achieve lateral inhibition, until the network converges to yield the sparse representation $\mathbf{a}$[25,212].

K-means clustering is another popular data-reduction method that partitions input vectors into K clusters with the nearest mean. The major operation, calculating the Euclidian distance between an input vector $\mathbf{u}$ and the centroid ($\mathbf{w}_i$) of cluster i, $\|\mathbf{u} - \mathbf{w}_i\|$, can be accelerated by using RSM crossbars. For example, a redox Ta_2O_{5-x} RSM crossbar was used to cluster the IRIS dataset[238].

Linear system and eigenvector solvers. Solving systems of linear equations such as $A \cdot \mathbf{x} = \mathbf{b}$ is crucial to scientific computing. Conjugate gradient methods use finite iterations to update the intermediate column vectors or scalar ($\mathbf{r}$, $\mathbf{v}$, s, $\mathbf{w}$) and the solution $\mathbf{x}$ sequentially in each iteration (FIG. 7d). A has been mapped to the conductance of a phase-change-RSM-based computational memory to compute the vector–matrix multiplication $\mathbf{w} = A \cdot \mathbf{v}$. Other steps were performed in digital circuits[27]. Such hybrid analogue–digital computing combines the high precision of digital computing with the area and energy efficiencies of analogue memcomputing.

Systems of linear equations may feature large but sparse matrices that can be partitioned into small blocks and represented by the conductance of RSM crossbars. The precision of an RSM crossbar can be improved by using multiple crossbars of different assigned significance so that each crossbar represents a given number of bits. These techniques enabled the Jacobi method to be used to iteratively improve the solution $\mathbf{x}$ to $A \cdot \mathbf{x} = \mathbf{b}$. The solution $\mathbf{x}$ in the previous iteration was multiplied with the conductance matrix (for example, that of a redox Ta_2O_{5-x} RSM crossbar[28]) that encoded $-R/d$ (where R is the non-diagonal part of A and d is the constant diagonal element of A) and bias $\mathbf{b}/d$, which effectively implemented $\mathbf{x} = \mathbf{b}/d - (R/d) \cdot \mathbf{x}$ (FIG. 7e).

Moreover, the linear system $A \cdot \mathbf{x} = \mathbf{b}$ was solved by forcing the analogue currents $\mathbf{I}$ (where $\mathbf{I} = \mathbf{b}$) flowing through the columns of a HfO_2 RSM crossbar of conductance matrix G (where $G = A$) and measuring the voltages $\mathbf{V} = G^{-1} \cdot \mathbf{I}$ (where $\mathbf{V} = \mathbf{x}$)[237] at the rows (FIG. 7a).

Similar to solving systems of linear equations, solving eigenvectors is central to numerous applications in scientific computing and could also greatly benefit from RSM crossbars. A square conductance matrix G with recurrent connections (or $\mathbf{V} = \mathbf{I}/g_\lambda$ in FIG. 7a, which could be implemented by linking the ith column of an RSM crossbar with the ith row through a transimpedance amplifier of a feedback conductance g_λ and an analogue inverter) could naturally solve the eigenvector $\mathbf{V}$ of $G \cdot \mathbf{V} = g_\lambda \cdot \mathbf{V}$ in a HfO_2 redox-RSM crossbar by measuring the row voltages $\mathbf{V}$ in the stable state[237].

Combinational optimization. For a spin glass in the state $\boldsymbol{\sigma}$ (a column vector in which the ith element of $\boldsymbol{\sigma}$ denotes the ith spin), the total energy is proportional to $-\boldsymbol{\sigma}^T \cdot (N \cdot \boldsymbol{\sigma})$, where the sparse and trigonal (accounting for the interactions of adjacent spins) coupling matrix N was sliced and programmed into TaO_x RSM crossbars for evaluation of the total energy change due to a spin flip[239]. The total energy of the system could be minimized by repeatedly flipping a random spin if it lowers the total energy or with a decaying probability if it increases the total energy.

Digital memcomputing

RSMs have also been proposed to implement digital logic. In these cases, Boolean logic states (1 and 0) are mapped to the resistance of the ON state and the OFF state of an RSM, enabling in-memory logic with the collocation of logic gates and latch memory[24,240–242]. Stateful and cascadable logic operations were based on voltage divisions between the RSM devices that represent the input bits (blue and orange lines in FIG. 7f) and the device that represents the output bit (green line in FIG. 7f). The RSM device for the output bit was switched according to the truth table, which implemented various important logic gates[24,241,242], such as the universal NAND gate with a redox HfO_2 RSM[242] (FIG. 7g).

RSM-based cybersecurity

Cybersecurity applications rely on hardware roots of trust that make use of process-induced variations in electronic devices as a source of cryptographic data[29]. RSMs feature intrinsic stochasticity in electrical operations, which offers a low-cost and easily accessible solution to hardware identities and random number generators.

Hardware identities

A physical identifier is the 'fingerprint' of a piece of hardware, such as a chip. RSM-based hardware identities have been physically embedded with their cryptographic processes, which are unpredictable and practically inimitable[29].

Physical unclonable functions. Physical unclonable functions (PUFs) are black boxes that produce responses (for example, a 1-bit output) upon receiving challenges (such as a 9-bit input). Each RSM PUF has a unique and unpredictable (but invariant for each hardware) way of mapping challenges to responses using device-to-device variations in switching-pulse amplitude and duration[243–245], sneak-path currents[246] and nonlinear current–voltage (I–V) relationships[29,247]. For instance, using a 3D-stacked passive Al_2O_3/TiO_{2-x} RSM crossbar with inherent variations in the I–V nonlinearity of different devices, the challenge (for example, 111010011) determined which rows and columns were biased (or floated) (FIG. 8a). The electrical current difference between the left-half and the right-half columns produced the
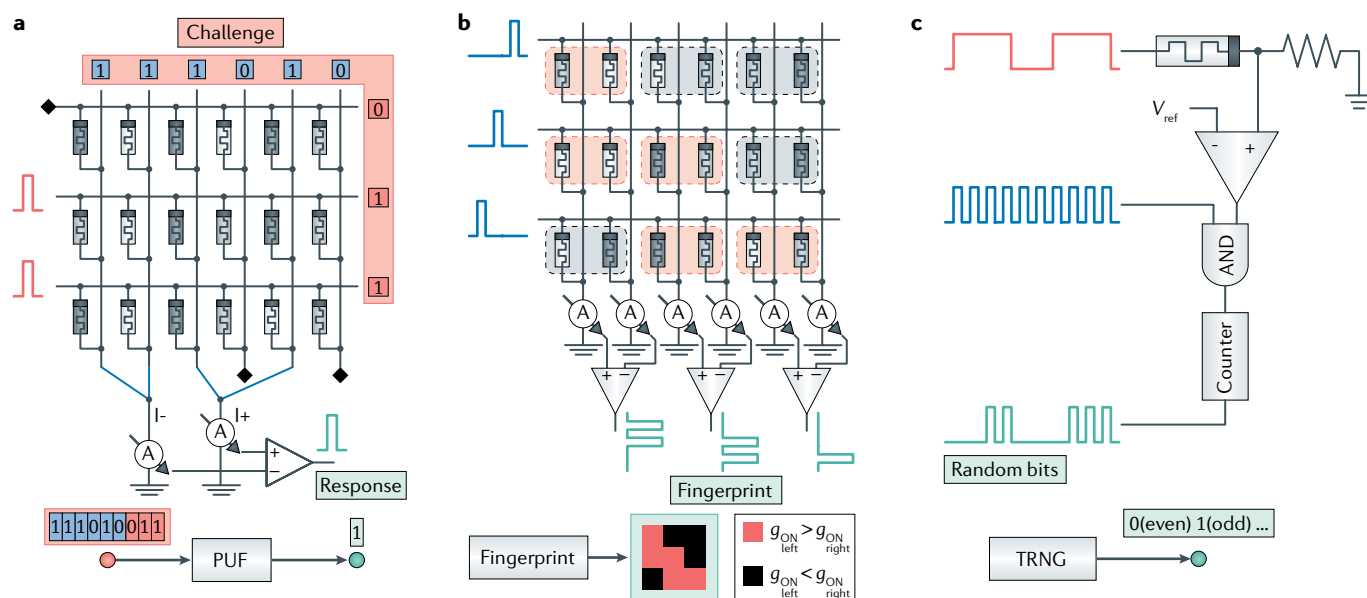
Fig. 8 | **RSM cybersecurity applications. a** | A resistive switching material (RSM) crossbar for a physical unclonable function (PUF). Using a 3D-stacked passive Al_2O_3/TiO_{2-x} RSM crossbar with inherent nonlinear variations in the current–voltage, the challenge (111010011) determines the rows that are biased or floated and the columns that are grounded or floated. The electrical current difference between the left-half and the right-half columns produces the response (1)[29]. **b** | An RSM crossbar fingerprint. By switching each cell of a HfO_x RSM crossbar to its ON state using identical programming parameters, the ON-state conductance of paired adjacent cells (red and black boxes) produces the fingerprint matrix, where the element of the fingerprint is red or black, depending on whether the left or right RSM cell, respectively, has a larger ON-state conductance (g_{ON})[30]. **c** | RSM-based true random number generator (TRNG). The diffusive RSM is switched ON by a voltage pulse (orange line) in each cycle. The variation in the delay in ON switching leads to the AND gate passing different numbers of clocks (blue line) to the counter in different cycles. The last bit of the counter output (green line) is the random bit of the cycle[252]. V_{ref}, reference voltage.

response (for example, 1)[29], which is repeatable each time the same challenge is presented to the same hardware and, thus, can be used to identify hardware.

Fingerprints. Unlike PUFs, which are black boxes with a large number of challenge–response pairs that cannot be enumerated, the intrinsic device-to-device variation of the ON-state conductance of redox-RSM crossbars could be an explicit hardware identifier or fingerprint. By switching each cell of a HfO_x RSM crossbar to its ON state using identical programming parameters, the ON-state conductance of paired adjacent cells (red and black boxes in FIG. 8b) produced a fingerprint matrix. The element of the fingerprint was red (black) if the left (right) RSM cell had a higher conductance[30].

Random number generators

Random numbers are used in most cryptographic applications, such as key generation. Random number generation could exploit the intrinsic stochasticity of redox, phase-change and magnetoresistive RSMs, including random telegraph noise[248], variations of the threshold voltages[249,250] or switching delays[251,252]. For instance, a random number generator based on a Ag-diffusive RSM was switched ON by a voltage pulse in each random bit generation cycle (orange line in FIG. 8c). The varied time delay in the ON switching enabled the AND gate to pass different numbers of clocks (blue line in FIG. 8c) to a counter in different cycles. The last bit of

the counter output (green line in FIG. 8c) was the random bit of the cycle[252].

Future perspectives

RSMs have shown their potential as the building blocks of unconventional computing architectures to circumvent the von Neumann bottleneck and the limits of Dennard scaling. However, to commercialize RSM-based computing technology, further application-specific research is needed on materials, devices, systems and algorithms.

Materials engineering

The switching and transport mechanisms of the RSMs are not completely understood and are, in many cases, controversial. In addition, the exploration of new materials and mechanisms may lead to improved electrical properties and, thus, computing performance.

Switching and transport mechanisms. Probing the operating mechanisms of redox and phase-change RSMs has been challenging because the switching involves fast and stochastic atomic rearrangements in a nanoscale volume. Quantitative rather than qualitative descriptions of switching and transport, which would enable elucidation of the electrical properties of material systems and formulation of compact device circuit models, is particularly lacking, especially for redox RSMs and domain-wall-transport ferroelectric RSMs.

High-spatiotemporal-resolution characterization of redox reactions[43], phase transitions[253] and domain evolutions[141] would likely provide crucial insight into the underlying mechanisms.

New materials and device structures. In addition to conventional RSMs, scaling RSMs to low dimensions may produce rich new physics to boost RSM performance. Such low-dimensional RSMs have recently been explored, including 1D nanowires[254] and 2D van der Waals RSMs (including redox[255,256], phase change[257], magnetoresistive[258] and ferroelectric[259]). For SOT MTJs, topological insulators with spin–momentum-locked surface states could give rise to a much higher SOC efficiency[260,261], even at room temperature[262–264].

Although resistive switching is usually embodied in two-terminal devices, three-terminal transistor-like devices, such as ferroelectric[265], carbon-nanotube[266] and ionic transistors[267], also feature switchable non-volatile channel conductance with projected in-memory computing capabilities. The third terminals of three-terminal devices may enable additional controls or computations to be performed beyond those available in the conventional crossbars of two-terminal devices, albeit at the expense of increased footprint. Moreover, resistance-based computing paradigms can be transplanted into other emerging circuit elements, inspiring concepts such as memcapacitors[268] and memtransistors[269].

CMOS compatibility and cost. The fabrication of RSM devices may involve steps that are difficult to integrate into CMOS-fabrication processes. For example, it is necessary to improve the CMOS compatibility of the inert-metal electrodes frequently used for redox RSMs. Scaling VCMA MTJs demands large VCMA coefficients (300–1,000 fJ V^{-1} m^{-1})[110], which relies on epitaxial deposition rather than industrially applicable sputtering processes. For SOT MTJs, the efficiency of charge-to-spin conversion is low in CMOS-compatible, heavy-metal-based devices[270], for which it is also challenging to generate perpendicular spin torque for directly switching a perpendicular MTJ. Therefore, an in-plane magnetic field is needed to deterministically switch the perpendicular MTJ. Field-free switching can be achieved through lateral symmetry breaking of structures[271], exchange bias[272,273], interlayer exchange coupling[274] or tilted magnetic anisotropy; however, practices for large-scale production of perpendicular MTJs with field-free switching need to be developed. For ferroelectric RSMs, polycrystalline doped-HfO_2-based and $(Hf,Zr)O_2$-based ferroelectrics that are compatible with atomic layer deposition and that have a comparable performance to that of epitaxial ferroelectrics (such as $BaTiO_3$ and $BiFeO_3$) are yet to be demonstrated.

Device optimization

Optimization of common device properties (such as information-representation capability, switching energy and speed, reliability and density) and application-specific properties (such as programming linearity and I–V linearity) could help to maximize the area and energy efficiencies of RSM-based computing.

Stochasticity and number of states. For redox and phase-change RSMs, the cycle-to-cycle and device-to-device variation in the switching process could be suppressed with increased confinement of the conducting channels[153,154] and homogenous switching (non-filamentary)[72,73]. For magnetoresistive RSMs, the theoretical ON/OFF ratio is >10,000%[275], but the highest ratio demonstrated at room temperature is only ~600% in in-plane MTJs[276] and ~250% in perpendicular MTJs[113], mainly owing to lattice imperfections. Techniques to suppress localized states induced by oxygen vacancies and grain boundaries[277,278] may help to increase the tunnel magnetoresistance ON/OFF ratio. For VCMA MTJs, the non-deterministic precessional switching could be compensated by special programming circuits[279,280] or other deterministic switching mechanisms[281]. Regularizing the domain-wall nucleation may improve the programming accuracy of ferroelectric RSMs. However, controllable stochasticity benefits certain neuromorphic and cybersecurity applications (discussed below).

Although it is possible to use a single RSM device to represent a computational parameter such as a synaptic weight, the conductance difference of a pair of RSM devices could reduce the post-processing involved in representing negative numbers[207] and increase the defect tolerance. Furthermore, multiple RSM devices could be grouped to represent a value of more bits[28,214,228] or a value of the same number of bits[88] of reduced programming error at the expense of the areal density of bits.

Switching energy and speed. The switching energy can be reduced by introducing a lower-energy potential barrier between RSM states, albeit at the expense of data stability. The small redox barriers and high ion mobilities of redox RSMs may yield low-energy switching. Low-energy phase changes could be achieved by sublithographic patterning or complex fabrication schemes to minimize the switching volumes[77,78]. The relatively large programming current of STT MTJs results in considerable ohmic dissipation and possible damage to the MgO barrier[282], which could be decreased by using a free layer with low damping and a small magnetic moment[163]. For VCMA MTJs and FTJs, the switching energy can be reduced by designing RSMs with smaller magnetic anisotropy and polarization energies, respectively.

In addition, all four types of RSM face the ==power–speed dilemma== (that is, fast switching is usually associated with a large driving power), which demands a balance between the power and speed.

Retention and endurance. Redox and phase-change RSMs have demonstrated data retention of >10 years but at the cost of the mobile species having a relatively large activation energy. By comparison, magnetoresistive and ferroelectric RSMs, which feature small switching-energy barriers, show comparatively poor retention. In addition, MTJs and FTJs, particularly STT RSMs, also suffer from retention degradation, owing

to reading perturbation, which could be alleviated by having smaller reading currents. This energy–retention dilemma might be overcome by borrowing the operating principles of a battery in a three-terminal ionic transistor[267,283,284] to decouple the programming and read operations.

In terms of endurance, redox, phase-change and magnetoresistive RSMs have all shown operations of >10^{11} cycles for discrete devices; such values may be sufficient for programming infrequent applications, such as those shown in FIG. 7b–e, but not for applications such as digital memcomputing (FIG. 7g). The endurance of ferroelectric RSMs is possibly limited by the charge-trapping-induced degradation[132].

Density. As redox and phase-change RSMs typically rely on the formation and rupture of localized conducting channels, the minimal device size is, in principle, limited by the dimensions of the channels. For MTJs and ferroelectric RSMs, a small dimension means a small energy barrier or poor data stability, which could be compensated by increasing the anisotropy energy and electric polarization density through materials engineering.

Programming linearity. For ANNs that are trained in situ by backpropagation, it is desirable to program RSMs in a linear and symmetric fashion, such that each electrical stimuli (such as the number or width of voltage pulses) induces the same amount of conductance change in both ON (potentiation) and OFF (depression) switching. Practically, there is usually nonlinearity and asymmetry in programming, although these could be reduced by waveform engineering[285]. In addition, work on pairing phase-change RSM devices with three-transistor–one-capacitor structures[23] and battery-like ionic transistors[267,283] could greatly improve programming linearity and symmetry upon application of trains of identical programming voltage spikes (see Note 1 in the Supplementary Information for further details). For ANNs trained ex situ and non-ANN applications, an efficient analogue programming method was developed to impose current compliance through series transistors[213] (see Note 1 in the Supplementary Information).

I–V nonlinearity. For crossbar-based applications, the sneak-path currents degrade the read voltage margin and incur a programming energy overhead, which could be circumvented by passive selectors of a large *I–V* nonlinearity. To date, diodes[286,287], tunnelling barriers[288], Ovonic electrical RSMs[289], Mott phase-change RSMs[290,291], mixed ionic–electronic conductors[292] and redox-diffusive RSMs[293,294] have been engineered for this purpose, although an ideal selector is yet to be demonstrated. Moreover, complementary resistive switches[295] may reduce errors caused by sneak-path currents in RSM-crossbar-based digital memcomputing[296]. Alternatively, active switches such as transistors could also be used to mitigate the sneak-path currents. Note that the introduction of a transistor effectively makes each RSM cell a three-terminal device by wiring a terminal of each RSM cell to the source or drain of the transistor (see Note 1 in the Supplementary Information).

System optimization

In a system-centric view, the overall performance of RSM-computing systems is also largely shaped by factors other than those of RSM devices, including the signal representation, the parasitic resistance and capacitance of the crossbars, the peripheral circuitry and the interconnections between crossbars.

Signal representation. Analogue voltages are possibly the most straightforward way of representing the signals of ANNs and analogue computing. Alternatively, the signals could also be encoded by the temporal width or number of voltage pulses of identical amplitude (see Supplementary Table 1). The latter approach has a larger time overhead as multiple temporal signals need to be sampled per computing operation but may avoid generating multiple bias levels within a narrow voltage range. More importantly, this approach could circumvent the requirement of ohmic or metallic transport of RSMs, which is usually available only in the relatively low-resistance regime, and thus makes it possible to use the high-resistance regime of RSMs to reduce energy consumption.

Crossbar parasitic resistance and capacitance. The scale of a single RSM crossbar is possibly limited by the wire resistance and the parasitic capacitance. The former may affect the accuracy of vector–matrix multiplications, whereas the latter may limit the maximum operating speed, which calls for mitigation techniques at the circuit level. For instance, an architecture consisting of multiple smaller crossbars with reconfigurable interconnections could possibly address this challenge. In addition, a low working conductance of RSMs can also minimize the effects of wire resistance, albeit at the cost of a larger RC constant.

Peripheral circuitry. The complexity of the peripheral circuitry depends on factors such as signal representation, computing precision and throughput. Different signal-generation and signal-acquisition schemes have been reported, with large variations in chip footprint and power. For instance, analogue voltage inputs could be sourced by digital-to-analogue converters, whereas the analogue currents could be sensed by transimpedance amplifiers and analogue-to-digital converters[17,209,212,213,217,226,229] (see Supplementary Table 1). An alternative is to use digital inputs and on-chip current-sense amplifiers, which has experimentally demonstrated an energy efficiency of >50 TOPS W^{-1} with redox-RSM arrays[211,227,228]. Therefore, the application-specific tailoring of the peripheral circuitry is crucial to optimize the overall energy and area efficiencies of RSM-based computing systems.

Reconfigurable fabric for multiple arrays. Real-world applications may require multiple RSM crossbars. Reconfigurable interconnections between RSM crossbars would be useful for flexible adaptation to a wide variety of RSM-network topologies[297].

Algorithm level

Besides hardware, the exploration of new applications or algorithms may better unleash the potential of RSMs.

For ANNs, inference-intensive and update-infrequent applications could better exploit the high-throughput and low-energy weighted sum of RSMs. Moreover, algorithms with relaxed requirements on the conductance precision may be more suitable for RSM-computing systems or even benefit from the imprecision, such as for harvesting stochasticity for natural regularization to avoid overfitting in ANNs[226], promoting policy finding in reinforcement learning[218] and representing parameters as probability distributions in Bayesian neural networks. In terms of ex situ-trained ANNs and memcomputing, there is a notable time and energy cost associated with the iterative programming of RSM crossbars; thus, there is a need for algorithms that balance the precision and speed. Moreover, for SNNs, bioinspired algorithms could intrinsically circumvent the stringent requirement on programming inaccuracy while supporting a high sparsity of signals. However, the practical application of bioinspired algorithms that can harness the strength of RSMs may demand interdisciplinary research and the co-optimization of algorithms and hardware, such as a hybrid ANN–SNN hardware platform[298].

For memcomputing, algorithms with iterative linear transformations that are robust to imprecision may be better suited to take advantage of the fast speed and low energy offered by analogue vector–matrix multiplications.

Conclusion

In addition to RSM-based memory (for example, Intel's Optane and Samsung's embedded magnetic random-access memory), prototypes of RSM-based computing systems developed by Panasonic[211] and the Taiwan Semiconductor Manufacturing Company (TSMC)[227,228] have already shown practical advantages in energy efficiency compared with state-of-the-art transistor-based hardware. Moreover, over the next few years, the efficiency may be further increased with scaled technology nodes and refined circuitry, possibly leading to an energy efficiency of >100 TOPS W^{-1} and increased area efficiency.

Nevertheless, whether these technologies can ultimately be commercialized, similar to how solid-state drives are replacing conventional hard disks, depends on how substantial the advantages are over alternative technologies and the effort invested by the community. In general, RSMs have exhibited even larger advantages over the alternative technologies in computing applications than in memory applications. However, the commercialization of RSM-based computing is still hindered by several challenges, particularly the energy consumption and large variability of redox and phase-change RSMs, the programming limitations (for example, high STT current, high VCMA error and low SOT efficiency) and the low ON/OFF ratio of magneto-resistive RSMs, and the poor reliability of ferroelectric RSMs. In addition to the directions of future research outlined above, including deeper understanding of the dynamic switching process at the atomic level and engineering materials systems for improved electrical properties, co-design at the circuit and algorithm levels may lead to innovative and effective solutions to mitigate device nonidealities. By combining these efforts, these transformative computing technologies may be realized in the coming decade.

Published online 13 January 2020

1. Moore, G. E. Cramming more components onto integrated circuits. *Proc. IEEE* **86**, 82–85 (1998).
2. Dennard, R. H., Gaensslen, F. H., Rideout, V. L., Bassous, E. & LeBlanc, A. R. Design of ion-implanted MOSFET's with very small physical dimensions. *IEEE J. Solid-State Circuits* **9**, 256–268 (1974).
3. Chi, P. et al. PRIME: a novel processing-in-memory architecture for neural network computation in ReRAM-based main memory. *Proc. 43rd Int. Symp. Computer Architecture* 27–39 (IEEE, 2016).
4. Pawlowski, J. T. Hybrid memory cube (HMC). *2011 IEEE Hot Chips 23 Symp. (HCS)* 1–24 https://doi.org/10.1109/HOTCHIPS.2011.7477494 (IEEE, 2011).
5. Waser, R. & Aono, M. Nanoionics-based resistive switching memories. *Nat. Mater.* **6**, 833–840 (2007).
6. Waser, R., Dittmann, R., Staikov, G. & Szot, K. Redox-based resistive switching memories - nanoionic mechanisms, prospects, and challenges. *Adv. Mater.* **21**, 2632–2663 (2009).
7. Wuttig, M. & Yamada, N. Phase-change materials for rewriteable data storage. *Nat. Mater.* **6**, 824–832 (2007).
8. Wong, H.-S. P. et al. Phase change memory. *Proc. IEEE* **98**, 2201–2227 (2010).
9. Zhang, W., Mazzarello, R., Wuttig, M. & Ma, E. Designing crystallization in phase-change materials for universal memory and neuro-inspired computing. *Nat. Rev. Mater.* **4**, 150–168 (2019).
10. Brataas, A., Kent, A. D. & Ohno, H. Current-induced torques in magnetic materials. *Nat. Mater.* **11**, 372–381 (2012).
11. Matsukura, F., Tokura, Y. & Ohno, H. Control of magnetism by electric fields. *Nat. Nanotechnol.* **10**, 209–220 (2015).
12. Garcia, V. & Bibes, M. Ferroelectric tunnel junctions for information storage and processing. *Nat. Commun.* **5**, 4289 (2014).
13. Martin, L. W. & Rappe, A. M. Thin-film ferroelectric materials and their applications. *Nat. Rev. Mater.* **2**, 16087 (2016).
14. Chua, L. Memristor-The missing circuit element. *IEEE Trans. Circuit Theory* **18**, 507–519 (1971).
15. Strukov, D. B., Snider, G. S., Stewart, D. R. & Williams, R. S. The missing memristor found. *Nature* **453**, 80–83 (2008).
16. Yang, J. J. et al. Memristive switching mechanism for metal/oxide/metal nanodevices. *Nat. Nanotechnol.* **3**, 429–433 (2008).
17. Prezioso, M. et al. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature* **521**, 61–64 (2015).
18. Burr, G. W. et al. Neuromorphic computing using non-volatile memory. *Adv. Phys. X* **2**, 89–124 (2016).
19. Ielmini, D. & Wong, H. S. P. In-memory computing with resistive switching devices. *Nat. Electron.* **1**, 333–343 (2018).
20. Tsai, H., Ambrogio, S., Narayanan, P., Shelby, R. M. & Burr, G. W. Recent progress in analog memory-based accelerators for deep learning. *J. Phys. D. Appl. Phys.* **51**, 283001 (2018).
21. Yu, S. Neuro-inspired computing with emerging nonvolatile memorys. *Proc. IEEE* **106**, 260–285 (2018).
22. Zidan, M. A., Strachan, J. P. & Lu, W. D. The future of electronics based on memristive systems. *Nat. Electron.* **1**, 22–29 (2018).
23. Ambrogio, S. et al. Equivalent-accuracy accelerated neural-network training using analogue memory. *Nature* **558**, 60–67 (2018).
24. Borghetti, J. et al. 'Memristive' switches enable 'stateful' logic operations via material implication. *Nature* **464**, 873–876 (2010).
25. Sheridan, P. M. et al. Sparse coding with memristor networks. *Nat. Nanotechnol.* **12**, 784–789 (2017).
   The first large-scale redox-RSM crossbar implementation of the locally competitive algorithm for sparse coding.
26. Li, C. et al. Analogue signal and image processing with large memristor crossbars. *Nat. Electron.* **1**, 52–59 (2018).
27. Le Gallo, M. et al. Mixed-precision in-memory computing. *Nat. Electron.* **1**, 246–253 (2018).
   The first large-scale phase-change RSM-based linear system solver.
28. Zidan, M. A. et al. A general memristor-based partial differential equation solver. *Nat. Electron.* **1**, 411–420 (2018).
29. Nili, H. et al. Hardware-intrinsic security primitives enabled by analogue state and nonlinear conductance variations in integrated memristors. *Nat. Electron.* **1**, 197–202 (2018).
   The first large-scale redox-RSM-crossbar-based PUF using variations in I–V nonlinearity.
30. Jiang, H. et al. A provable key destruction scheme based on memristive crossbar arrays. *Nat. Electron.* **1**, 548–554 (2018).
31. Hickmott, T. W. Low-frequency negative resistance in thin anodic oxide films. *J. Appl. Phys.* **33**, 2669–2682 (1962).
32. Sawa, A. Resistive switching in transition metal oxides. *Mater. Today* **11**, 28–36 (2008).
33. Akinaga, H. & Shima, H. Resistive random access memory (ReRAM) based on metal oxides. *Proc. IEEE* **98**, 2237–2251 (2010).
34. Valov, I., Waser, R., Jameson, J. R. & Kozicki, M. N. Electrochemical metallization memories—fundamentals, applications, prospects. *Nanotechnology* **22**, 254003 (2011).
35. Jeong, D. S. et al. Emerging memories: resistive switching mechanisms and current status. *Rep. Prog. Phys.* **75**, 076502 (2012).
36. Wong, H. S. P. et al. Metal-oxide RRAM. *Proc. IEEE* **100**, 1951–1970 (2012).
37. Yang, J. J., Strukov, D. B. & Stewart, D. R. Memristive devices for computing. *Nat. Nanotechnol.* **8**, 13–24 (2013).

38. Goldfarb, I. et al. Electronic structure and transport measurements of amorphous transition-metal oxides: observation of Fermi glass behavior. *Appl. Phys. A* **107**, 1–11 (2012).

39. Szot, K., Speier, W., Bihlmayer, G. & Waser, R. Switching the electrical resistance of individual dislocations in single-crystalline SrTiO₃. *Nat. Mater.* **5**, 312–320 (2006).

40. Kwon, D.-H. et al. Atomic structure of conducting nanofilaments in TiO₂ resistive switching memory. *Nat. Nanotechnol.* **5**, 148–153 (2010).

41. Yang, J. J. et al. The mechanism of electroforming of metal oxide memristive switches. *Nanotechnology* **20**, 215201 (2009).

42. Strachan, J. P. et al. Direct identification of the conducting channels in a functioning memristive device. *Adv. Mater.* **22**, 3573–3577 (2010).

43. Baeumer, C. et al. Quantifying redox-induced Schottky barrier variations in memristive devices via in operando spectromicroscopy with graphene electrodes. *Nat. Commun.* **7**, 12398 (2016).

44. Nallagatla, V. R. et al. Topotactic phase transition driving memristive behavior. *Adv. Mater.* **31**, 1903191 (2019).

45. Kumar, S. et al. Conduction channel formation and dissolution due to oxygen thermophoresis/diffusion in hafnium oxide memristors. *ACS Nano* **10**, 11205–11210 (2016).

46. Kumar, S. et al. Direct observation of localized radial oxygen migration in functioning tantalum oxide memristors. *Adv. Mater.* **28**, 2772–2776 (2016).

47. Li, C. et al. Direct observations of nanofilament evolution in switching processes in HfO₂-based resistive random access memory by in situ TEM studies. *Adv. Mater.* **29**, 1602976 (2017).

48. Yang, Y. et al. Probing nanoscale oxygen ion motion in memristive systems. *Nat. Commun.* **8**, 15173 (2017).

49. Cooper, D. et al. Anomalous resistance hysteresis in oxide ReRAM: oxygen evolution and reincorporation revealed by in situ TEM. *Adv. Mater.* **29**, 1700212 (2017).

50. Du, H. et al. Nanosized conducting filaments formed by atomic-scale defects in redox-based resistive switching memories. *Chem. Mater.* **29**, 3164–3173 (2017).

51. Park, G.-S. et al. In situ observation of filamentary conducting channels in an asymmetric Ta₂O₅₋ₓ/TaO₂₋ₓ bilayer structure. *Nat. Commun.* **4**, 2382 (2013).

52. Miao, F. et al. Anatomy of a nanoscale conduction channel reveals the mechanism of a high-performance memristor. *Adv. Mater.* **23**, 5633–5640 (2011).

53. Yang, Y. et al. Electrochemical dynamics of nanoscale metallic inclusions in dielectrics. *Nat. Commun.* **5**, 4232 (2014).

54. Hirose, Y. & Hirose, H. Polarity-dependent memory switching and behavior of Ag dendrite in Ag-photodoped amorphous As₂S₃ films. *J. Appl. Phys.* **47**, 2767–2772 (1976).

55. Guo, X., Schindler, C., Menzel, S. & Waser, R. Understanding the switching-off mechanism in Ag⁺ migration based resistively switching model systems. *Appl. Phys. Lett.* **91**, 133513 (2007).

56. Wang, Z. et al. Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nat. Mater.* **16**, 101–108 (2017).

57. Xu, Z., Bando, Y., Wang, W., Bai, X. & Golberg, D. Real-time in situ HRTEM-resolved resistance switching of Ag₂S nanoscale ionic conductor. *ACS Nano* **4**, 2515–2522 (2010).

58. Liu, Q. et al. Real-time observation on dynamic growth/dissolution of conductive filaments in oxide-electrolyte-based ReRAM. *Adv. Mater.* **24**, 1844–1849 (2012).

59. Yang, Y. et al. Observation of conducting filament growth in nanoscale resistive memories. *Nat. Commun.* **3**, 732 (2012).

60. Valov, I. et al. Atomically controlled electrochemical nucleation at superionic solid electrolyte surfaces. *Nat. Mater.* **11**, 530–535 (2012).

61. Hubbard, W. A. et al. Nanofilament formation and regeneration during Cu/Al₂O₃ resistive memory switching. *Nano Lett.* **15**, 3983–3987 (2015).

62. Yuan, F. et al. Real-time observation of the electrode-size-dependent evolution dynamics of the conducting filaments in a SiO₂ layer. *ACS Nano* **11**, 4097–4104 (2017).

63. Wang, W. et al. Surface diffusion-limited lifetime of silver and copper nanofilaments in resistive switching devices. *Nat. Commun.* **10**, 81 (2019).

64. Valov, I. et al. Nanobatteries in redox-based resistive switches require extension of memristor theory. *Nat. Commun.* **4**, 1771 (2013).

65. Onofrio, N., Guzman, D. & Strachan, A. Atomic origin of ultrafast resistance switching in nanoscale electrometallization cells. *Nat. Mater.* **14**, 440–446 (2015).

66. Tian, X. et al. Bipolar electrochemical mechanism for mass transfer in nanoionic resistive memories. *Adv. Mater.* **26**, 3649–3654 (2014).

67. Chae, B. G. et al. Nanometer-scale phase transformation determines threshold and memory switching mechanism. *Adv. Mater.* **29**, 1701752 (2017).

68. Wedig, A. et al. Nanoscale cation motion in TaOₓ, HfOₓ and TiOₓ memristive systems. *Nat. Nanotechnol.* **11**, 67–74 (2016).

69. Sawa, A., Fujii, T., Kawasaki, M. & Tokura, Y. Interface resistance switching at a few nanometer thick perovskite manganite active layers. *Appl. Phys. Lett.* **88**, 232112 (2006).

70. Kim, K. M. et al. A detailed understanding of the electronic bipolar resistance switching behavior in Pt/TiO₂/Pt structure. *Nanotechnology* **22**, 254010 (2011).

71. Baikalov, A. et al. Field-driven hysteretic and reversible resistive switch at the Ag–Pr₀.₇Ca₀.₃MnO₃ interface. *Appl. Phys. Lett.* **83**, 957–959 (2003).

72. Herpers, A. et al. Spectroscopic proof of the correlation between redox-state and charge-carrier transport at the interface of resistively switching Ti/PCMO devices. *Adv. Mater.* **26**, 2730–2735 (2014).

73. Baek, K. et al. In situ TEM observation on the interface-type resistive switching by electrochemical redox reactions at a TiN/PCMO interface. *Nanoscale* **9**, 582–593 (2017).

74. Wang, Y. et al. Mott-transition-based RRAM. *Mater. Today* **28**, 63–80 (2019).

75. Mott, N. F. & Davis, E. A. *Electronic Processes in Non-Crystalline Materials* 2nd edn (Clarendon Press, 2012).

76. Szot, K., Bihlmayer, G. & Speier, W. Nature of the resistive switching phenomena in TiO₂ and SrTiO₃: origin of the reversible insulator–metal transition. *Solid State Phys.* **65**, 353–559 (2014).

77. Raoux, S. et al. Phase-change random access memory: a scalable technology. *IBM J. Res. Dev.* **52**, 465–479 (2008).

78. Burr, G. W. et al. Phase change memory technology. *J. Vac. Sci. Technol. B* **28**, 223–262 (2010).

79. Burr, G. W. et al. Recent progress in phase-change memory technology. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **6**, 146–162 (2016).

80. Waldecker, L. et al. Time-domain separation of optical properties from structural transitions in resonantly bonded materials. *Nat. Mater.* **14**, 991–995 (2015).

81. Le Gallo, M., Krebs, D., Zipoli, F., Salinga, M. & Sebastian, A. Collective structural relaxation in phase-change memory devices. *Adv. Electron. Mater.* **4**, 1700627 (2018).

82. Sebastian, A., Le Gallo, M. & Krebs, D. Crystal growth within a phase change memory cell. *Nat. Commun.* **5**, 4314 (2014).

83. Salinga, M. et al. Measurement of crystal growth velocity in a melt-quenched phase-change material. *Nat. Commun.* **4**, 2371 (2013).

84. Nirschl, T. et al. Write strategies for 2 and 4-bit multi-level phase-change memory. *2007 IEEE Int. Electron Devices Meeting* 461–464 (IEEE, 2007).

85. Papandreou, N. et al. Programming algorithms for multilevel phase-change memory. *2011 IEEE Int. Symp. Circuits and Systems (ISCAS)* 329-332 (IEEE, 2011).

86. Burr, G. W. et al. 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* **62**, 3498–3507 (2015).
The first in situ training of a large-scale, fully connected ANN based on a phase-change RSM for MNIST classification.

87. Sebastian, A. et al. Temporal correlation detection using computational phase-change memory. *Nat. Commun.* **8**, 1115 (2017).

88. Boybat, I. et al. Neuromorphic computing with multi-memristive synapses. *Nat. Commun.* **9**, 2514 (2018).

89. Siegrist, T. et al. Disorder-induced localization in crystalline phase-change materials. *Nat. Mater.* **10**, 202–208 (2011).

90. Zhang, W. et al. Role of vacancies in metal-insulator transitions of crystalline phase-change materials. *Nat. Mater.* **11**, 952–956 (2012).

91. Zhang, W. et al. Density-functional theory guided advances in phase-change materials and memories. *MRS Bull.* **40**, 856–869 (2015).

92. Nardone, M., Simon, M., Karpov, I. V. & Karpov, V. G. Electrical conduction in chalcogenide glasses of phase change memory. *J. Appl. Phys.* **112**, 071101 (2012).

93. Wuttig, M., Deringer, V. L., Gonze, X., Bichara, C. & Raty, J. Y. Incipient metals: functional materials with a unique bonding mechanism. *Adv. Mater.* **30**, 1803777 (2018).

94. Chappert, C., Fert, A. & Van Dau, F. N. The emergence of spin electronics in data storage. *Nat. Mater.* **6**, 813–823 (2007).

95. Miyazaki, T. & Tezuka, N. Giant magnetic tunneling effect in Fe/Al₂O₃/Fe junction. *J. Magn. Magn. Mater.* **139**, L231–L234 (1995).

96. Moodera, J. S., Kinder, L. R., Wong, T. M. & Meservey, R. Large magnetoresistance at room temperature in ferromagnetic thin film tunnel junctions. *Phys. Rev. Lett.* **74**, 3273–3276 (1995).

97. Baibich, M. N. et al. Giant magnetoresistance of (001) Fe/(001) Cr magnetic superlattices. *Phys. Rev. Lett.* **61**, 2472–2475 (1988).

98. Binasch, G., Grünberg, P., Saurenbach, F. & Zinn, W. Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange. *Phys. Rev. B* **39**, 4828–4830 (1989).

99. Parkin, S. S. P. et al. Giant tunnelling magnetoresistance at room temperature with MgO (100) tunnel barriers. *Nat. Mater.* **3**, 862–867 (2004).

100. Yuasa, S., Nagahama, T., Fukushima, A., Suzuki, Y. & Ando, K. Giant room-temperature magnetoresistance in single-crystal Fe/MgO/Fe magnetic tunnel junctions. *Nat. Mater.* **3**, 868–871 (2004).

101. Ikeda, S. et al. A perpendicular-anisotropy CoFeB–MgO magnetic tunnel junction. *Nat. Mater.* **9**, 721–724 (2010).

102. Berger, L. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B* **54**, 9353–9358 (1996).

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

104. Golonzka, O. et al. MRAM as embedded non-volatile memory solution for 22FFL FinFET technology. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 18-1 (2018).

105. Lee, K. et al. 22-nm FD-SOI embedded MRAM technology for low-power automotive-grade-I MCU applications. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 27-1 (2018).

106. Song, Y. J. et al. Demonstration of highly manufacturable STT-MRAM embedded in 28nm logic. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 18-2 (2018).

107. Duan, C.-G. et al. Surface magnetoelectric effect in ferromagnetic metal films. *Phys. Rev. Lett.* **101**, 137201 (2008).

108. Grezes, C. et al. Ultra-low switching energy and scaling in electric-field-controlled nanoscale magnetic tunnel junctions with high resistance-area product. *Appl. Phys. Lett.* **108**, 012403 (2016).

109. Wang, W.-G., Li, M., Hageman, S. & Chien, C. L. Electric-field-assisted switching in magnetic tunnel junctions. *Nat. Mater.* **11**, 64–68 (2011).

110. Nozaki, T. et al. Large voltage-induced changes in the perpendicular magnetic anisotropy of an MgO-based tunnel junction with an ultrathin Fe layer. *Phys. Rev. Appl.* **5**, 044006 (2016).

111. Li, X. et al. Enhancement of voltage-controlled magnetic anisotropy through precise control of Mg insertion thickness at CoFeB|MgO interface. *Appl. Phys. Lett.* **110**, 052401 (2017).

112. Liu, T., Zhang, Y., Cai, J. W. & Pan, H. Y. Thermally robust Mo/CoFeB/MgO trilayers with strong perpendicular magnetic anisotropy. *Sci. Rep.* **4**, 5895 (2014).

113. Wang, M. et al. Current-induced magnetization switching in atom-thick tungsten engineered perpendicular magnetic tunnel junctions with large tunnel magnetoresistance. *Nat. Commun.* **9**, 671 (2018).

114. Manchon, A., Koo, H. C., Nitta, J., Frolov, S. M. & Duine, R. A. New perspectives for Rashba spin–orbit coupling. *Nat. Mater.* **14**, 871–882 (2015).

115. Mihai Miron, I. et al. Current-driven spin torque induced by the Rashba effect in a ferromagnetic metal layer. *Nat. Mater.* **9**, 230–234 (2010).

116. Liu, L. et al. Spin-torque switching with the giant spin Hall effect of tantalum. *Science* **336**, 555–558 (2012).

117. Julliere, M. Tunneling between ferromagnetic films. *Phys. Lett. A* **54**, 225–226 (1975).

118. Slonczewski, J. C. Conductance and exchange coupling of two ferromagnets separated by a tunneling barrier. *Phys. Rev. B* **39**, 6995–7002 (1989).

119. Hoffmann, M. et al. Unveiling the double-well energy landscape in a ferroelectric layer. *Nature* **565**, 464–467 (2019).
120. Scott, J. F. *Ferroelectric Memories* Vol. 3 (Springer, 2013).
121. McAdams, H. P. et al. A 64-Mb embedded FRAM utilizing a 130-nm 5LM Cu/FSG logic process. *IEEE J. Solid-State Circuits* **39**, 667–677 (2004).
122. Böscke, T. S., Müller, J., Bräuhaus, D., Schröder, U. & Böttger, U. Ferroelectricity in hafnium oxide thin films. *Appl. Phys. Lett.* **99**, 102903 (2011).
123. Müller, J. et al. Ferroelectricity in yttrium-doped hafnium oxide. *J. Appl. Phys.* **110**, 114113 (2011).
124. Mueller, S. et al. Incipient ferroelectricity in Al-doped HfO_2 thin films. *Adv. Funct. Mater.* **22**, 2412–2417 (2012).
125. Park, M. H. et al. Study on the degradation mechanism of the ferroelectric properties of thin $Hf_{0.5}Zr_{0.5}O_2$ films on TiN and Ir electrodes. *Appl. Phys. Lett.* **105**, 072902 (2014).
126. Esaki, A. L., Laibowitz, R. B. & Stiles, P. J. Polar switch. *IBM Tech. Discl. Bull.* **13**, 114–116 (1971).
127. Seidel, J. et al. Conduction at domain walls in oxide multiferroics. *Nat. Mater.* **8**, 229–234 (2009).
128. Boyn, S. et al. Learning through ferroelectric domain dynamics in solid-state synapses. *Nat. Commun.* **8**, 14736 (2017).
129. Rojac, T. et al. Domain-wall conduction in ferroelectric $BiFeO_3$ controlled by accumulation of charged defects. *Nat. Mater.* **16**, 322–327 (2017).
130. Sharma, P. et al. Nonvolatile ferroelectric domain wall memory. *Sci. Adv.* **3**, e1700512 (2017).
131. Ma, J. et al. Controllable conductive readout in self-assembled, topologically confined ferroelectric domain walls. *Nat. Nanotechnol.* **13**, 947–952 (2018).
132. Jiang, J. et al. Temporary formation of highly conducting domain walls for non-destructive read-out of ferroelectric domain-wall resistance switching memories. *Nat. Mater.* **17**, 49–56 (2018).
133. Catalan, G., Seidel, J., Ramesh, R. & Scott, J. F. Domain wall nanoelectronics. *Rev. Mod. Phys.* **84**, 119–156 (2012).
134. Stengel, M., Vanderbilt, D. & Spaldin, N. A. Enhancement of ferroelectricity at metal-oxide interfaces. *Nat. Mater.* **8**, 392–397 (2009).
135. Farokhipoor, S. & Noheda, B. Conduction through 71° domain walls in $BiFeO_3$ thin films. *Phys. Rev. Lett.* **107**, 127601 (2011).
136. Maksymovych, P. et al. Dynamic conductivity of ferroelectric domain walls in $BiFeO_3$. *Nano Lett.* **11**, 1906–1912 (2011).
137. Palai, R. et al. β phase and γ–β metal-insulator transition in multiferroic $BiFeO_3$. *Phys. Rev. B* **77**, 014110 (2008).
138. Jiang, A.-Q., Lee, H. J., Hwang, C. S. & Tang, T.-A. Resolving the Landauer paradox in ferroelectric switching by high-field charge injection. *Phys. Rev. B* **80**, 024119 (2009).
139. Ducharme, S. et al. Intrinsic ferroelectric coercive field. *Phys. Rev. Lett.* **84**, 175–178 (2000).
140. Jia, C. L. et al. Atomic-scale study of electric dipoles near charged and uncharged domain walls in ferroelectric films. *Nat. Mater.* **7**, 57–61 (2008).
141. Jia, C.-L., Urban, K. W., Alexe, M., Hesse, D. & Vrejoiu, I. Direct observation of continuous electric dipole rotation in flux-closure domains in ferroelectric $Pb(Zr,Ti)O_3$. *Science* **331**, 1420–1423 (2011).
142. Nelson, C. T. et al. Spontaneous vortex nanodomain arrays at ferroelectric heterointerfaces. *Nano Lett.* **11**, 828–834 (2011).
143. Chanthbouala, A. et al. Solid-state memories based on ferroelectric tunnel junctions. *Nat. Nanotechnol.* **7**, 101–104 (2011).
144. Chanthbouala, A. et al. A ferroelectric memristor. *Nat. Mater.* **11**, 860–864 (2012).
145. Yoong, H. Y. et al. Epitaxial ferroelectric $Hf_{0.5}Zr_{0.5}O_2$ thin films and their implementations in memristors for brain-inspired computing. *Adv. Funct. Mater.* **28**, 1806037 (2018).
146. Chen, L. et al. Ultra-low power $Hf_{0.5}Zr_{0.5}O_2$ based ferroelectric tunnel junction synapses for hardware neural network applications. *Nanoscale* **10**, 15826–15833 (2018).
147. Liu, S., Grinberg, I. & Rappe, A. M. Intrinsic ferroelectric switching from first principles. *Nature* **534**, 360–363 (2016).
148. Meyer, R., Contreras, J. R., Petraru, A. & Kohlstedt, H. On a novel ferro resistive random access memory (FRRAM): basic model and first experiments. *Integr. Ferroelectr.* **64**, 77–88 (2004).
149. Black, C. T. & Welser, J. J. Electric-field penetration into metals: consequences for high-dielectric-constant capacitors. *IEEE Trans. Electron Devices* **46**, 776–780 (1999).
150. Wen, Z., Li, C., Wu, D., Li, A. & Ming, N. Ferroelectric-field-effect-enhanced electroresistance in metal/ferroelectric/semiconductor tunnel junctions. *Nat. Mater.* **12**, 617–621 (2013).
151. Yamada, H. et al. Ferroelectric control of a Mott insulator. *Sci. Rep.* **3**, 2834 (2013).
152. Kohlstedt, H., Pertsev, N., Contreras, J. R. & Waser, R. Theoretical current-voltage characteristics of ferroelectric tunnel junctions. *Phys. Rev. B* **72**, 125341 (2005).
153. Choi, S. et al. SiGe epitaxial memory for neuromorphic computing with reproducible high performance based on engineered dislocations. *Nat. Mater.* **17**, 335–340 (2018).
154. Ding, K. et al. Phase-change heterostructure enables ultralow noise and drift for memory operation. *Science* **366**, 210–215 (2019).
155. Shiota, Y. et al. Reduction in write error rate of voltage-driven dynamic magnetization switching by improving thermal stability factor. *Appl. Phys. Lett.* **111**, 022408 (2017).
156. Cheng, C.-H., Tsai, C., Chin, A. & Yeh, F. High performance ultra-low energy RRAM with good retention and endurance. *2010 Int. Electron Devices Meeting* 19-4 (IEEE, 2010).
157. Strachan, J. P., Torrezan, A. C., Medeiros-Ribeiro, G. & Williams, R. S. Measuring the switching dynamics and energy efficiency of tantalum oxide memristors. *Nanotechnology* **22**, 505402 (2011).
158. Xiong, F., Liao, A. D., Estrada, D. & Pop, E. Low-power switching of phase-change materials with carbon nanotube electrodes. *Science* **332**, 568–570 (2011).
159. Liang, J., Jeyasingh, R. G. D., Chen, H.-Y. & Wong, H.-S. P. A 1.4 µA reset current phase change memory cell with integrated carbon nanotube electrodes for cross-point memory application. *2011 Symp. VLSI Technology-Digest of Technical Papers* 100–101 (IEEE, 2011).
160. Choi, B. J. et al. High-speed and low-energy nitride memristors. *Adv. Funct. Mater.* **26**, 5290–5296 (2016).
161. Rao, F. et al. Reducing the stochasticity of crystal nucleation to enable subnanosecond memory writing. *Science* **358**, 1423–1427 (2017).
162. Loke, D. et al. Breaking the speed limits of phase-change memory. *Science* **336**, 1566–1569 (2012).
163. Thomas, L. et al. STT-MRAM devices with low damping and moment optimized for LLC applications at 0x nodes. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 27.3.1–27.3.4 (IEEE, 2018).
164. Zhao, H. et al. Sub-200 ps spin transfer torque switching in in-plane magnetic tunnel junctions with interface perpendicular anisotropy. *J. Phys. D* **45**, 025001 (2011).
165. Jiang, H. et al. Sub-10 nm Ta channel responsible for superior performance of a HfO_2 memristor. *Sci. Rep.* **6**, 28525 (2016).
166. Navarro, G. et al. Trade-off between SET and data retention performance thanks to innovative materials for phase-change memory. *2013 IEEE Int. Electron Devices Meeting* 21-5 (IEEE, 2013).
167. Lee, Y. K. et al. Embedded STT-MRAM in 28-nm FDSOI logic process for industrial MCU/IoT application. *2018 IEEE Symp. VLSI Technology* 181–182 (2018).
168. Yamada, H. et al. Giant electroresistance of super-tetragonal $BiFeO_3$-based ferroelectric tunnel junctions. *ACS Nano* **7**, 5385–5390 (2013).
169. Chen, B. et al. Physical mechanisms of endurance degradation in TMO-RRAM. *2011 Int. Electron Devices Meeting* 12-3 (IEEE, 2011).
170. Lee, M. J. et al. A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta_2O_{5-x}/TaO_{2-x} bilayer structures. *Nat. Mater.* **10**, 625–630 (2011).
171. Padilla, A. et al. Voltage polarity effects in $Ge_2Sb_2Te_5$-based phase change memory devices. *J. Appl. Phys.* **110**, 054501 (2011).
172. Raoux, S. et al. Phase change materials and their application to random access memory technology. *Microelectron. Eng.* **85**, 2330–2333 (2008).
173. Pedersen, T. P. L. et al. Mechanical stresses upon crystallization in phase change materials. *Appl. Phys. Lett.* **79**, 3597–3599 (2001).
174. Kim, I. et al. High performance PRAM cell scalable to sub-20nm technology with below $4F^2$ cell size, extendable to DRAM applications. *2010 Symp. VLSI Technology* 203–204 (IEEE, 2010).
175. Sato, H. et al. 14ns write speed 128Mb density embedded STT-MRAM with endurance>10^{10} and 10yrs retention@85°C using novel low damage MTJ integration process. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 27-2 (2018).
176. Shiokawa, Y. et al. High write endurance up to 10^{12} cycles in a spin current-type magnetic memory array. *AIP Adv.* **9**, 035236 (2019).
177. Boyn, S. et al. High-performance ferroelectric memory based on fully patterned tunnel junctions. *Appl. Phys. Lett.* **104**, 052909 (2014).
178. Govoreanu, B. et al. $10×10nm^2$ Hf/HfO_x crossbar resistive RAM with excellent performance, reliability and low-energy operation. *2011 Int. Electron Devices Meeting* 31-6 (IEEE, 2011).
179. Pi, S. et al. Memristor crossbar arrays with 6-nm half-pitch and 2-nm critical dimension. *Nat. Nanotechnol.* **14**, 35–39 (2019).
180. Golonzka, O. et al. Non-volatile RRAM embedded into 22FFL FinFET technology. *2019 Symp. VLSI Technology* T230–T231 (2019).
181. Choi, Y. et al. A 20nm 1.8 V 8Gb PRAM with 40MB/s program bandwidth. *2012 IEEE Int. Solid-State Circuits Conf.* 46–48 (IEEE, 2012).
182. Luo, Q. et al. 8-Layers 3D vertical RRAM with excellent scalability towards storage class memory applications. *2017 IEEE Int. Electron Devices Meeting (IEDM)* 2–7 (IEEE, 2017).
183. Yoon, K. J., Kim, Y. & Hwang, C. S. What will come after V-NAND—vertical resistive switching memory? *Adv. Electron. Mater.* **5**, 1800914 (2019).
184. Park, C. et al. Low RA magnetic tunnel junction arrays in conjunction with low switching current and high breakdown voltage for STT-MRAM at 10 nm and beyond. *2018 IEEE Symp. VLSI Technology* 185–186 (2018).
185. Sakhare, S. et al. Enablement of STT-MRAM as last level cache for the high performance computing domain at the 5nm node. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 18-3 (2018).
186. Gao, X. S., Liu, J. M., Au, K. & Dai, J. Y. Nanoscale ferroelectric tunnel junctions based on ultrathin $BaTiO_3$ film and Ag nanoelectrodes. *Appl. Phys. Lett.* **101**, 142905 (2012).
187. Jo, S. H. et al. Nanoscale memristor device as synapse in neuromorphic systems. *Nano Lett.* **10**, 1297–1301 (2010).
188. Ohno, T. et al. Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. *Nat. Mater.* **10**, 591–595 (2011).
189. 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. Elect. Dev.* **58**, 2729–2737 (2011).
190. 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.* **12**, 2179–2186 (2012).
191. Kim, S. et al. Experimental demonstration of a second-order memristor and its ability to biorealistically implement synaptic plasticity. *Nano Lett.* **15**, 2203–2211 (2015).
192. Vincent, A. F. et al. Spin-transfer torque magnetic memory as a stochastic memristive synapse for neuromorphic systems. *IEEE Trans. Biomed. Circuits Syst.* **9**, 166–174 (2015).
193. 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.* **25**, 4290–4299 (2015).
194. Suri, M. et al. Phase change memory as synapse for ultra-dense neuromorphic systems: application to complex visual pattern extraction. *2011 IEEE Int. Electron Devices Meeting* 4-4 (IEEE, 2011).
195. Srinivasan, G., Sengupta, A. & Roy, K. Magnetic tunnel junction based long-term short-term stochastic synapse for a spiking neural network with on-chip STDP learning. *Sci. Rep.* **6**, 29545 (2016).
196. Tuma, T., Pantazi, A., Le Gallo, M., Sebastian, A. & Eleftheriou, E. Stochastic phase-change neurons. *Nat. Nanotechnol.* **11**, 693–699 (2016).
197. Lee, D. et al. Various threshold switching devices for integrate and fire neuron applications. *Adv. Electron. Mater.* **5**, 1800866 (2019).
198. Cai, J. et al. Voltage-controlled spintronic stochastic neuron based on a magnetic tunnel junction. *Phys. Rev. Appl.* **11**, 034015 (2019).
199. Wu, M.-H. et al. Extremely compact integrate-and-fire STT-MRAM neuron: a pathway toward all-spin artificial deep neural network. *2019 Symp. VLSI Technology* T34-T35 (IEEE, 2019).
200. Wang, Z. et al. Fully memristive neural networks for pattern classification with unsupervised learning. *Nat. Electron.* **1**, 137–145 (2018).

201. Pickett, M. D., Medeiros-Ribeiro, G. & Williams, R. S. A scalable neuristor built with Mott memristors. *Nat. Mater.* **12**, 114–117 (2013).

202. Kumar, S., Strachan, J. P. & Williams, R. S. Chaotic dynamics in nanoscale NbO₂ Mott memristors for analogue computing. *Nature* **548**, 318–321 (2017).

203. Lim, H. et al. Relaxation oscillator-realized artificial electronic neurons, their responses, and noise. *Nanoscale* **8**, 9629–9640 (2016).

204. Torrejon, J. et al. Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017).

205. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).

206. Hu, M. et al. Memristor-based analog computation and neural network classification with a dot product engine. *Adv. Mater.* **30**, 1705914 (2018).

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

208. Park, S. et al. Electronic system with memristive synapses for pattern recognition. *Sci. Rep.* **5**, 10123 (2015).

209. Bayat, F. M. et al. Implementation of multilayer perceptron network with highly uniform passive memristive crossbar circuits. *Nat. Commun.* **9**, 2331 (2018).

210. Yao, P. et al. Face classification using electronic synapses. *Nat. Commun.* **8**, 15199 (2017).

211. Mochida, R. et al. A 4M synapses integrated analog ReRAM based 66.5 TOPS/W neural-network processor with cell current controlled writing and flexible network architecture. *2018 IEEE Symp. VLSI Technology* 175-176 (IEEE, 2018).

212. Cai, F. et al. A fully integrated reprogrammable memristor–CMOS system for efficient multiply–accumulate operations. *Nat. Electron.* **2**, 290–299 (2019).

213. Li, C. et al. Efficient and self-adaptive in-situ learning in multilayer memristor neural networks. *Nat. Commun.* **9**, 2385 (2018).

214. Yu, S. et al. Binary neural network with 16 Mb RRAM macro chip for classification and online training. *2016 IEEE Int. Electron Devices Meeting (IEDM)* 16-2 (IEEE, 2016).

215. Sanger, T. D. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Netw.* **2**, 459–473 (1989).

216. Choi, S., Shin, J. H., Lee, J., Sheridan, P. & Lu, W. D. Experimental demonstration of feature extraction and dimensionality reduction using memristor networks. *Nano Lett.* **17**, 3113–3118 (2017).

217. Wang, Z. et al. Reinforcement learning with analogue memristor arrays. *Nat. Electron.* **2**, 115–124 (2019).

218. Berdan, R. et al. In-memory reinforcement learning with moderately-stochastic conductance switching of ferroelectric tunnel junctions. *2019 Symp. VLSI Technology* T22-T23 (2019).

219. Lin, Y. et al. Demonstration of generative adversarial network by intrinsic random noises of analog RRAM devices. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 3–4 (IEEE, 2018).

220. Tavanaei, A., Ghodrati, M., Kheradpisheh, S. R., Masquelier, T. & Maida, A. Deep learning in spiking neural networks. *Neural Netw.* **111**, 47–63 (2019).

221. Serb, A. et al. Unsupervised learning in probabilistic neural networks with multi-state metal-oxide memristive synapses. *Nat. Commun.* **7**, 12611 (2016).

222. Shi, Y. et al. Neuroinspired unsupervised learning and pruning with subquantum CBRAM arrays. *Nat. Commun.* **9**, 5312 (2018).

223. Kim, S. et al. NVM neuromorphic core with 64k-cell (256-by-256) phase change synaptic array with on-chip neuron circuits for continuous in-situ learning. *2015 IEEE Int. Electron Devices Meeting (IEDM)* 17–1 (IEEE, 2015).

224. Pantazi, A., Wozniak, S., Tuma, T. & Eleftheriou, E. All-memristive neuromorphic computing with level-tuned neurons. *Nanotechnology* **27**, 355205 (2016).

225. Wang, W. et al. Learning of spatiotemporal patterns in a spiking neural network with resistive switching synapses. *Sci. Adv.* **4**, eaat4752 (2018).

226. Wang, Z. et al. In situ training of feed-forward and recurrent convolutional memristor networks. *Nat. Mach. Intell.* **1**, 434–442 (2019).
The first in situ training of a large-scale redox-RSM-based convolutional ANN and convolutional recurrent ANN for pattern and video classifications, respectively.

227. Chen, W.-H. et al. CMOS-integrated memristive non-volatile computing-in-memory for AI edge processors. *Nat. Electron.* **2**, 420–428 (2019).

228. Xue, C.-X. et al. A 1Mb multibit ReRAM computing-in-memory macro with 14.6 ns parallel MAC computing time for CNN based AI edge processors. *2019 IEEE Int. Solid-State Circuits Conf. (ISSCC)* 388–390 (IEEE, 2019).
One of the highest energy efficiencies reported to date for a convolutional ANN implemented by an integrated chip of redox-RSM crossbars and peripheral circuits for CIFAR-10 classification.

229. Li, C. et al. Long short-term memory networks in memristor crossbar arrays. *Nat. Mach. Intell.* **1**, 49–57 (2019).
The first in situ training of a large-scale redox-RSM-based recurrent ANN for walking gait classification.

230. Tsai, H. et al. Inference of long-short term memory networks at software-equivalent accuracy using 2.5 M analog phase change memory devices. *2019 Symp. VLSI Technology*. T82-T83 (2019).

231. Hu, S. G. et al. Associative memory realized by a reconfigurable memristive Hopfield neural network. *Nat. Commun.* **6**, 7522 (2015).

232. Zhou, Y. et al. Associative memory for image recovery with a high-performance memristor array. *Adv. Funct. Mater.* **29**, 1900155 (2019).

233. Yan, B. et al. RRAM-based spiking nonvolatile computing-in-memory processing engine with precision-configurable in situ nonlinear activation. *2019 Symp. VLSI Technology* T86–T87 (2019).
The first large-scale redox-RSM-based SNN for CIFAR-10 classification.

234. Du, C. et al. Reservoir computing using dynamic memristors for temporal information processing. *Nat. Commun.* **8**, 2204 (2017).

235. Moon, J. et al. Temporal data classification and forecasting using a memristor-based reservoir computing system. *Nat. Electron.* **2**, 480–487 (2019).

236. Midya, R. et al. Reservoir computing using diffusive memristors. *Adv. Intell. Syst.* **1**, 1900084 (2019).

237. Sun, Z. et al. Solving matrix equations in one step with cross-point resistive arrays. *Proc. Natl Acad. Sci. USA* **116**, 4123–4128 (2019).

238. Jeong, Y., Lee, J., Moon, J., Shin, J. H. & Lu, W. D. K-means data clustering with memristor networks. *Nano Lett.* **18**, 4447–4453 (2018).

239. Shin, J. H., Jeong, Y. J., Zidan, M. A., Wang, Q. & Lu, W. D. Hardware acceleration of simulated annealing of spin glass by RRAM crossbar array. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 3-3 (IEEE, 2018).

240. Cassinerio, M., Ciocchini, N. & Ielmini, D. Logic computation in phase change materials by threshold and memory switching. *Adv. Mater.* **25**, 5975–5980 (2013).

241. Huang, P. et al. Reconfigurable nonvolatile logic operations in resistance switching crossbar array for large-scale circuits. *Adv. Mater.* **28**, 9758–9764 (2016).

242. Sun, Z., Ambrosi, E., Bricalli, A. & Ielmini, D. Logic computing with stateful neural networks of resistive switches. *Adv. Mater.* **30**, 1802554 (2018).

243. Liu, R., Wu, H., Pang, Y., Qian, H. & Yu, S. A highly reliable and tamper-resistant RRAM PUF: Design and experimental validation. *2016 IEEE Int. Symp. Hardware Oriented Security Trust (HOST)* 13–18 (IEEE, 2016).

244. Zhang, R. et al. Nanoscale diffusive memristor crossbars as physical unclonable functions. *Nanoscale* **10**, 2721–2726 (2018).

245. Mazady, A., Rahman, M. T., Forte, D. & Anwar, M. Memristor PUF—a security primitive: theory and experiment. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **5**, 222–229 (2015).

246. Gao, L., Chen, P.-Y., Liu, R. & Yu, S. Physical unclonable function exploiting sneak paths in resistive cross-point array. *IEEE Trans. Electron Devices* **63**, 3109–3115 (2016).

247. Kim, J. et al. A physical unclonable function with redox-based nanoionic resistive memory. *IEEE Trans. Inf. Forensics Secur.* **13**, 437–448 (2018).

248. Huang, C.-Y., Shen, W. C., Tseng, Y.-H., King, Y.-C. & Lin, C.-J. A contact-resistive random-access-memory-based true random number generator. *IEEE Electron Device Lett.* **33**, 1108–1110 (2012).

249. Balatti, S., Ambrogio, S., Wang, Z. & Ielmini, D. True random number generation by variability of resistive switching in oxide-based devices. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **5**, 214–221 (2015).

250. Balatti, S. et al. Physical unbiased generation of random numbers with coupled resistive switching devices. *IEEE Trans. Electron Devices* **63**, 2029–2035 (2016).

251. Gallo, M. L., Tuma, T., Zipoli, F., Sebastian, A. & Eleftheriou, E. Inherent stochasticity in phase-change memory devices. *2016 46th European Solid-State Device Research Conf. (ESSDERC)* 373–376 (2016).

252. Jiang, H. et al. A novel true random number generator based on a stochastic diffusive memristor. *Nat. Commun.* **8**, 882 (2017).

253. Zalden, P. et al. Femtosecond x-ray diffraction reveals a liquid–liquid phase transition in phase-change materials. *Science* **364**, 1062–1067 (2019).

254. Milano, G. et al. Self-limited single nanowire systems combining all-in-one memristive and neuromorphic functionalities. *Nat. Commun.* **9**, 5151 (2018).

255. Zhu, X., Li, D., Liang, X. & Lu, W. D. Ionic modulation and ionic coupling effects in MoS₂ devices for neuromorphic computing. *Nat. Mater.* **18**, 141–148 (2018).

256. Wang, M. et al. Robust memristors based on layered two-dimensional materials. *Nat. Electron.* **1**, 130–136 (2018).

257. Yoshida, M., Suzuki, R., Zhang, Y., Nakano, M. & Iwasa, Y. Memristive phase switching in two-dimensional 1T-TaS₂ crystals. *Sci. Adv.* **1**, e1500606 (2015).

258. Cobas, E., Friedman, A. L., van't Erve, O. M., Robinson, J. T. & Jonker, B. T. Graphene as a tunnel barrier: graphene-based magnetic tunnel junctions. *Nano Lett.* **12**, 3000–3004 (2012).

259. Fei, Z. et al. Ferroelectric switching of a two-dimensional metal. *Nature* **560**, 336–339 (2018).

260. Fan, Y. et al. Magnetization switching through giant spin–orbit torque in a magnetically doped topological insulator heterostructure. *Nat. Mater.* **13**, 699–704 (2014).

261. Fan, Y. et al. Electric-field control of spin–orbit torque in a magnetically doped topological insulator. *Nat. Nanotechnol.* **11**, 352–359 (2016).

262. Han, J. et al. Room-temperature spin-orbit torque switching induced by a topological insulator. *Phys. Rev. Lett.* **119**, 077702 (2017).

263. Mahendra, D. C. et al. Room-temperature high spin–orbit torque due to quantum confinement in sputtered BiₓSe₍₁₋ₓ₎ films. *Nat. Mater.* **17**, 800–807 (2018).

264. Khang, N. H. D., Ueda, Y. & Hai, P. N. A conductive topological insulator with large spin Hall effect for ultralow power spin–orbit torque switching. *Nat. Mater.* **17**, 808–813 (2018).

265. Jerry, M. et al. Ferroelectric FET analog synapse for acceleration of deep neural network training. *2017 IEEE Int. Electron Devices Meeting (IEDM)* 6-2 (IEEE, 2017).

266. Danesh, C. D. et al. Synaptic resistors for concurrent inference and learning with high energy efficiency. *Adv. Mater.* **31**, 1808032 (2019).

267. van de Burgt, Y. et al. A non-volatile organic electrochemical device as a low-voltage artificial synapse for neuromorphic computing. *Nat. Mater.* **16**, 414–418 (2017).

268. Wang, Z. et al. Capacitive neural network with neuro-transistors. *Nat. Commun.* **9**, 3208 (2018).

269. Sangwan, V. K. et al. Multi-terminal memtransistors from polycrystalline monolayer molybdenum disulfide. *Nature* **554**, 500–504 (2018).

270. Sinova, J., Valenzuela, S. O., Wunderlich, J., Back, C. H. & Jungwirth, T. Spin hall effects. *Rev. Mod. Phys.* **87**, 1213–1259 (2015).

271. Yu, G. et al. Switching of perpendicular magnetization by spin–orbit torques in the absence of external magnetic fields. *Nat. Nanotechnol.* **9**, 548–554 (2014).

272. Fukami, S., Zhang, C., DuttaGupta, S., Kurenkov, A. & Ohno, H. Magnetization switching by spin–orbit torque in an antiferromagnet–ferromagnet bilayer system. *Nat. Mater.* **15**, 535–541 (2016).

273. Oh, Y.-W. et al. Field-free switching of perpendicular magnetization through spin–orbit torque in antiferromagnet/ferromagnet/oxide structures. *Nat. Nanotechnol.* **11**, 878–884 (2016).

274. Lau, Y.-C., Betto, D., Rode, K., Coey, J. M. D. & Stamenov, P. Spin–orbit torque switching without an external field using interlayer exchange coupling. *Nat. Nanotechnol.* **11**, 758–762 (2016).

275. Zhang, X. G. & Butler, W. H. Large magnetoresistance in bcc Co/MgO/Co and FeCo/MgO/FeCo tunnel junctions. *Phys. Rev. B* **70**, 172407 (2004).

276. Ikeda, S. et al. Tunnel magnetoresistance of 604% at 300K by suppression of Ta diffusion in CoFeB/MgO/CoFeB pseudo-spin-valves annealed at high temperature. *Appl. Phys. Lett.* **93**, 082508 (2008).

277. Schleicher, F. et al. Localized states in advanced dielectrics from the vantage of spin-and symmetry-polarized tunnelling across MgO. *Nat. Commun.* **5**, 4547 (2014).

278. Bean, J. J. & McKenna, K. P. Stability of point defects near MgO grain boundaries in FeCoB/MgO/FeCoB magnetic tunnel junctions. *Phys. Rev. Mater.* **2**, 125002 (2018).

279. Dong, Q. et al. A 1Mb 28nm STT-MRAM with 2.8 ns read access time at 1.2 V VDD using single-cap offset-cancelled sense amplifier and in-situ self-write-termination. *2018 IEEE Int. Solid-State Circuits Conf. (ISSCC)* 480−482 (2018).

280. Grezes, C. et al. Write error rate and read disturbance in electric-field-controlled magnetic random-access memory. *IEEE Magn. Lett.* **8**, 3102705 (2017).

281. Kanai, S. et al. Magnetization switching in a CoFeB/MgO magnetic tunnel junction by combining spin-transfer torque and electric field-effect. *Appl. Phys. Lett.* **104**, 212406 (2014).

282. Van Beek, S. et al. Impact of self-heating on reliability predictions in STT-MRAM. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 25-2 (2018).

283. Fuller, E. J. et al. Li-ion synaptic transistor for low power analog computing. *Adv. Mater.* **29**, 1604310 (2016).

284. Tang, J. et al. ECRAM as scalable synaptic cell for high-speed, low-power neuromorphic computing. *2018 IEEE Int. Electron Devices Meeting (IEDM)* 13−1 (IEEE, 2018).

285. Chen, P.-Y. et al. Mitigating effects of non-ideal synaptic device characteristics for on-chip learning. *Proc. IEEE/ACM Int. Conf. Computer-Aided Design* 194−199 (IEEE Press, 2015).

286. Kim, G. H. et al. 32 × 32 crossbar array resistive memory composed of a stacked Schottky diode and unipolar resistive memory. *Adv. Funct. Mater.* **23**, 1440−1449 (2013).

287. Yoon, K. J. et al. Double-layer-stacked one diode-one resistive switching memory crossbar array with an extremely high rectification ratio of 10^9. *Adv. Electron. Mater.* **3**, 1700152 (2017).

288. Choi, B. J. et al. Trilayer tunnel selectors for memristor memory cells. *Adv. Mater.* **28**, 356−362 (2016).

289. Lee, M. J. et al. A plasma-treated chalcogenide switch device for stackable scalable 3D nanoscale memory. *Nat. Commun.* **4**, 2629 (2013).

290. Cha, E. et al. Nanoscale (~ 10nm) 3D vertical ReRAM and NbO_2 threshold selector with TiN electrode. *2013 IEEE Int. Electron Devices Meeting (IEDM)* 10−5 (2013).

291. Luo, Q. et al. $Nb_{1-x}O_2$ based universal selector with ultra-high endurance ($>10^{12}$), high speed (10ns) and excellent Vth stability. *2019 Symp. VLSI Technology* T236-T237 (2019).

292. Virwani, K. et al. Sub-30nm scaling and high-speed operation of fully-confined access-devices for 3D crosspoint memory based on mixed-ionic-electronic-conduction (MIEC) materials. *2012 IEEE Int. Electron Devices Meeting (IEDM)* 2−7 (2012).

293. Midya, R. et al. Anatomy of Ag/hafnia-based selectors with 10^{10} nonlinearity. *Adv. Mater.* **29**, 1604457 (2017).

294. Sun, L. et al. Self-selective van der Waals heterostructures for large scale memory array. *Nat. Commun.* **10**, 3161 (2019).

295. Linn, E., Rosezin, R., Kugeler, C. & Waser, R. Complementary resistive switches for passive nanocrossbar memories. *Nat. Mater.* **9**, 403−406 (2010).

296. Yang, Y., Mathew, J., Pradhan, D. K., Ottavi, M. & Pontarelli, S. Complementary resistive switch based stateful logic operations using material implication. *2014 Design, Automation & Test in Europe Conference & Exhibition (DATE)* 185 (IEEE, 2014).

297. Kataeva, I. et al. Towards the development of analog neuromorphic chip prototype with 2.4 M integrated memristors. *2019 IEEE Int. Symp. Circuits and Systems (ISCAS)* 1−5 (IEEE, 2019).

298. Pei, J. et al. Towards artificial general intelligence with hybrid Tianjic chip architecture. *Nature* **572**, 106−111 (2019).

299. Cheng, H. et al. A high performance phase change memory with fast switching speed and high temperature retention by engineering the $Ge_xSb_yTe_z$ phase change material. *2011 Int. Electron Devices Meeting* 3−4 (IEEE, 2011).

300. Kim, Y.-B. et al. Bi-layered RRAM with unlimited endurance and extremely uniform switching. *2011 Symp. VLSI Technology-Digest of Technical Papers* 52−53 (IEEE, 2011).

Author contributions
All authors contributed to the discussion of content and reviewed and edited the manuscript prior to submission. Z.W. and J.J.Y. researched data for the article. Z.W., G.W.B., C.S.H., K.L.W., Q.X. and J.J.Y. wrote the article.

Competing interests
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information
Supplementary information is available for this paper at https://doi.org/10.1038/s41578-019-0159-3.