Synaptic electronics and neuromorphic computing

Article *in* Sciece China. Information Sciences · June 2016

DOI: 10.1007/s11432-016-5565-1

CITATIONS
14

READS
282

3 authors:

Navnidhi Upadhyay
University of Massachusetts Amherst
7 PUBLICATIONS **83** CITATIONS
SEE PROFILE

Saumil Joshi
Micron Technology, Inc.
29 PUBLICATIONS **650** CITATIONS
SEE PROFILE

Jianhua Joshua Yang
HP Inc.
198 PUBLICATIONS **9,623** CITATIONS
SEE PROFILE

SCIENCE CHINA
Information Sciences

• REVIEW •
Special Focus on Advanced Microelectronics Technology

June 2016, Vol. 59 061404:1–061404:26
doi: 10.1007/s11432-016-5565-1

Synaptic electronics and neuromorphic computing

Navnidhi K. UPADHYAY, Saumil JOSHI & J. Joshua YANG*

Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA 01003*, USA*

Received January 26, 2016; accepted February 23, 2016; published online May 11, 2016

Abstract In order to map the computing architecture and intelligent functions of the human brain on hardware, we need electronic devices that can emulate biological synapses and even neurons, preferably at the physical level. Beginning with the history of neuromorphic computation, in this article, we will briefly review the architecture of the brain and the learning mechanisms responsible for its plasticity. We will also introduce several memristive devices that have been used to implement electronic synapses, presenting some important milestones in this area of research and discussing their advantages, disadvantages, and future prospects.

Keywords memristors, neuromorphic engineering, RRAM, synapses, synaptic devices

Citation Upadhyay N K, Joshi S, Yang J J. Synaptic electronics and neuromorphic computing. Sci China Inf Sci, 2016, 59(6): 061404, doi: 10.1007/s11432-016-5565-1

1 Introduction

Alan Turing, widely considered the father of theoretical computer science and artificial intelligence, published a seminal paper titled 'On computable numbers' in 1936 [1]. The paper presented a formal proof that any mathematical computation that is representable as an algorithm can be performed by a machine. Such a machine, known as 'Turing machine', is the basis for today's digital computers. Modern computers are based on the von Neumann architecture, proposed by John von Neumann in 1945 [2], and although remarkably successful in handling many problems, these machines have a very complex instruction-processing unit that sequentially and frequently accesses a memory bank to perform a set of instructions, therefore suffering from the so-called von Neumann bottleneck.

Interestingly, Turing also anticipated for the first time a computing system inspired by neural systems. In 1948, he described a machine that consisted of artificial neurons connected together in an arbitrary pattern with modifier devices between them [3]. Such neural networks consist of simple processing units, and due to their intrinsic analog capabilities and massive parallel connections, can potentially solve many problems that are computationally intensive for von Neumann systems. Examples of such problems are adaptive behaviour, learning by association, pattern recognition, fuzzy logic, etc. Solving these problems becomes more and more important in the so-called big-data era. Consequently, neural network machine architectures, many of which are inspired by biological systems such as the human brain, are currently under intensive investigation world-wide. One of the major goal behind these efforts is to build a "thinking machine" based on neurophysiological models [4].

* Corresponding author (email: jjyang@umass.edu)

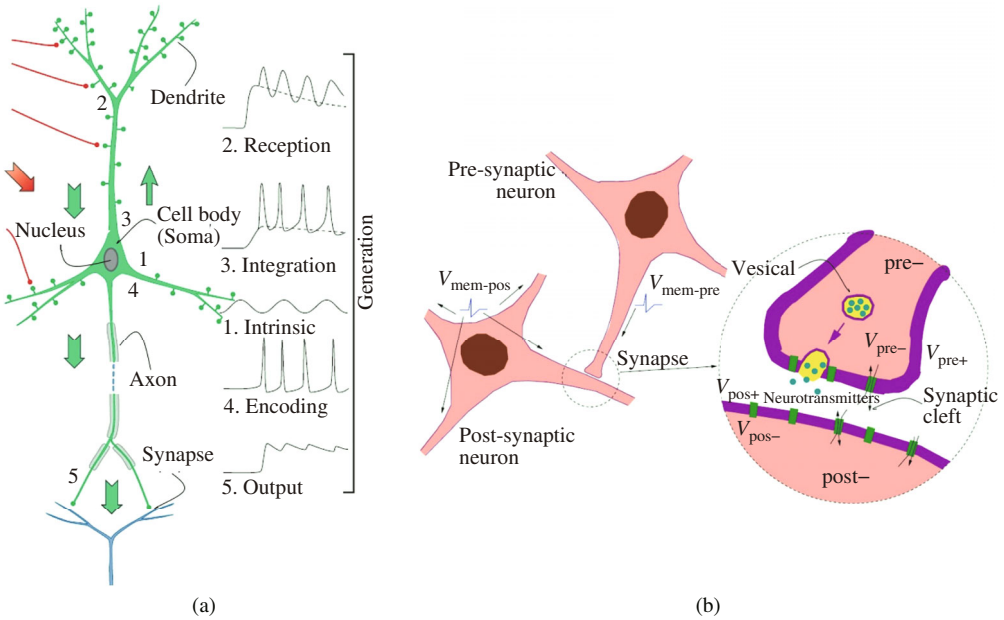


Figure 1 (Color online) Structure of the neuron and the synapse. (a) Nerve cells have four main regions (dendrite, cell body, axon, and synapse) and five main functions (numbered). Reprinted with permission from [5]. Copyright 2008, Academic Press. (b) Illustration of the synapse junction between the pre-synaptic neuron and the post-synaptic neuron, and its schematic structure. Reprinted with permission from [8]. Copyright 2011, Frontiers in Neuroscience.

2 Architecture of the brain

The human brain is likely the most efficient computing system so far, which is not surprising considering that the brain has evolved for millions of years, while the history of computers is still far below a century. The brain consists of many different types of cells, but the primary functional unit is a cell called neuron. The signals generated by neurons are responsible for emotions, memories, movements, thinking, and feelings. A human brain has approximately 10^{11} neurons. As shown in Figure 1(a), each neuron consists of a cell body called soma, an axon, and multiple dendrites [5]. Axon carries information from soma to a junction, where it is picked-up by dendrites of other neurons. The junction is called synapse (Greek, syn: union, association) and the strength of the synapse (synaptic weight) decides the bonding between two neurons, which can be altered by neural activities. This process is known as synaptic plasticity, and is believed to be the reason behind learning and memorization capabilities of the brain. The synaptic inputs picked up by the dendrites of other neurons are then integrated by their own cells, encoded in the form of action potentials, and distributed to even more neurons from their axon terminals. Typically, each neuron is connected with about 5000–10000 other neurons, resulting in a large number (about 10^{15}) of biological synapses in the human brain. The operating frequency of our brain is in the range of 1–10 Hz and it consumes around 10–100 W of power for its functioning [6], yielding an energy consumption of approximately 1–10 fJ per synaptic event.

In order to understand how a neuron works, it is necessary to know how the different parts of the neuron interact. Each neuron carries out five basic functions as described in Figure 1(a): (1) Generate intrinsic membrane activity in the neuron; (2) Receive synaptic inputs in dendrites; (3) Combine synaptic inputs with the intrinsic membrane activity; (4) Generate outputs in the form of action potentials; (5) Distribute the outputs from axon terminals.

There are two types of synaptic transmissions identified between the neurons, namely, electrical and chemical. At electrical synapses, the cytoplasm of adjacent cells are directly connected by clusters of intercellular channels called gap junctions [7]. These gap junctions allow passive current flow between cell membranes. The chemical synapse between two neurons is illustrated in Figure 1(b) [8]. At these synapses, information transfer from one neuron to another occurs through the release of neurotransmitter by one neuron (pre-synaptic neuron) and detection of the neurotransmitter by an adjacent neuron (post-

synaptic neuron) [9]. Electrical synapses generally maintain a stable functional structure and strength, whereas chemical synapses have the ability to exhibit plasticity [10]. The synaptic plasticity shown by chemical synapses is one of the most important neurochemical foundations of learning and memory in the brain. Chemical synapses can be of two types: excitatory or inhibitory. Whether a synapse is excitatory or inhibitory depends entirely on the receptors on the postsynaptic membrane. In the brain, the receptors for glutamate neurotransmitter are typically excitatory, whereas the receptors for GABA (Gamma-Aminobutyric acid) neurotransmitter are typically inhibitory. An excitatory receptor results in an Excitatory Postsynaptic Potential (EPSP) and drives the postsynaptic neuron closer to the depolarization threshold which makes the cell "fire" an action potential. An inhibitory receptor results in an Inhibitory Postsynaptic Potential (IPSP) and drives the postsynaptic neuron further from the depolarization threshold. Axon terminals from many neurons can connect to a given neuron and release a variety of neurotransmitters which impinge on excitatory and inhibitory receptors to produce EPSPs and IPSPs. The postsynaptic neuron behaves like a tiny computer, summing the EPSPs and IPSPs which determine whether or not it will "fire". When a neuron fires, the resulting action potential travels towards the synaptic cleft through its axon. Arrival of the action potential at the axon terminal results in the merging of neurotransmitter vesicles with the presynaptic membrane and a subsequent release of the neurotransmitters into the synaptic cleft. The neurotransmitter diffuses through the synaptic cleft, binds to and activates a receptor in the postsynaptic membrane, modifying the plasticity of the connection, i.e., the synaptic weight. Hereon the chemical synapse will be simply referred to as a synapse, unless otherwise stated.

3 Emulating the brain using hardware: neuromorphic systems

Supercomputers can store more information than the human brain and can compute equations faster, but even the biggest and fastest supercomputers in the world cannot match the overall processing power of the human brain in performing many tasks, such as pattern recognition, perception, motor control, flexibility in changing environments, learning, and ultimately, intelligent cognition. Still, they are almost a million times larger and consume a million times more power than a human brain. Figure 2 [11] illustrates the basic differences in the architecture of the brain and a conventional computer.

Due to the amazing capabilities of the brain, it is very attractive to study its structure and working mechanisms in order to mimic it using electronic circuits. The study of the brain and its inspiration in developing computing systems has led to a new form of computer architecture, known as neuromorphic architecture. This field combines knowledge from different disciplines such as biology, physics, mathematics, computer science, and engineering in order to design artificial neural systems [12]. The physical architecture and design principles of these artificial systems are based on those of biological nervous systems.

Although the behaviour and connections between neurons can be partially simulated on a conventional computer, such a system will consume excessive power and is not capable of exploiting the architecture of the brain due to the fundamental differences between these two systems. As a result, a race is on to develop new types of devices and hardware architectures that can better resemble bio-intelligent systems at the physical level, and thus more efficiently emulate the brain at the functional level. For instance, the so-called neuromorphic circuit is built from devices that behave like neurons, transmitting and responding to information sent in the form of spikes rather than continuously varying voltages. The concept of neuromorphic engineering was first developed by Carver Mead in the late 1980s [13]. Mead described it as "using VLSI systems containing electronic analog circuits to mimic neurobiological architectures present in the nervous system".

As the brain consumes very low energy, the first crucial step in realising these hardware systems is to achieve a suitable device that can function as a synapse with low power consumption and desired plasticity. Accordingly, researchers have explored a variety of device systems with programmable conductance, including FET based devices [14], phase change memory (PCM) [15], resistive change random-access

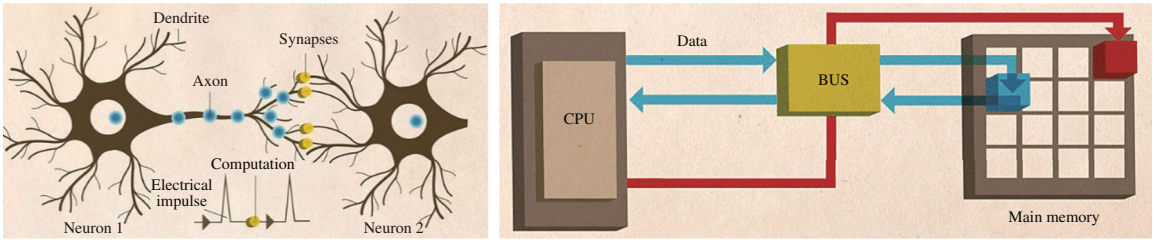


Figure 2 (Color online) Illustration of the architecture of the brain compared to a traditional computer. (a) Computing architecture of the brain consisting of neurons communicating through junctions called synapses which store memory in the form of connection strength (synaptic weight), and (b) traditional computer architecture with a central processing unit (CPU) communicating with the main memory unit through a bus. Reprinted with permission from [11]. Copyright 2010, IEEE Spectrum.

memory (RRAM) [16], and others.

4 Implementing brain-like architecture

Several research efforts have been directed towards developing architectures for neuromorphic computing, the majority of them using established analog and digital technologies with CMOS devices [17,18]. They are summarized in Figure 3. DARPA funded the SyNAPSE (Systems of Neuromorphic Adaptive Plastic Scalable Electronics) program in 2008 [19] with a goal of building a system matching a mammalian brain. Under this program, IBM introduced the TrueNorth chip in 2014 using 5 billion transistors making up 256 million synapses. It consumes 68 mW power with a 400×240 pixel video input at 30 frames per second [20,21]. Manchester University is developing the SpiNNaker (Spiking Neural Network Architecture), a highly parallel neuromorphic supercomputer, using more than a million ARM processors. The machine will simulate the behaviour of 1 million neurons, ten times more than the number of neurons in a mouse [22]. The Blue Brain Project is aimed at simulating sections of the rat brain using the IBM BlueGene supercomputer [23]. All these architectures would consume significantly large power if scaled up to the same complexity level as the human brain. Reducing the energy consumption would remain extremely challenging as long as the conventional CMOS technology is used. In addition, none of the implementations truly mimic the biological synapse.

Compared to these digital implementations, the BrainScaleS project employs analog circuits, which are faster and consume much less energy [24]. Similarly, the Neurogrid project at Stanford University utilizes analog computation to simulate ion-channel activity and digital communication to wire the synaptic connections. This hybrid approach reduces the power consumption by five orders of magnitude compared to a supercomputer performing the same task [25, 26]. However, these architectures cannot match the compactness or parallelism of biological systems.

5 Synaptic plasticity mechanism and learning

Neuroplasticity, or brain plasticity, is the capability of the brain to modify existing synaptic strengths or create completely new connections as it needs during the process of learning and memorizing new information. Sometimes brain annihilates the neural connections that are no longer needed and strengthens the necessary ones, in order to tune itself to meet new requirements. Over the last several years, neuroscientists have extensively studied the neural connections present in the brain and recognized many molecular mechanisms that are responsible for neuroplasticity.

Plasticity in the brain can be rate dependent or timing dependent. Several mechanisms have been proposed to explain the different forms of plasticity in the brain. In 1949 Donald Hebb proposed the basic mechanism behind synaptic plasticity also known as Hebbian learning rule [27]: "When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A's efficiency as one of the cells

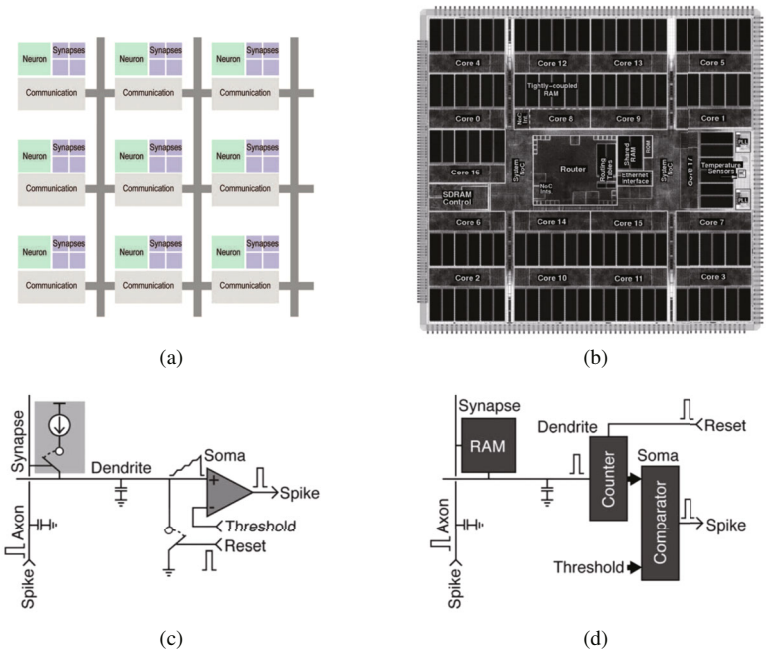


Figure 3 (Color online) Various techniques to implement brain-inspired computing. (a) Schematic illustration of the brain-like architecture used in TrueNorth chip [21], where memory, computation, and communication blocks are integrated as modules operating in parallel with each other using an event-driven network. Reprinted with permission from [21]. Copyright 2014, The American Association for the Advancement of Science. (b) A die of the SpiNNaker computer with 18 ARM cores, fabricated on a 130 nm CMOS process. Reprinted with permission from [22]. Copyright 2013, IEEE. (c) Analog implementation of neuron presented in [26], where spikes on the wire labeled 'axon' are integrated by the capacitance of the wire labeled 'dendrite'. The resulting voltage is compared with a threshold using a comparator and a spike is generated when the voltage exceeds threshold. (d) Digital implementation of a neuron consists of a counter which is incremented every time an incoming spike triggers a 1 out of the bit cell. The output of the counter is compared with the threshold and a spike is generated when the count exceeds the threshold, after which the counter is reset. Reprinted with permission from [26]. Copyright 2014, IEEE.

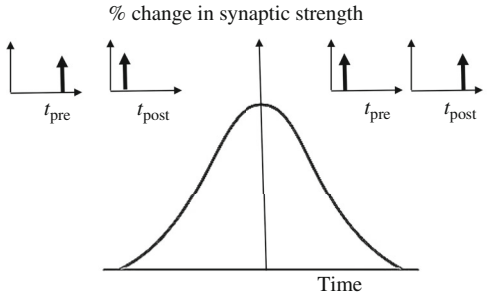


Figure 4 The curve corresponding to Hebbian learning rule, showing the percentage change in synaptic strength as a function of the time difference between the pre and post-synaptic spikes.

firing B, is increased". This mechanism has often been summarized as "cells that fire together, wire together". Hebbian learning rule is described in Figure 4. As shown, the curve is symmetric and the temporal order of the input and output has no influence on the change in synaptic strength. However, this picture is incomplete as observed in many experiments conducted over the last two decades [28–32]. Other rules, such as the BCM rule [33], the covariance rule [34], etc. have been proposed as extensions or alternatives to the Hebb rule. No single rule has been successful in explaining all varieties of synaptic plasticity observed in nature. In the following paragraphs, we will discuss some of the synaptic plasticity mechanisms that are crucial for realization of neuromorphic hardware.

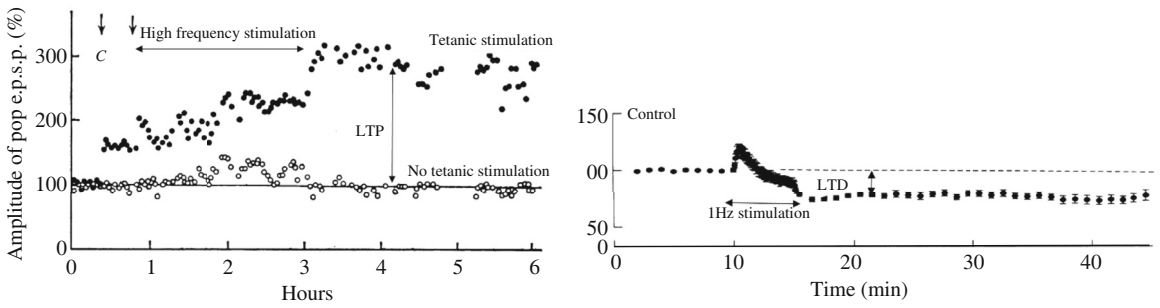


Figure 5 Experimental demonstration of long-term potentiation (LTP) and long-term depression (LTD) in a biological synapse showing synaptic conductance as a function of time. (a) The solid circle corresponds to the conductance due to high frequency tectanic stimulation which results in LTP, and the hollow circle corresponds to the conductance in the absence of tectanic stimulation. Reprinted with permission from [36]. Copyright 1970, John Wiley and Sons. (b) Plot showing the change in synaptic conductance due to low frequency stimulation, resulting in LTD. The dotted horizontal line shows the conductance level when there is no stimulation applied. Reprinted with permission from [37]. Copyright 1993, The American Association for the Advancement of Science.

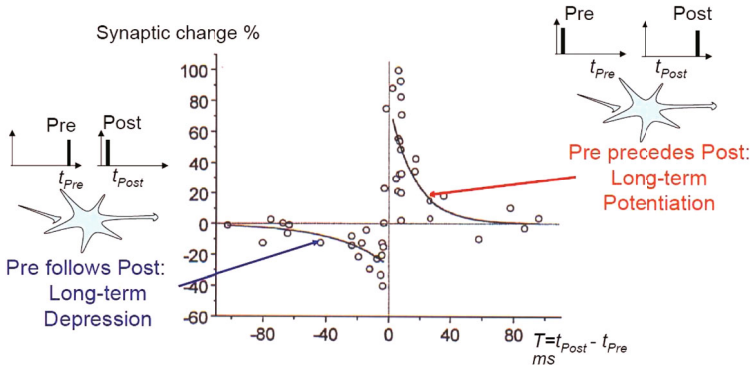


Figure 6 (Color online) The percentage change in the synaptic weight as a function of the time difference between the pre and post-synaptic spikes across a biological synapse. If pre- follows post-spike, the synapse undergoes LTD, conversely if pre-precedes post-spike, synapse undergoes LTP. Reprinted with permission from [32]. Copyright 1998, Journal of Neuroscience.

5.1 Long-term plasticity

Long-term plasticity refers to persistent activity-dependent changes in the synaptic strength. Due to its durability, this form of synaptic plasticity is widely believed to be the reason for learning and memory [35]. Some patterns of synaptic activity in pre and post-synaptic neurons produce a long-lasting increase in synaptic strength known as long-term potentiation (LTP), whereas other patterns of activity produce a long-lasting decrease in synaptic strength, known as long-term depression (LTD). In such rate dependent synapses, LTP occurs due to high frequency stimulations and low frequency stimulations induce LTD, as demonstrated in Figure 5 [36, 37].

Spike-timing-dependent plasticity (STDP) is an asymmetric form of Hebbian learning which depends on the temporal correlations between pre and post-synaptic spikes [38]. The relative timing between the pre and post-synaptic spikes determines the direction and the magnitude of changes in synaptic weights. Post-synaptic spiking that occurs repetitively within a time window after presynaptic activation results in LTP, whereas post-synaptic spiking within a time window before pre-synaptic activation leads to LTD, as shown in Figure 6. STDP can be divided into two types: additive STDP and multiplicative STDP. If the learning function depends only on the time interval (ΔT) between pre and post-neuron spikes, but not on the actual weight (w), this type of weight-independent STDP learning rule is usually called "additive STDP". Additive STDP requires the weight values to be bounded to an interval because weights will stabilize at one of their boundary values. On the other hand, in "multiplicative STDP" the learning function is also a function of the actual weight value $f(w, \Delta T)$ [39–41]. In multiplicative STDP weights can stabilize to intermediate values inside the boundary definitions. Thus, it is often not even necessary

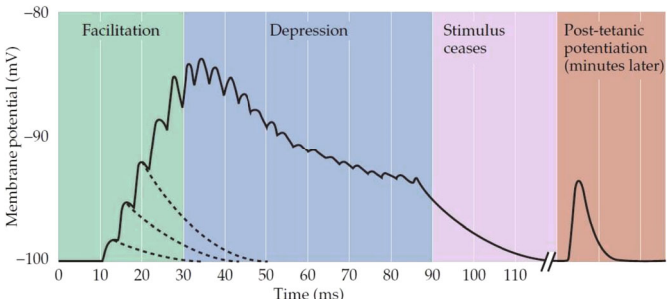


Figure 7 (Color online) Illustration of short-term plasticity at the neuromuscular synapse. Presynaptic motor nerve was stimulated by train of electrical pulses. In the beginning facilitation of the EPP occurs and is followed by depression of the EPP. After the removal of stimulation EPP goes back to resting state. Reprinted with permission from [42]. Copyright 2001, Sunderland Sinauer Assc.

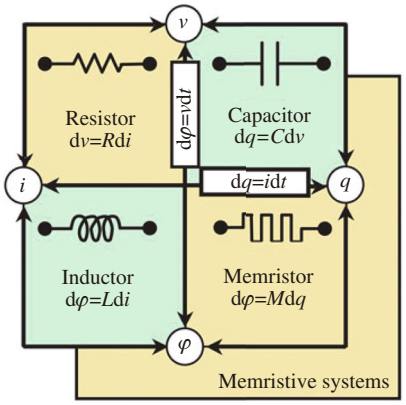


Figure 8 (Color online) Four basic electrical circuit elements: Relationships between the four circuit variables and four fundamental two-terminal circuit elements: resistor, capacitor, inductor and memristor. Reprinted with permission from [49]. Copyright 2008, Nature Publishing Group.

to enforce boundary conditions for the weight values [41]. STDP has been widely used by computational neuroscientist because of its simplicity, biological plausibility, and computational power.

5.2 Short-term plasticity

Short-term plasticity acts on a shorter time scale, typically in the order of tens of milliseconds to a few minutes. The modification in synaptic strength due to short-term plasticity is temporary, such that in the absence of continued presynaptic activity, synaptic strength quickly returns to its initial level, as shown in Figure 7 [42]. Compared to long-term plasticity, which is an experience-dependent modification, in case of short-term plasticity the response of post synaptic neuron reflects the history of pre-synaptic activity. Two types of short-term plasticity, with opposite effects on synaptic strength, have been observed in biological synapses: short-term depression and short-term facilitation.

Short-term facilitation is a transient increase in synaptic strength (Figure 7) that occurs when two or more action potentials arrive at presynaptic terminal very close in time, due to which calcium builds up in presynaptic side. As a result, more neurotransmitters are released by a subsequent presynaptic action potential. A high-frequency train of presynaptic firings (referred to as tetanus) can cause an even more sustained rise of presynaptic calcium levels which results in another form of synaptic plasticity called post-tetanic potentiation. Repeated presynaptic activities can also cause continuous depletion of the synaptic vesicles available for release into the synaptic cleft, which results in decrease of synaptic strength or depression. Synaptic strength decreases until the pool of neurotransmitters replenished via the process involved in recycling of synaptic vesicles.

6 Realization of synaptic devices

The recent progress in experimental realization of nanoscale memristive devices [43–49] has opened possibilities for implementing true brain architectures based on hardware with plastic synaptic connections. A memristor is any passive electronic circuit element that displays a pinched hysteresis loop in its I-V characteristics [50]. The device was first proposed by Leon Chua in 1971. He observed that there should be six different relationships between the four circuit variables: charge, current, voltage, and magnetic flux. While two relationships were covered by basic physical laws and three were given by the definitions of the then known circuit elements (resistor, capacitor, and inductor), one relation was left unaccounted for (see Figure 8). This led him to postulate the fourth fundamental passive circuit element, called

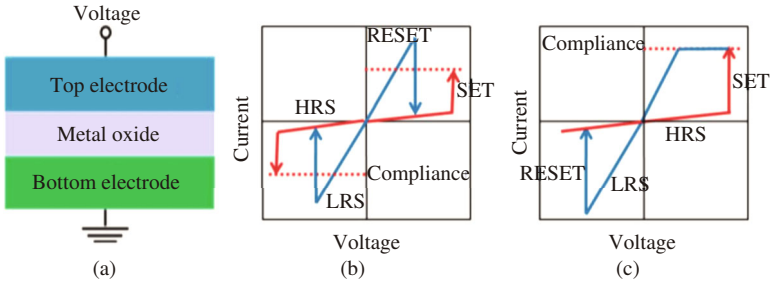


Figure 9 (Color online) Properties of resistance switches without electrochemically reactive metal. (a) Schematic of the metal-insulator-metal (MIM) structure for metal–oxide RRAM, and schematic of the device I-V characteristics, showing two modes of operation: (b) unipolar and (c) bipolar. Reprinted with permission from [56]. Copyright 2012, IEEE.

the memristor. The memristance (M) describes the remaining relation between charge q and flux φ, $\mathrm{d}\varphi = M\mathrm{d}q$.

Although Chua demonstrated that the memristor has many interesting and valuable circuit properties, and pinched hysteresis loops had been observed in the I-V characteristics of oxide based resistance switches around the same time [51,52], the link between the memristor theory and experimental demonstration was not established until 2008 [49]. Memristors have attracted significant attention since then due to a number of promising applications, including non-volatile random access memory (NVRAM) and neuromorphic computing. The concept of memristors has also been generalized to cover devices beyond the resistance switches (often called RRAM or ReRAM for memory applications), including phase change memory devices, spintronic devices, Spin Torque Transfer (STT) MRAM, ferroelectric devices [53], etc.

Many of these device technologies have been explored as electronic synapses with programmable conductance. As expected, different types of devices have different characteristics. To mimic the ability of the brain to perform fuzzy, fault-tolerant, and stochastic computation, without sacrificing either its space or power efficiency, these devices should have incremental (or analog) conductance change, under short pulses (e.g., 100 ns) with low voltages (e.g., 0.5 V) and low currents (e.g., 100 nA). Neuromorphic systems based on such emerging nanoscale devices can potentially improve density and power consumption by at least a factor of 10, as compared with the conventional CMOS implementations [54]. In the following sections we will review the research efforts on synaptic devices, categorizing them according to their dominant switching mechanisms.

6.1 Resistance switches without electrochemically active metals

Resistance switches are made using a conductor/insulator/conductor structure. Depending on whether an electrochemically active metal (e.g., Cu or Ag) layer is used as the conductor layer or not, the switching mechanism can be fairly different. Without such active metal layer, the devices switch via valence change mechanism (or electron trapping/de-trapping in some cases)which will be discussed in this section. These devices are often called RRAM or ReRAM devices. With an active metal the switching mechanism is typically electrochemical metallization. These devices, also called RRAM, are more frequently referred to as Programmable Metallization Cell (PMCs), which will be discussed separately in the next section. The resistance switches are one of the leading candidates for non-volatile memory (NVM) technology due to their good scalability, performance, and low-power consumption. They do not suffer from the scaling limits of the dynamic random access memory (DRAM) and Flash memory [43]. However, they are still susceptible to some of the problems associated with scaling nanodevices, such as high defect rates, high device variability, and device aging [55,56]. These issues are of greater concern in memory applications than in neuromorphic computing applications.

A schematic of the valence change memory (VCM) device is shown in Figure 9. The exact switching and electron transport mechanism varies significantly with materials used, but it is generally true in all types of devices that electric field and/or Joule heating causes ionic motion and local structural changes in the switching layer, resulting in a measurable change in the device resistance. Information can be stored as the resistance states of the device.

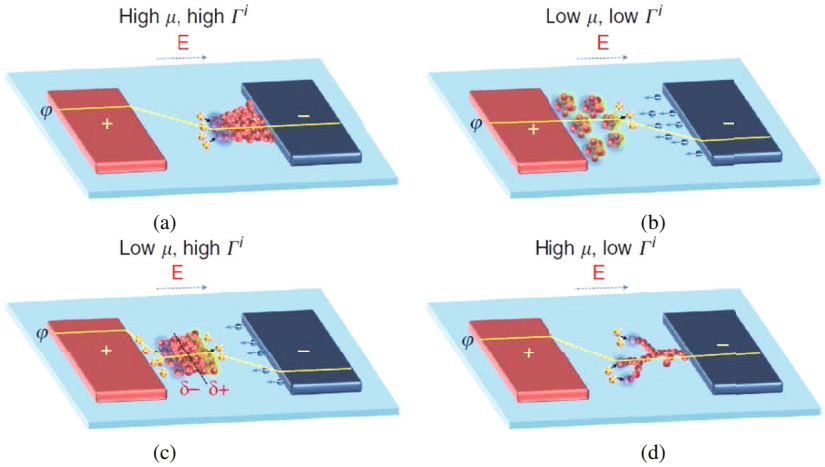


Figure 10 (Color online) Illustration of filament growth dynamics controlled by kinetic parameters. Filament growth (a) when both m and r are high, (b) both m and r are low, (c) m is low but r is high, and (d) m is high but r is low. Reprinted with permission from [58]. Copyright 2014, Nature Publishing Group.

The devices can have multiple resistance levels between two boundary resistance states, i.e., Lower Resistance State (LRS) and Higher Resistance State (HRS). The process of switching from HRS to LRS is called 'set', and a switch from LRS to HRS is called 'reset'. Usually an as-prepared sample needs a higher voltage for its initial set process, which is called "electroforming", with the corresponding voltage being called the "electroforming voltage" [57]. These devices can be further categorized as unipolar or bipolar, based on the voltage polarity used for the switching operations. In the unipolar case (Figure 9(b)), irrespective of voltage polarity, set and reset operations only depend on the magnitude of applied voltages or currents. In the case of bipolar switching (Figure 9(c)), the set and reset have to be done with opposite voltage polarities.

The fundamental physics and mechanism of resistive-switching in RRAM devices are surprisingly divergent and has not yet been completely understood. Based on current understanding of the switching mechanisms, RRAM devices can be loosely divided into two categories: *cation-based switching* and *anion-based switching*.

Cation-based switching systems form metallic filament inside dielectric medium. Yang et al. [58] presented a unified framework for the filament growth dynamics in these types of memristive devices. Depending on the kinetic factors (ion mobility m and redox rate r) filament growth dynamics may be divided into four categories (Figure 10):

(a) *High and homogeneous* m and r: In this case ions can reach the inert electrode without reducing or clustering inside the dielectric layer, so filament growth will begin from the inert electrode. Due to high r, ion supply is high which will lead to an inverted cone shaped filament (Figure 10(a)). This kind of filament growth can be found in conventional ECM devices (Subsection 6.2).

(b) *Low and inhomogeneous* m *and low* r: In this case ions can nucleate inside the dielectric medium and further filament growth will happen by means of cluster displacement and repeated splitting-merging process (Figure 10(b)).

(c) *Low* m *and high* r: In this case also, ions can nucleate inside the dielectric medium, but due to high r, incoming ions from the cathode will reduce at the site of nucleation. This process will continue till the gap between nucleation site and cathode is filled and the process will repeat again, giving an effective filament growth towards inert electrode (Figure 10(c)).

(d) *High* m *and low* r: Due to high m, there may not be any nucleation inside the dielectric and hence ions will reduce at the surface of the inert electrode. Due to limited ion supply (low r) reduction occurs at the edges with high electric field strengths, resulting in a branched growth of the filament towards the active electrode (Figure 10(d)).

Based on the above discussion it can be concluded that the nucleation sites and the direction of filament growth are controlled by the ion mobility m, while the shape and geometry of filament will depend on

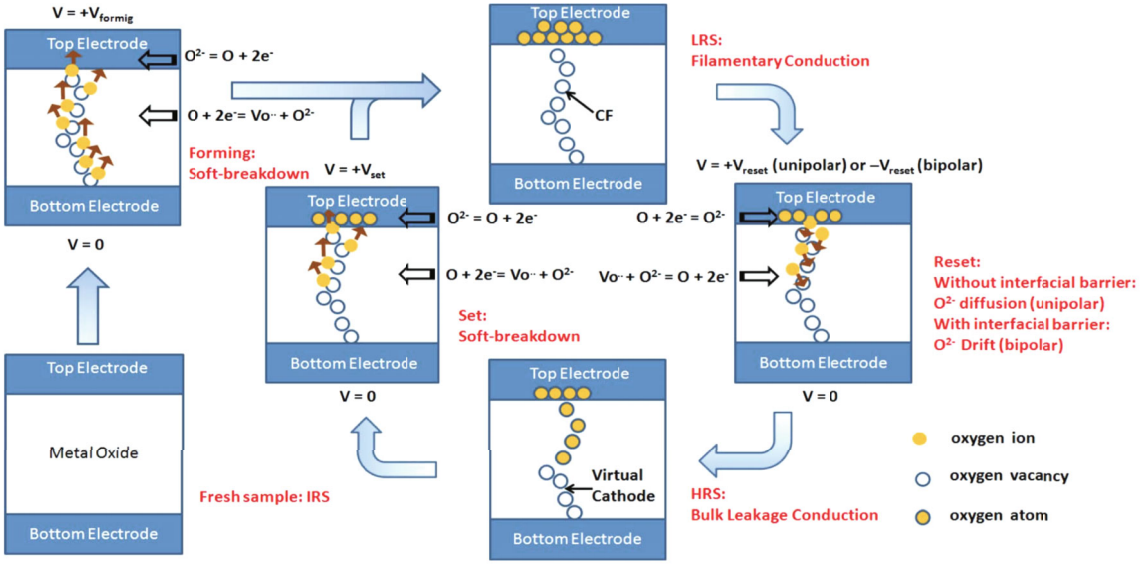


Figure 11 (Color online) Schematic illustration of the switching process for anion-based RRAM devices. Reprinted with permission from [56]. Copyright 2012, IEEE.

Table 1 Summary of the synaptic properties of RRAM devices presented in literature

Device structure	Size	Operating voltage	Programming time	Energy consumption	R_{on}/R_{off}	Switching mechanism	Gradual set reset
$Pt/Al_2O_3/Ni$ [59]	100 μm	+/−0.5 V	-	-	-	Filamentary	Yes
PdH_x/Nafion/PdH_x [60]	∼100 μm	-	∼1 ms	30 nJ	-	Non-filamentary	-
$Pt/WO_{3-x}/Pt$ [61]	25 μm	+/−3 V	-	-	-	Filamentary	Yes
Pt/TiN/PCMO/Pt [62]	-	+/−3 V	10 ms	-	-	Filamentary	Yes
W/TiN/Mn:HfO_2/Ru [63]	20 nm	+/−1 V	10 ms	482 fJ	-	Filamentary	Yes
Ti/Pt/Gd:HfO_2/TiN [64]	10 μm	+/−3 V	0.1–1 μs		1/100	Filamentary	Multi bit
Ta/TaO_x/TiO_2/Ti [65]	100 μm	+/−5 V				Non-filamentary	Yes
TiN/HfO_x/AlO_x/Pt [66,67]	0.5 μm	+/−3 V	10 ns	6 pJ	1/120	Filamentary	7: LRS 6: HRS
Pt/GdO_x+Cu-MoO_x/Pt [68]	100 nm	+/−4 V		-	1/20	Filamentary (forming free)	Analog memory

the rate of reduction.

Anion-based switching systems operate using the following sequence of events: soft breakdown of the dielectric and creation of conductive filaments consisting of oxygen vacancies, which sets the device. If the polarity of the applied voltage changes, a recombination of oxygen vacancies with oxygenions takes place eventually breaking the conducting filaments (or changing the composition of the filaments) and resetting the device (Figure 11).

Table 1 [59–68] summarizes the device structure, size, operating voltage, programming time, energy consumption, R_{on}/R_{off} ratio, switching mechanism, and gradual set/reset characteristics of some of the synaptic devices present in the literature. Neuromorphic computing in general is less demanding on device performance requirements than memory applications, especially on device variability and sneak-path current issues. Most valence change based devices show filamentary type of switching which has the following limitations:

1. Due to the filament evolving process inside the switching layer, most devices exhibit gradual reset but abrupt set, which limits the synaptic device to behave like an excitatory and inhibitory synapse at the same time.

2. Filament formation is a stochastic process and it is facilitated by the defects present in the switching layer. As the devices are scaled down to smaller dimensions, the defects required for filament formation

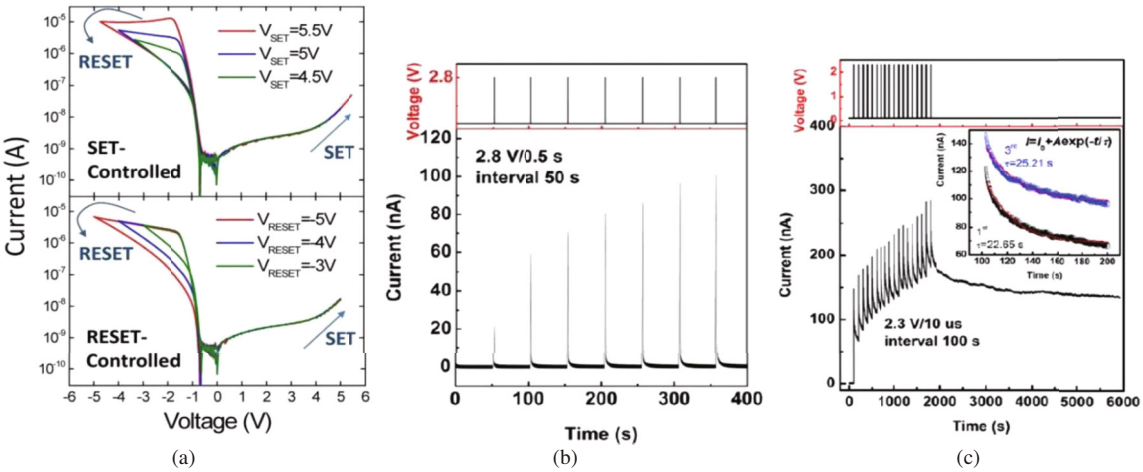


Figure 12 (Color online) Multilevel change in the conductance of devices. (a) *I*-*V* curves of the Ta/TaOx/TiO2/Ti device. The upper panel shows SET-controlled multilevel resistance switching in which SET voltage was gradually increased. The lower panel shows a RESET-controlled mode in which RESET voltage was increased gradually. Reprinted with permission from [65]. Copyright 2015, Nature Publishing Group. (b) Change in volatile conductance in the WO_{3-x} device after applying seven consecutive pulses of 2.8 V for 0.5 s with an interval of 50 s. The read voltage was 0.5 V. (c) Non-volatile conductance and subsequent state retention of the WO_{3-x} device after application of 2.3 V pulses for 10 μs. Current decay curves after the first and third pulses are shown in the inset. Reprinted with permission from [61]. Copyright 2013, IOP Publishing Limited.

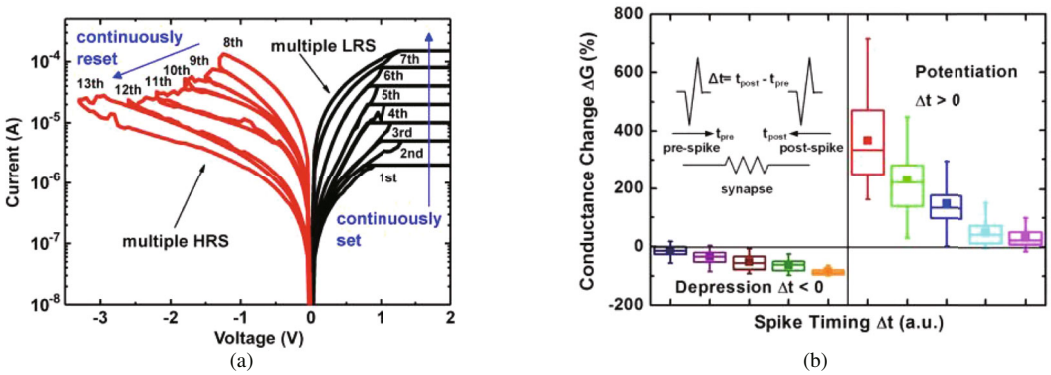
will decrease, increasing the randomness in the number of defects and resulting in large conductance fluctuations, especially in the low conductance regime.

Although the neuromorphic system is more fault-tolerant, a high degree of conductance fluctuation may still adversely affect computational accuracy of the system. Wang et al. [65] tried to address this problem by adopting non-filamentary switches, such as the Ta/TaO_x/TiO_2/Ti structure and argued that these devices can change conductance gradually because the resistance change in the devices was determined by a homogeneous barrier modulation (HBM) induced by oxygen ion motion. As shown in Figure 12(a), they observed multilevel set and reset in the device.

Multilevel resistance switching has been demonstrated in numerous types of valence change based devices [64, 66–68]. Yang et al. [61] have demonstrated synaptic behaviour of WO_{3-x} based devices. As shown in the case of biological synapses discussed in Section 5, they were able to realize rate based synaptic changes in the device. Upon application of low frequency pulse, the device displays short-term facilitation, as shown in Figure 12(b), and it shows long-term potentiation for high frequency pulses (see Figure 12(c)).

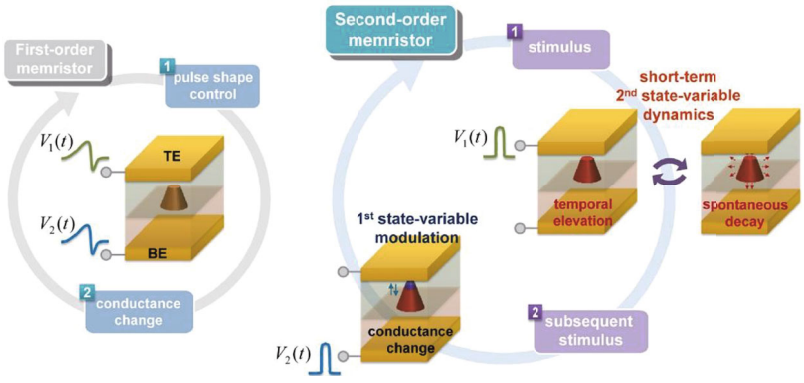
Yu et al. [66] have shown multilevel switching in TiN/HfO_x/AlO_x/Pt based resistive switching cells by controlling the compliance current and reset stop voltages, as shown in Figure 13(a). The property of gradual resistance change using voltage pulses was exploited to demonstrate the spike-timing-dependent plasticity (STDP) learning rule, as in Figure 13(b). They were able to achieve low (sub-Pico joule) energy consumption per operation, which is a critical parameter for using these devices in large scale neuromorphic systems.

For the purpose of neuromorphic computing, resistive change devices can be divided into two categories: *first order systems* and *second order systems*. In order to mimic synaptic learning rules, first order systems require engineering the shape of the applied spikes, such that the relative timing of the spikes from the pre and post-synaptic neurons can be encoded into effective programming pulse amplitude or duration resulting from the two overlapping spikes [55, 63, 69]. The reason for this requirement is the fact that the modulation of the memristor conductance, determined by an internal state-variable (e.g., conduction channel size w), is solely controlled by the input (e.g., the voltage V applied to the device) and the channel size. However, pulse modulation and precise timing control requires additional circuitry, making the system bulky and complicated.



(a) (b)

Figure 13 (Color online) Gradual resistance change in an RRAM device and demonstration of STDP characteristics by Yu et al. [66]. (a) *I-V* characteristics of the HfOx/AlOx synapse device. Continuous set states were obtained by consecutive increase in the value of current compliance from 1 μA to 200 μA, and continuous reset states were achieved using consecutive increase in the reset stop voltages from −1.3 V to −3.3 V. Multiple states were obtained due to continuous change in resistance. (b) Plot of percentage conductance change of the HfOx/AlOx synapse device as a function of the spike timing difference, showing an STDP-like curve. Reprinted with permission from [66]. Copyright 2011, IEEE.



(a) (b)

Figure 14 (Color online) Schematic diagram of the mechanism behind first and second-order memristors. (a) Operation of a first-order memristor, and (b) operation of a second-order memristor. Reprinted with permission from [71]. Copyright 2015, American Chemical Society.

To avoid this complexity and to emulate the synaptic behavior more directly without the strict requirement of overlapping pulses, Kim et al. [70, 71] used the so-called second-order memristors (Figure 14), where the conductance and the associated primary state variable are also regulated by a secondary state-variable (e.g., temperature of the channel T) in addition to the first state-variable (e.g., conduction channel size w), enabling the device to implement complex and potentially bio-realistic dynamic effects. An example of such a second-order memristor is mathematically described in [71] as

$$\frac{\mathrm{d}w}{\mathrm{d}t} = f(w, T, V, t).$$

In this device, the local temperature T near the filament increases due to Joule heating which affects the drift and diffusion processes of the conducting filaments (CF) and determines the growth and dissolution of the filaments. More importantly, the dynamics of T is a function of time and provides an internal timing mechanism; these internal dynamics in second-order memristors provide a possible solution to equip the memristors with timing-based synaptic plasticity effects, such as STDP.

6.2 Resistance switches based on electrochemical metallization

Resistance switches based on electrochemical metallization are normally called PMC, which is also known as solid electrolyte memory, nano-ionic resistive memory, electrochemical memory, or conductive bridging

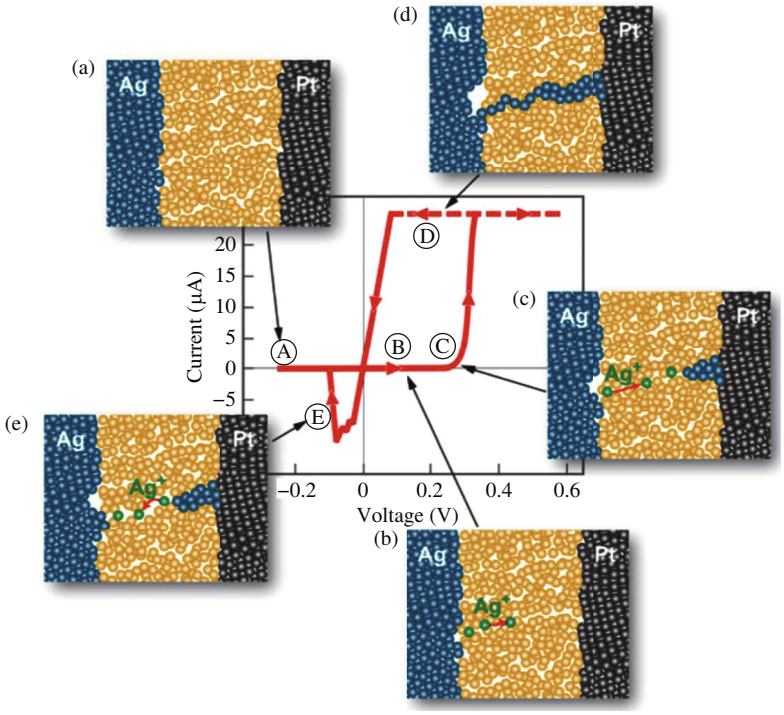


Figure 15 (Color online) Schematic of the operating principle of a PMC consisting of silver (Ag) and platinum (Pt) electrodes. (a) High resistance OFF state of the PMC. (b) Ag dissolves in the solid electrolyte as Ag+ and ions drift towards the Pt electrode when a high electric field is applied. (c) Ag+ ions reduce on the surface of the Pt electrode leading to electrodeposition of Ag, decreasing the resistance of the device. (d) Ag filament is formed and the device turns ON. (e) When electric field of opposite polarity is applied, Ag filament dissolves in the electrolyte and resets the device to the OFF state.

RAM. It is another promising non-volatile memory device [72], with a number of advantages, including low operation current and voltage, relatively high-speed switching, multiple resistance levels, high device yield, and relatively high endurance [73, 74]. Some early studies on these devices were reported on Ag doped Ag_2S_3 films by Hirose in 1976 [75]. Much more significant progress has been made in the past decade, including integration using an industry standard process [76].

Each programmable memory cell has a thin solid state electrolyte layer sandwiched between an inert cathode layer and an electrochemically active anode. Typically, the anode layer is Ag or Cu, and the electrolyte layer is a doped amorphous material that can conduct ions. When a positive bias is applied on the anode, redox reactions drive the metal ions from the anode into the electrolyte layer. The metal ions, drifting through the electrolyte layer, are electrochemically reduced and deposited on the cathode. A conducting bridge thus forms between the two electrodes, leading to the low resistance state, i.e., the ON state. The device can be switched OFF by applying an opposite voltage [77], causing a similar process where the cations migrate in the opposite direction. The above switching process is schematically shown in Figure 15 [72]. In other cases, especially when the dielectric medium is an oxide (e.g., HfO_x) instead of a traditional electrolyte material (e.g., GeS_x), the filament growth direction could be opposite to the schematic shown in Figure 15 [78].

The PMC device shares some similarities with the biological synapse in ionic transport at the physical level. Several materials have been used as electrolytes for switching operation [77] in PMCs, among which silver based chalcogenides have been the most popular one so far for synapse electronics. In 2011, Ohno et al. [79] demonstrated an Ag_2S inorganic synapse showing a transition from short-term to long-term potentiation by controlling the repetition time of the input pulses (see Figure 16). For a lower repetition rate, the device showed a temporary increase in conductance and the conductance then decayed spontaneously over time, whereas a persistent enhancement was achieved with a frequent repetition of the stimuli. In addition, the device mimicked the multistore model of human memory [80] where information

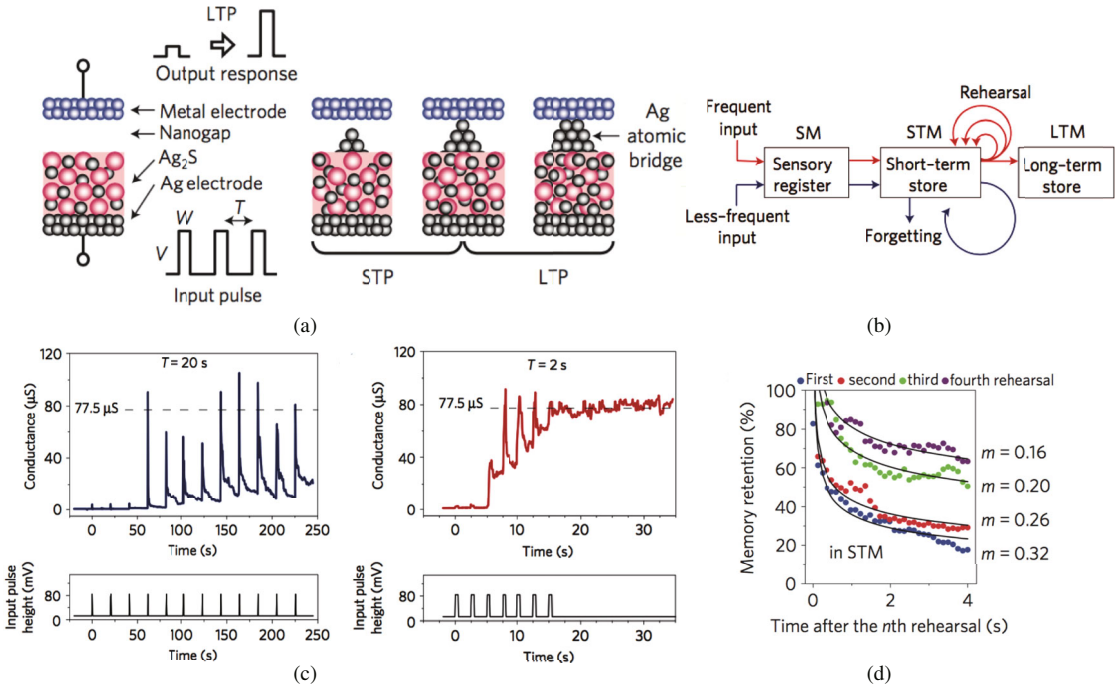


Figure 16 (Color online) Demonstration of a PMC as an electronic synapse by Ohno et al. [79]. (a) Schematic diagram of the cell consisting of a Ag electrode and Ag_2S electrolyte separated by a nanogap from the intert metal electrode. Also shown is the process of formation of a bridge between electrodes when repetitive voltage pulses are applied, leading to transition from short to long-term potentiation. (b) Block diagram of the multi-store model of human memory, showing that information is first stored in the sensory registor and then sent to a short-term memory where frequent rehearsal stores the information in the long-term store, whereas less frequent or no rehearsal results in forgetting. (c) Experimental results showing device conductance when a sequence of low frequency voltage pulses is applied across the device. The conductance increases with voltage but decreases when the voltage is removed, demonstrating short-term memory. At high frequency, the conductance increases and stays high at the value of 77.5 μS which is the conductance of a single filament channel, demonstrating long-term memory. (d) Plot of memory retention of the device as a function of time for one to four rehearsals. More rehearsals result in higher retention for the same amount of time elapsed. Reprinted with permission from [79]. Copyright 2011, Macmillan Publishers Ltd.

is transferred from sensory memory to the short-term memory, and finally some information becomes long-term through a repetition rehearsal process.

In 2010, Yu et al. [81, 82] modelled a CBRAM device showing that the resistance of the PMC can be tuned by gradually varying the pulse amplitude, and proposed a method to implement STDP, as shown in Figure 17. However, varying pulse amplitude requires additional overhead in the neuron circuitry because each neuron must generate pulses with increasing amplitude [83]. In 2014, Mahalanabis et al. [84] showed gradual change in PMC resistance by applying positive and negative voltage pulses with constant magnitude for program and erase (see Figure 17(d)) to achieve an average resistance change of 3% with each pulse operation, close to the DARPA goal of 1% change in weight by each operation.

Another class of PMC is Ag doped amorphous Si, with which Jo et al. [85] demonstrated STDP in 2010. With application of consecutive DC sweeps in this demonstration, they could gradually change the resistance for both set and reset process, as shown in Figure 18(a). They also implemented the synaptic learning rules LTP, LTD (Figure 18(b)) and STDP (Figure 18(c)) with these devices. LTP and LTD were demonstrated by applying consecutive potentiating or depressing pulses while STDP was observed upon application of two programming pulses on the two terminals of the device. The same device was used as a synapse in [86]. Neurons were implemented using CMOS circuits.

There have been reports of the possibility of biological synapses being binary [87, 88]. Such synapses can be emulated using cation-based programmable memory cells with sharp turn ON/OFF switching characteristics. Here the switching states are digital in nature, and therefore the probability of switching can be used as the weighting parameter [89]. Synapses emulated using such devices are expected to be

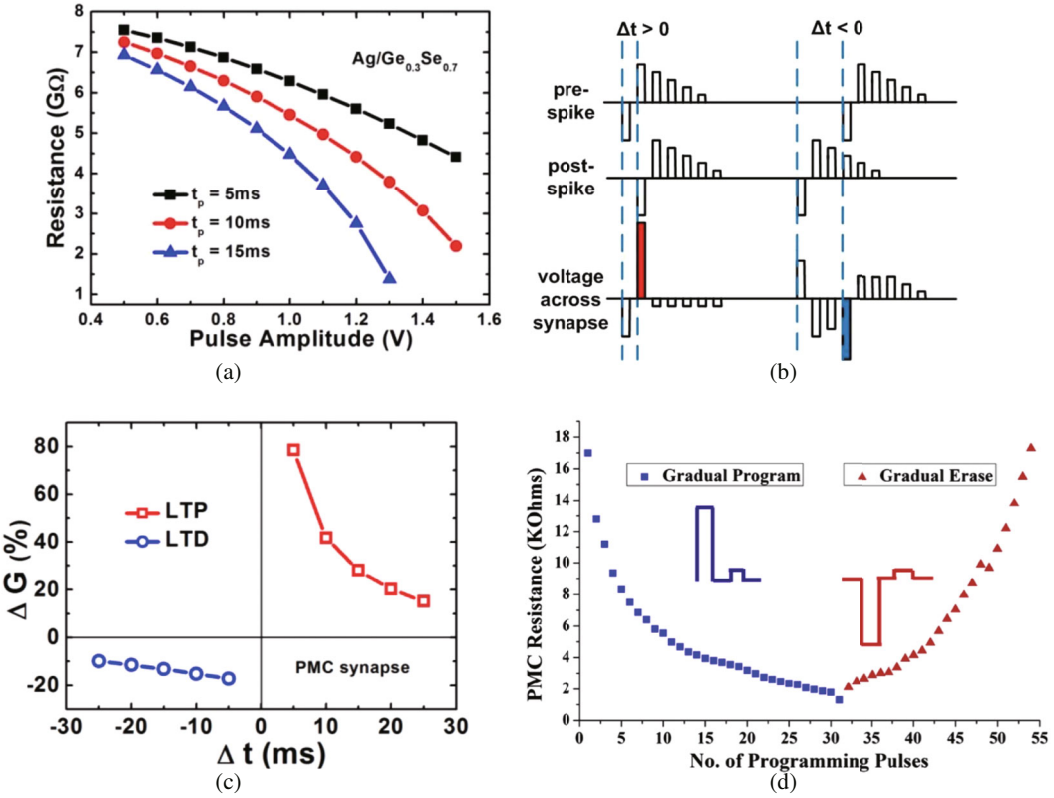


Figure 17 (Color online) Gradual tuning of the resistance of a PMC device, and simulated STDP characteristics. (a) Simulated change in the off-state resistance of the $Ag/Ge_{0.3}Se_{0.7}$ device in [81] using varying pulse amplitudes. The pulse width is denoted by t_p. (b) Programming scheme used in [81] to investigate the STDP properties of an $Ag/Ge_{0.3}Se_{0.7}$ device. The difference in the timing between the pre and post-spikes is converted to varying pulse amplitudes across the device. The result is that when pre preceeds post, the device is set, whereas it is reset if post preceeds pre. (c) Simulated STDP response of the $Ag/Ge_{0.3}Se_{0.7}$ device using the pulse scheme in (c), and showing LTP and LTD. Reprinted with permission from [81]. Copyright 2010, IEEE. (d) Programming of an Ag-Ge-Se device by Mahalanabis et al. [84] using multiple 100 µs, 1.5 V write pulses and 200 µs, 1.5 V erase pulses. The resistance was read using a 100 mV read pulse. The device shows gradual resistance change of one order of magnitude on applying $\sim$ 30 set and reset pulses. Reprinted with permission from [84]. Copyright 2014, Elsevier Ltd.

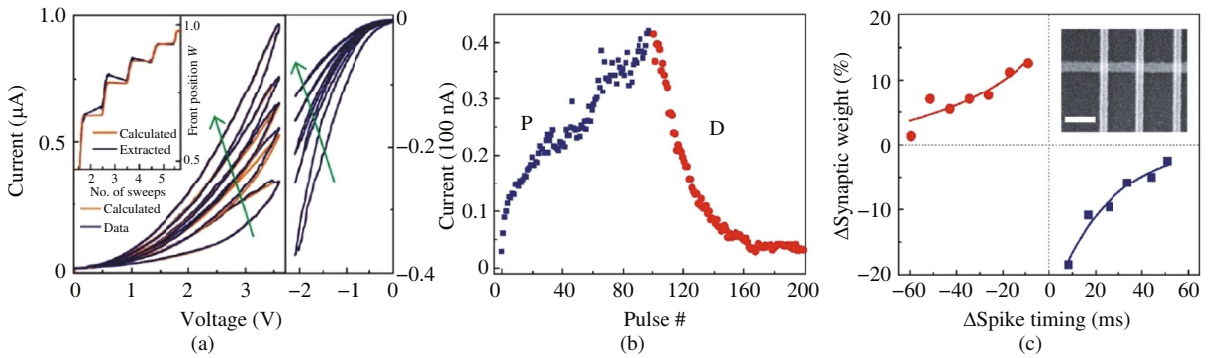
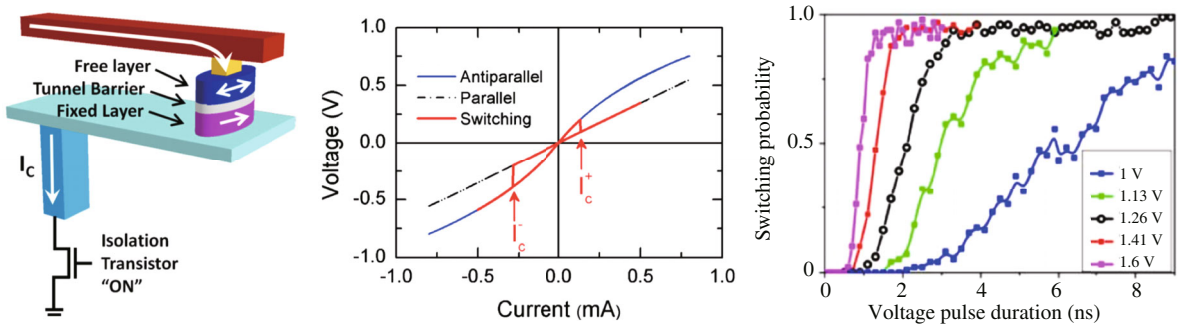


Figure 18 (Color online) Characteristics of the Ag-doped Si based PMC device and the demonstration of spike-timing dependent plasticity (STDP) in [85]. (a) Plot of the measured (blue lines) and calculated (orange lines) $I(V)$ characteristics. The conductance of the device increases with the number of positive sweeps and decreases with increasing number of negative sweeps. The inset shows the simulated and extracted values of the position of the Ag front with increasing number of positive sweeps. (b) Plot of the device current response to positive (3.2 V, 300 µs) and negative (−2.8 V, 300 µs) programming pulses showing incremental increase (potentiation) and decrease (depression) of device conductance. (c) STDP demonstration showing normalized change in the synaptic weight i.e., conductance of the cell vs. the timing ($\Delta\tau$) between the pre and post-spikes. Inset shows a scanning electron microscope of the 100 nm × 100 nm crossbar device. Reprinted with permission from [85]. Copyright 2010, American Chemical Society.



(a) (b) (c)

Figure 19 (Color online) Structure and properties of the magnetic tunnel junction (MTJ) device. (a) Schematic diagram of the basic structure showing different layers in the MTJ. The arrows in the free and fixed layers indicate the direction of the magnetization. Reprinted with permission from [96]. Copyright 2013, IOP Publishing. (b) $I(V)$ characteristics of the device when the magnetizations are parallel (dashed black) and antiparallel (solid blue), and the switching between the states (solid red). Reprinted with permission from [102]. Copyright 2011, American Institute of Physics. (c) Plot of the probability of the switching from one state to another as a function of the pulse duration and voltage amplitude. Higher voltages result in greater probability of switching for shorter pulses. Reprinted with permission from [101]. Copyright 2014, IEEE.

more reliable and resistant to noise compared to electronic synapses made using analog memristors.

Several other reports have investigated the capability of PMC and its variations for emulating the synapse [90, 91]. The conducting bridge in the PMC can be as narrow as a few nanometers, which indicates a good device scalability and cell density [92]. Programming currents and voltages have been shown to be as low as 1 μA and a few hundred mV [74], respectively. The turn ON time can be as fast as 10 ns [93], resulting in a low-energy ($\sim$ 10 fJ) operation. By controlling the programming current of an individual cell (if it can be done in a large array as well), the device can be set at different resistance states giving it multivalued conductance characteristics [94]. All these characteristics make them a promising candidate as synapse in neuromorphic circuits.

6.3 Spin transfer torque MRAM

Magnetoresistive Random Access Memory (MRAM) is another emerging non-volatile memory that has relatively high read and write speeds, and essentially unlimited endurance [95, 96].

STT MRAM uses the electron spin to switch the MRAM which consists of a thin insulating layer (tunnel barrier) sandwiched between a fixed magnetic layer and a free magnetic layer. Passing a current through the fixed layer produces a spin-polarized current which can change the magnetization orientation of the free layer. When the magnetization orientations of the free and fixed layers are parallel, the device is in a low resistance state. When antiparallel, the device is in the high resistance state [97, 98]. The switching is however stochastic in nature [99, 100], meaning that the time required for programming the device is somewhat random, requiring longer write times to ensure switching. This switching randomness can be utilized to implement synapse behaviour [101]. If a programming pulse of a duration Δt is applied to the device, the probability of switching is exponential in form,

$$P = 1 - \mathrm{e}^{-\Delta t/\langle\tau\rangle}.$$

Here $\langle\tau\rangle$ is the mean switching time, which can be controlled over several decades by adjusting the current. Therefore, both pulse duration and current can be tuned to adjust the switching probability, as shown in Figure 19 [96, 101, 102], which is similar to other memristors in this regard and can be employed for synaptic functions. Symmetry in the voltages used for P $\rightarrow$ AP and AP $\rightarrow$ P transitions means simple pulse waveforms, an advantage over other devices that are usually asymmetric. Using these properties of the STT MRAM, Vincent et al. [101] simulated a spiking neural network to implement unsupervised learning using the STDP rule to count vehicles in a six-lane freeway. The detection rate was 99.0% for the four inward lanes, whereas it was 76.3% and 54.7% for the two outward lanes.

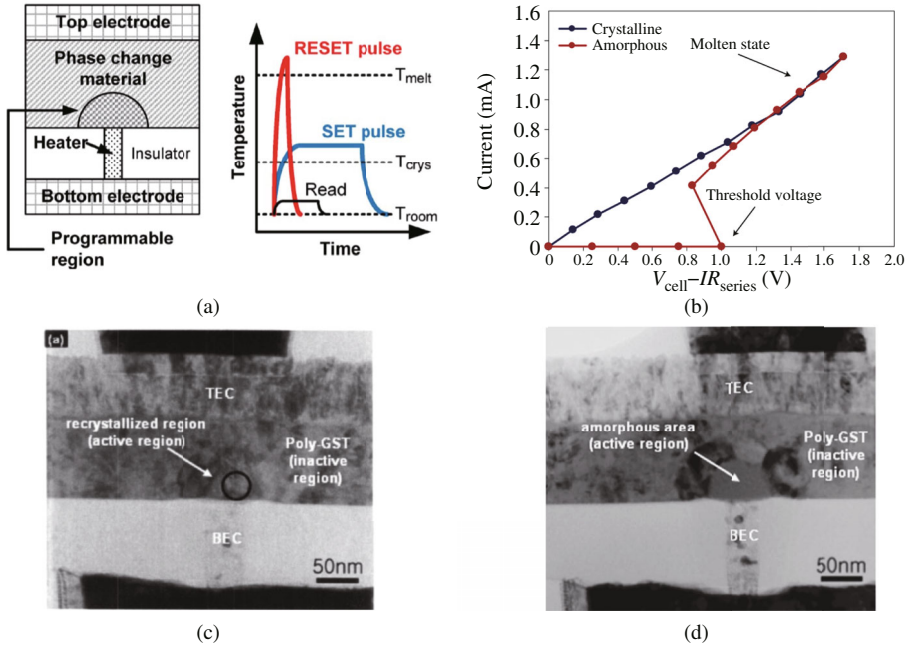


Figure 20 (Color online) Structure and characteristics of phase change memory (PCM) devices. (a) Schematic of the cross-section of the PCM cell consisting of a phase change material sandwiched between the bottom and top electrodes. The region in the phase change material near the heater layer is the programmable region, also called the mushroom cell. A short high voltge reset pulse changes the temperature of the material rapidly to convert the crystalline region to amorphous phase, which RESETs the device. A long low voltage pulse is used to crystallize the amorphous phase to SET the device. Reprinted with permission from [106]. Copyright 2010, IEEE. (b) $I(V)$ characteristics of a typical PCM cell showing the increase in device current at the threshold voltage due to the transition from amorphous to crystalline phase. Reprinted with permission from [108]. Copyright 2005, Nature Publishing Group. Cross-sectional transmission electron microscopy images of a GST-based PCM cell showing the active region in (c) the crystalline phase, and (d) amorphous phase. Reprinted with permission from [109]. Copyright 2007, The Electrochemical Society.

Because the switching pulse can be as short as a few ns and the current is around tens of μA at voltages < 1 V, the energy consumption for each operation can be in the range of 100 fJ. Although still higher than PMCs, some new device designs have been recently proposed to further reduce the energy consumption [103]. The devices can be scaled down to a certain degree, giving a high packing density. However, thermal fluctuations can make the magnetization unstable for small sizes. This limitation in device scaling and the corresponding ultimate packing density is called the super paramagnetic limit [104]. In addition, the energy required for switching the device does not proportionally scale with the device area if the resistance is kept constant. Despite these issues, STT MRAMs are interesting candidates for synaptic electronics due to a few attractive properties, especially the nearly unlimited endurance. Experimental realization of STDP using these devices is yet to be demonstrated.

6.4 Phase change memory

Phase change memory (PCM) was developed in the 1960s [105] and has been under development over the last several decades [46]. Among the emerging non-volatile memory technologies, PCM is likely the most mature one for a large scale product so far. It is non-volatile with a large ON/OFF ratio and relatively long endurance [106]. PCM typically uses a chalcogenide that can be electrically switched from amorphous (high resistance, reset) phase to crystalline (low resistance, set) phase and vice versa [107]. The transition to the amorphous phase is achieved by applying a large but short voltage pulse, whereas amorphous to crystalline change is achieved using a long but medium level current pulse to anneal the material above its crystalline temperature, as shown in Figure 20 [106, 108, 109]. Typical reset pulse durations are of the order of ns, whereas set pulse durations are in the range of μs, although recently sub-ns set switching speeds have been demonstrated [110].

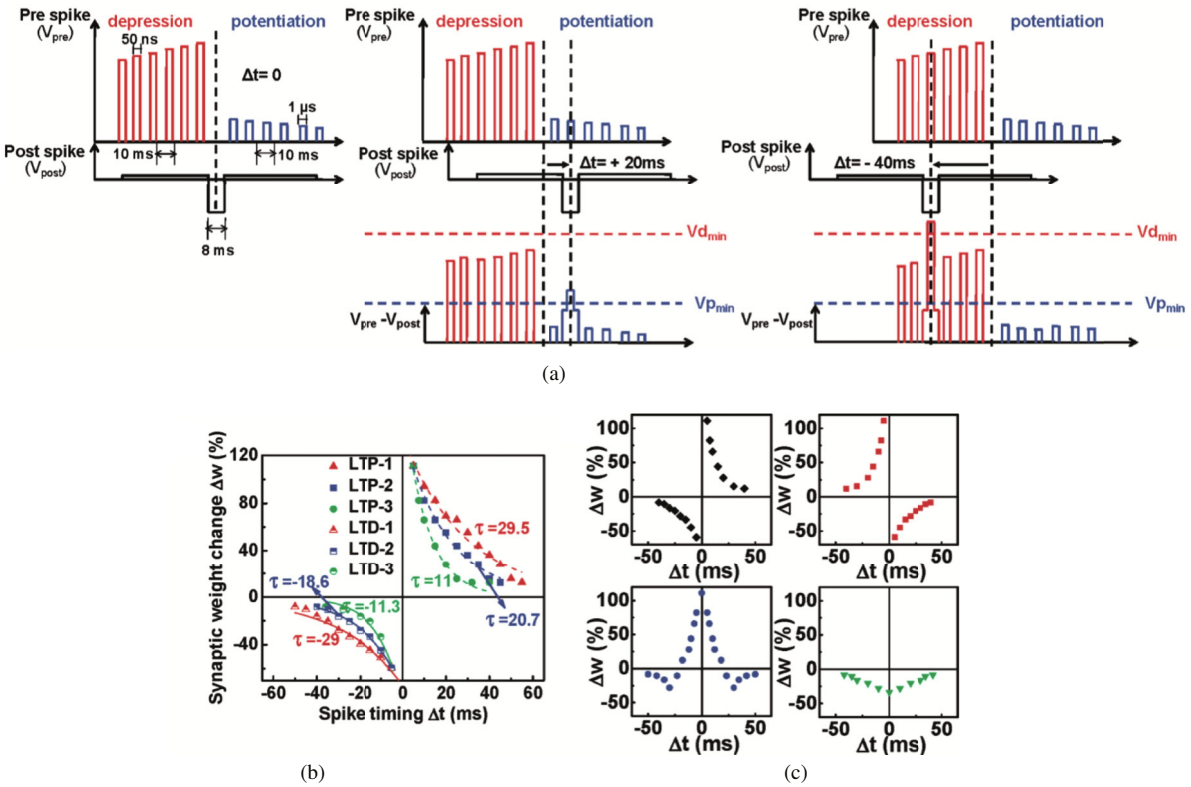


Figure 21 (Color online) Implementation of STDP using phase change memory devices. (a) Programming scheme used in [47] to implement STDP. The pre-spike consists of a sequence of pulses of increasing and decreasing amplitudes. The post-spike consists of a single pulse. A Δt of 20 ms sets the device whereas for $\Delta t = -40$ ms, the device is reset. (b) STDP characteristics of the synaptic device for different prespike configurations (LTP1, LTP2, and LTP3) and different post-spike configurations (LTD1, LTD2, LTD3) resulting in varying potentiation and depression time constants. (c) Demonstration of different types of STDP learning rules: asymmetric (upper left, positive Δt potentiation), asymmetric (upper right, negative Δt potentiation), symmetric (bottom left, potentiation), and symmetric (bottom right, depression). Reprinted with permission from [47]. Copyright 2010, American Chemical Society.

Several groups have reported experiments using PCM devices for synaptic applications [111]. Using GST based PCM cells, Kuzum et al. [47] obtained multiple intermediate resistance levels between set and reset, and achieved an order of magnitude change in the resistance through 100 voltage steps. They used a pulsing scheme, shown in Figure 21, similar to their gradual set/reset experiments in order to induce the desired STDP characteristics. The response agreed with the biological data measured in hippocampal glutamatergic synapses. They achieved control of the STDP time constant by adjusting the amplitude and time spacing between individual pulses in the spikes, and implemented both asymmetric and symmetric STDP schemes. The reset operation required 50 pJ of energy, while the set operation required a much higher energy, although using set current data extrapolated from device scaling they predicted a lower energy consumption of 0.675 pJ for smaller device sizes.

Suri et al. [112] utilized the nucleation dominated behaviour in GST to achieve a gradual increase in conductance with multiple identical voltage pulses, and compared it with GeTe which, due to growth dominated behaviour, shows an abrupt conductance increase. They noticed that in contrast to the conductance increase operations (potentiation) which corresponds to the crystallization process of the PCM, identical pulses did not lead to a gradual decrease in conductance (depression) and varying voltage pulses are needed for this purpose. To avoid using varying voltage pulses, they proposed a 2-PCM synapse consisting of a first PCM that is managed to contribute negatively to the output neuron (equivalent to LTD) through circuit design, and the second PCM that contributes positively to the output (equivalent to LTP). In this 2-PCM synapse, only crystallization operation is needed to realize either LTD or LTP, circumventing the abrupt reset characteristic and potentially resulting in a lower energy consumption per

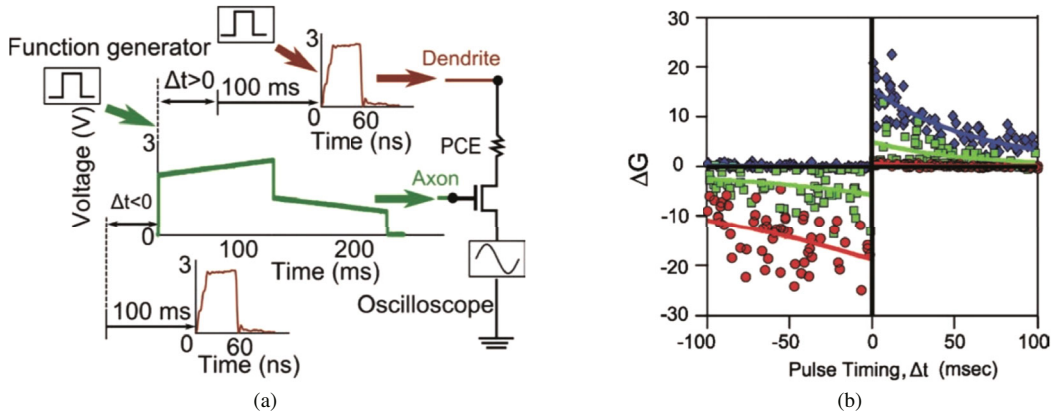


Figure 22 (Color online) STDP demonstration and energy consumption in 40 nm mushroom PCM devices. (a) Pulsing scheme used for the implementation of STDP. An asymmetric 200 ms pulse is applied to the gate terminal (axon) of the transistor followed by a 60 ns pulse to the PCM terminal (dendrite), 100 ms later. The asymmetry in the gating pulse is such that for positive Δt the conductance of the PCM increases, and it decreases for $\Delta t < 0$. (b) STDP response of the device from 1000 pulses with random pulse timings. The change in conductance (ΔG) is the ratio of the difference in final and initial conductance to the minimum of the two. The colors correspond to different initial conductance values (blue diamonds: < 0.5 μS, green squares: > 0.5 μs and < 5 μS, red circles: > 5 μS). Reprinted with permission from [113]. Copyright 2013, ACM.

synaptic event.

Jackson et al. [113] also implemented STDP using GST PCM cells that were pore and mushroom shaped, but used a field effect transistor (FET) in series with the PCM cell. In this implementation (shown in Figure 22), a FET (axon terminal) was gated using a long asymmetric pulse and short pulses were applied to the PCM (dendrite) to melt-quench and recrystallize the GST for depression and potentiation, respectively. The timing of the pulses with respect to the gate pulse resulted in an STDP-like response. The energy consumed in the depression and facilitation processes was 58 pJ and 18 pJ, respectively. They also proposed a method using RC delay as the timer that negates the need of an access device and allows for more frequent synaptic updates, exceeding 100 KHz. Fast synapses could be valuable for high speed detection systems. Other recent reports of using PCM based devices [114] focus on implementing grids and reducing the energy consumption.

6.5 Silicon-based devices

As Moore's law continued its streak [115], CMOS devices have become smaller and smaller in size, consume significantly less power compared to a few decades back.

The idea of using CMOS devices for neuromorphic computing developed in the 1980s, when Carver Mead proposed the use of floating gate memory, resistor networks, and amplifiers to implement analog computing using neuromorphic systems [116]. He used two examples, i.e., retinal computing with automatic gain control and an adaptive retina, to introduce the principles behind neural systems. Following this, Mead and others developed nFET and pFET based non-volatile synapse transistors with simultaneous memory reading and writing capability [117]. The synapses computed the product of stored memory (weight) and the input voltage to implement learning functions, like the biological synapse. They demonstrated 2×2 arrays and showed that they can address individual array nodes with a good selectivity. Since then, a number of groups have worked on CMOS based networks for synaptic circuits [118, 119].

More recently, Bartolozzi and Indiveri have described analog VLSI synaptic circuits that could be integrated into large VLSI spike-based neural systems [120]. They presented experimental data for a diff-pair integrator (DPI) device showing realistic dynamics and how it could be connected to multiple modules for implementing other synaptic functions and recent biological findings, such as short and long-term depression and potentiation.

However, CMOS and transistor based circuits still face scaling limits [121]. Because it was envisioned that current technology is difficult to extend below 10 nm gate length, Likharev developed the concept

of CMOL circuits [122–124], a hybrid CMOS/nanowire/MOLecular device that would combine a level of advanced CMOS devices fabricated by lithographic patterning, layers of parallel nanowire arrays, and a level of self-assembled [125] molecular devices. Once implemented to its full capacity, the device density in CMOL circuits can be as high as 10^{12} cm^{-2} and may provide up to 10^{20} operations per cm^2 per second with manageable power consumption [126]. Self-assembly of devices is extremely challenging due to poor yield issues, and is still an area of active research.

7 Realization of the artificial neuron

We can identify three generations of artificial neural networks: the first generation uses the McCulloch-Pitts neurons, also known as a perceptron or a threshold-gate, as the basic unit of computation. The multilayer perceptron, which is a first generation model, uses binary or bipolar digital input and output, and with a single hidden layer can be used to compute any Boolean function[1)]. The second generation uses an activation function, such as a sigmoid or the hyperbolic tangent, as its basis computational unit (neuron). These neural networks can also compute any Boolean function by using a threshold, and can approximate any continuous analog function with one hidden layer. They also support learning algorithms based on gradient descent, such as error back-propagation [127].

The third generation of artificial neural networks is based on spiking neural networks (SNN), which are inspired by neurophysiology and use temporal coding to pass information between neurons. Like second generation neurons, they can also approximate continuous functions by using temporally encoded inputs and outputs [128]. In the following paragraph we will discuss two of the most widely used spiking neuron models:

(a) *The leaky integrate-and-fire* (*LIF*) *neuron*: This is one of the simplest spiking neuron models, but it is very popular due to the ease with which it can be analysed and simulated. The neuron is modelled as a leaky integrator of its input current $I(t)$:

$$\tau_m \frac{\mathrm{d}v}{\mathrm{d}t} = -v(t) + RI(t).$$

Here $v(t)$ represents the membrane potential at time t, τ_m is the membrane time-constant and R is the membrane resistance. This equation describes a simple resistor-capacitor (RC) circuit where the leakage term is due to the resistor and the integration of $I(t)$ is due to the capacitor that is in parallel with the resistor.

The LIF neuron model, although very simple, is not biologically plausible and cannot explain actual spike generation. The Hodgkin-Huxley model, proposed by Alan Hodgkin and Andrew Huxley in 1952, has a solid foundation in physiology.

(b) *Hodgkin-Huxley Model* [129]: The Hodgkin-Huxley model, shown in Figure 23, for neural dynamics is one of the most successful models in computational neuroscience. Based on voltage-clamp experiments on the giant squid axon, the model incorporates voltage-sensitive ion channels into the circuit model of the membrane to describe the generation and propagation of action potentials. Mathematically, the current flowing through the membrane (I_c) can be expressed as

$$I_c = C_m \frac{\mathrm{d}V_m}{\mathrm{d}t},$$

and the current through a given ion channel is the product:

$$I_i = g_i(V_m - V_i),$$

where V_i is the reversal potential of the ith ion channel. For a cell with sodium and potassium channels, the total current through the membrane is given by

$$I = C_m \frac{dV_m}{\mathrm{d}t} + g_K(V_m - V_K) + g_{Na}(V_m - V_{Na}) + g_l(V_m - V_l),$$

1) http://www.cs.cornell.edu/courses/cs4700/2011fa/lectures/13_Ann.pdf. 17.

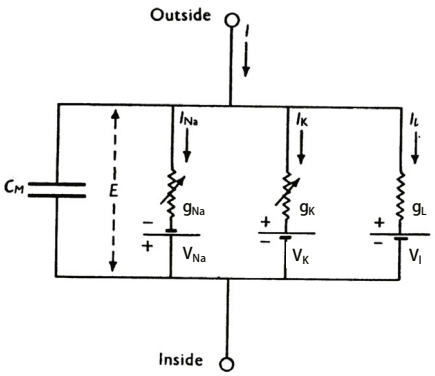


Figure 23 Equivalent circuit diagram representation of the Hodgkin-Huxley model. The model represents the lipid bilayer membrane between the interior and exterior of the nerve cell. Adapted from [129]. Copyright 1952, The Physiological Society.

where,

I: total membrane current per unit area,

C_m: the membrane capacitance per unit area,

g_K, g_{Na}, g_l : the potassium, sodium and leakage conductance per unit area, respectively,

V_K, V_{Na}, V_l: the potassium, sodium and leakage reversal potentials, respectively.

Electronic neuron implementation using memristors: Neuristor. Traditionally electronic neurons were implemented either using a computer program [130–132] or using specialized analogue MOSFET circuits [133–136]. There was always a need of constructing a physical system that can mimic biological functionality more directly, with the goal of improving efficiency and scale. If one wants to realize an entire neuromorphic system on a single chip, scaling of traditional hardware neurons will prove to be a bottleneck, even though memristor based synapses will be scalable. In 2013, HP Labs proposed a scalable neuron built using two memristors [137]. A neuromorphic system having both the neurons and synapses implemented using memristors can help to achieve a massive parallel and an extremely compact computational system.

Figure 24 depicts the neuristor circuit, which closely mimics the Hodgkin and Huxley model, implemented using Mott memristors. A Mott memristor is a type of threshold switching device made with an insulating-metal transition (IMT) material. Using two identical Mott memristors, each having a parallel capacitor attached to it (C_1 and C_2), the device is powered by two DC sources with opposite polarities, which represent the sodium and potassium channels of the neuron. Similar to Hodgkin & Huxley neuron model this circuit can be modelled using four state variables. Figure 24(b) shows the bistable current-voltage characteristics of the memristors used to construct the neuristor.

8 Concluding remarks and future outlook

Although the idea of mimicking the biological brain using electronic circuits has been around for decades, recent developments in the field of microelectronics and emerging nanoelectronics seem to have brought this dream closer to reality. The ultimate goal is to conquer the last computing frontier, i.e., building a brain-like computer that consumes low power, has high computational density per unit volume, and can perform complex tasks at high-speed. This is obviously still far from realization, primarily due to our still very limited knowledge of the intelligent biosystems and the immaturity of emerging devices.

It is reasonable to assume that it will take decades or longer for us to significantly improve our knowledge about the human brain. Understanding the working principle of neurons and synapses is one of the most important tasks in order to capture some of the functionalities vital to truly mimicking them on a circuit. However, the development of artificial intelligent systems does not have to stop or wait. In contrast, the exploration of artificial intelligent systems can benefit the understanding of their bio-counterparts by providing feedbacks based on the emulating results.

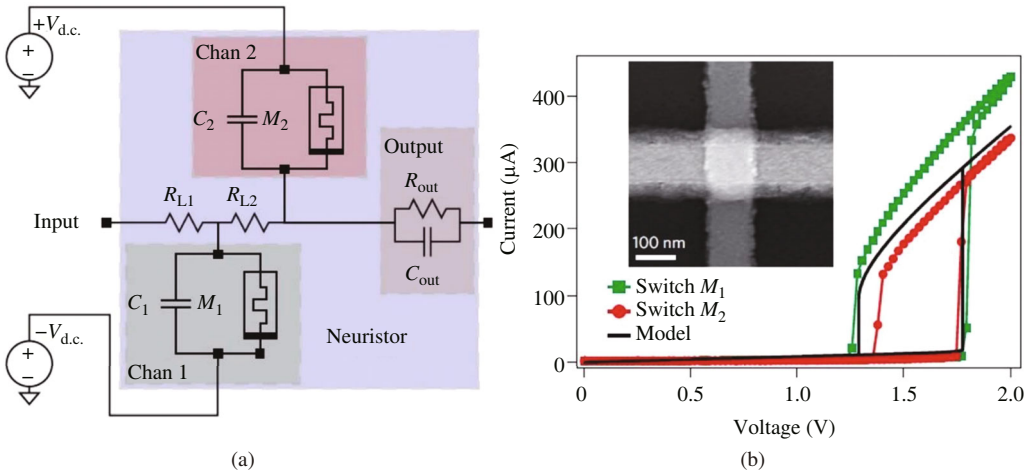


Figure 24 (Color online) Electronic implementation of the neuron using a neuristor device. (a) Circuit diagram of the neuristor, (b) The I-V curves of the two niobium dioxide crosspoint devices. The green and red curves correspond to the experimental I-V characteristics of devices M_1 and M_2, respectively. The black curve represents the quasi-static simulation result obtained using the model. The scanning electron microscope image of a typical device is shown in the inset. Reprinted with permission from [137]. Copyright 2013, Macmillan Publishers Limited.

On the other hand, substantial improvements are still needed in order to provide building blocks for the artificial intelligent systems, including devices and circuits that can emulate synapses and neurons in the brain with characteristics such as low-energy, scalability, integratability, and dynamics. We have discussed some of these devices which are suitable for mimicking synapses, such as resistive change memory, programmable memory cells, MRAM devices, phase change memory, etc. These emerging devices have demonstrated significant potential in this direction, and can emulate some of the crucial functionalities of synapses such as LTP, LTD and STDP. However, most of the demonstrations so far were realized by engineering voltage pulses at the expense of circuit and operation complexity. For instance, sparse spikes without overlapping are known to change synaptic weight in a bio-synapse, while majority of the STDP demonstrations made till date using emerging devices require overlapping spikes. This is due to the lack of appropriate dynamics in those devices, which is extremely important for the bio-system. Therefore, developing nanodevices with some intrinsic timing mechanism to implement the dynamics of bio-systems is critical for the future research.

Compared to synaptic devices, research efforts into devices that can mimic neural functionalities have lagged behind. Neuristors built with two memristors have shown some neural function in generating electrical pulses, which is an encouraging step towards implementing an all-memristor neuromorphic system. In addition, more efforts are needed towards building and testing integrated systems using emerging devices [138].

Before integrated systems using emerging devices are available for neuromorphic function demonstrations at a system level [139], those purely based on traditional CMOS devices will prove very useful to implement and test a variety of algorithms that can later be transferred to the low-energy and high-density intelligent systems based on emerging devices. In order to realize such intelligent systems, an interdisciplinary approach is required with efforts from researchers working in diverse areas such as biology, physics, materials engineering, electrical engineering, and computer science.

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Turing A. On computable numbers, with an application to the entscheidungs problem. Proc London Math Soc, 1936, 42: 230–265

2 von Neumann J. First draft of a report on the EDVAC. IEEE Ann Hist Comput, 1993, 15: 11–21

3 Turing A. Intelligent machinery. In: Copeland B J, ed. The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life plus The Secrets of Enigma. New York: Oxford University

Press, 2004. 406–443
4 Anderson H C. Neural network machines. IEEE Potentials, 1989, 8: 13–16
5 Squire L R, Berg D, Bloom F, et al. Fundamental neuroscience. Curr Opin Neurobiol, 2008, 10: 649–654
6 Kandel E R, Schwartz J H, Jessell T M. Principles of Neural Science. 4th ed. New York: McGraw-Hill Medical, 2000
7 Bennett M V L, Zukin R S. Electrical coupling and neuronal synchronization in the mammalian brain. Neuron, 2004, 41: 495–511
8 Zamarreño-Ramos C, Camuñas-Mesa L A, Pérez-Carrasco J A, et al. On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex. Front Neurosci, 2011, 5: 1–22
9 Pereda A E. Electrical synapses and their functional interactions with chemical synapses. Nat Rev Neurosci, 2014, 15: 250–63
10 Noback C R, Ruggiero D A, Demarest R J, The Human Nervous System: Structure and Function, 6th ed. Totowa: Humana Press, 2005
11 Versace M, Chandler B. MoNETA: a mind made from memristors. IEEE Spectr, 2010. http://spectrum.ieee.org/robotics/artificial-intelligence/moneta-a-mind-made-from-memristors
12 Chua L, Adamatzky A. Memristor Networks. Switzerland: Springer International Publishing, 2013
13 Mead C. Neuromorphic electronic systems. Proc IEEE, 1990, 78: 1629–1636
14 Diorio C, Hasler P, Minch B A, et al. Single-transistor silicon synapse. IEEE Trans Electron Dev, 1996, 43: 1972–1980
15 Wong H-S P, Raoux S, Kim S, et al. Phase change memory. Proc IEEE, 2010, 98: 2201–2227
16 Waser R, Aono M. Nanoionics-based resistive switching memories. Nat Mater, 2007, 6: 833–840
17 Versace M, Chandler B. The brain of a new machine. IEEE Spectr, 2010, 47: 30–37
18 Snider G. Amerson R, Carter D, et al. From synapses to circuitry: using memristive memory to explore the electronic brain. Computer, 2011, 44: 21–28
19 Hylton T. DARPA SyNAPSE Project. Arlington, 2009
20 Ananthanarayanan R, Esser S K, Simon H D, et al. The cat is out of the bag: cortical simulations with 109 neurons, 1013 synapses. In: Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, Portland, 2009. 1–12
21 Merolla P A, Arthur J V, Alvarez-Icaza R, et al. A million spiking-neuron integrated circuit with a scalable communication network and interface. Science, 2014, 345: 668–673
22 Furber S B, Lester D R, Plana L A, et al. Overview of the SpiNNaker system architecture. IEEE Trans Comput, 2013, 62: 2454–2467
23 Markram H. The Blue Brain Project. Nat Rev Neurosci, 2006, 7: 153–160
24 Schemmel J, Grubl A, Hartmann S, et al. Live demonstration: a scaled-down version of the BrainScaleS wafer-scale neuromorphic system. In: Proceedings of 2012 IEEE International Symposium on Circuits and Systems, Seoul, 2012. 702
25 Boahen K. Neurogrid: Emulating a Million Neurons in the Cortex. In: Proceedings of 28th IEEE Engineering in Medicine and Biology Society Annual International Conference, New York, 2006. Supp: 6702
26 Benjamin B V, Gao P, McQuinn E, et al. Neurogrid: a mixed-analog-digital multichip system for large-scale neural simulations. Proc IEEE, 2014, 102: 699–716
27 Hebb D O. The first stage of perception: growth of the assembly. In: The Organization of Behavior. Hoboken: John Wiley & Sons Inc., 1949. 60–78
28 Markram H, Gerstner W, Sjöström P J. A history of spike-timing-dependent plasticity. Front Synaptic Neurosci, 2011, 3: 1–24
29 Markram H, Lübke J, Frotscher M, et al. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science, 1997, 275: 213–215
30 Levy W B, Steward O. Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. Neuroscience, 1983, 8: 791–797
31 Cooper L N, Bear M F. The BCM theory of synapse modification at 30: interaction of theory with experiment. Nat Rev Neurosci, 2012, 13: 798–810
32 Bi G Q, Poo M M. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J Neurosci, 1998, 18: 10464–10472
33 Bienenstock E L, Cooper L N, Munro P W. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J Neurosci, 1982, 2: 32–48
34 Sejnowski T, Chattarji S, Sfanton P. Induction of synaptic plasticity by hebbian covariance in the hippocampus. In: The Computing Neuron. Boston: Addison-Wesley Longman Publishing Co., 1989. 105–124
35 Lynch M A. Long-term potentiation and memory. Physiol Rev, 2004, 84: 87–136
36 Bliss T V P, Lomo T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J Physiol, 1973, 232: 331–356
37 Mulkey R, Herron C, Malenka R. An essential role for protein phosphatases in hippocampal long-term depression. Science, 1993, 261: 1051–1055
38 Sjostrom P J, Gerstner W. Spike-timing-dependent plasticity. Scholarpedia, 2010, 5: 1362
39 Gütig R, Aharonov R, Rotter S, et al. Learning input correlations through nonlinear temporally asymmetric Hebbian

plasticity. J Neurosci, 2003, 23: 3697–3714
40 Rubin J, Lee D D, Sompolinsky H. Equilibrium properties of temporally asymmetric Hebbian plasticity. Physl Rev Lett, 2001, 86: 364–367
41 van Rossum M C, Bi G Q, Turrigiano G G. Stable Hebbian learning from spike timing-dependent plasticity. J Neurosci, 2000, 20: 8812–8821
42 Purves D, Augustine G J, Fitzpatrick D, et al. Neuroscience. 2nd ed. Sunderland: Sinauer Associates, 2001
43 Lee M-J, Lee C B, Lee D, et al. A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta2O5-x/TaO2-x bilayer structures. Nat Mater, 2011, 10: 625–630
44 Chanthbouala A, Garcia V, Cherifi R O, et al. A ferroelectric memristor. Nat Mater, 2012, 11: 860–864
45 Yang J J, Pickett M D, Li X, et al. Memristive switching mechanism for metal/oxide/metal nanodevices. Nat Nanotechnol, 2008, 3: 429–433
46 Wuttig M, Yamada N. Phase-change materials for rewriteable data storage. Nat Mater, 2007, 6: 824–832
47 Kuzum D, Jeyasingh R G D, Lee B, et al. Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. Nano Lett, 2012, 12: 2179–2186
48 Yang J J, Strukov D B, Stewart D R. Memristive devices for computing. Nat Nanotechnol, 2013, 8: 13–24
49 Strukov D B, Snider G S, Stewart D R, et al. The missing memristor found. Nature, 2008, 453: 80–83
50 Chua L O, Kang S M. Memristive devices and systems. Proc IEEE, 1976, 64: 209–223
51 Hickmott T W, Hiatt W R. Bistable switching in Niobium oxide diodes. Appl Phys Lett, 1965, 6: 106–108
52 Hickmott T W. Low-frequency negative resistance in thin anodic oxide films. J Appl Phys, 1962, 33: 2669
53 Chua L. Resistance switching memories are memristors. Appl Phys A-Mater Sci Process, 2011, 102: 765–783
54 Rajendran B, Liu Y, Seo J S, et al. Specifications of nanoscale devices and circuits for neuromorphic computational systems. IEEE Trans Electron Dev, 2013, 60: 246–253
55 Snider G S. Spike-timing-dependent learning in memristive nanodevices. In: Proceedings of 2008 IEEE/ACM International Symposium on Nanoscale Architectures NANOARCH 2008, Anaheim, 2008. 85–92
56 Wong H S P, Lee H Y, Yu S M, et al. Metal-oxide RRAM. Proc IEEE, 2012, 100: 1951–1970
57 Yang J J, Miao F, Pickett M D, et al. The mechanism of electroforming of metal oxide memristive switches. Nanotechnology, 2009, 20: 215201
58 Yang Y, Gao P, Li L, et al. Electrochemical dynamics of nanoscale metallic inclusions in dielectrics. Nat Commun, 2014, 5: 4232
59 Sarkar B, Lee B, Misra V. Understanding the gradual reset in Pt/Al_2O_3 /Ni RRAM for synaptic applications. Semicond Sci Technol, 2015, 30: 105014
60 Rolandi M, Josberger E E, Deng Y X. Two-terminal proton conducting devices with synaptic behavior and memory. In: Proceedings of 72nd Device Research Conference, Santa Barbara, 2014. 245–246
61 Yang R, Terabe K, Yao Y, et al. Synaptic plasticity and memory functions achieved in a WO_{3-x}-based nanoionics device by using the principle of atomic switch operation. Nanotechnology, 2013, 24: 384003
62 Jung J-W, Park S, Jeong Y-H. ReRAM-based synaptic device for neuromorphic computing. In: Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS), Melbourne VIC, 2014. 1054–1057
63 Mandal S, El-Amin A, Alexander K, et al. Novel synaptic memory device for neuromorphic computing. Sci Rep, 2014, 4: 5333
64 Gao B, Liu L, Kang J. Investigation of the synaptic device based on the resistive switching behavior in hafnium oxide. Prog Nat Sci Mater Int, 2015, 25: 47–50
65 Wang Y-F, Lin Y-C, Wang I-T, et al. Characterization and modeling of nonfilamentary Ta/TaOx/TiO2/Ti analog synaptic device. Sci Rep, 2015, 5: 10150
66 Yu S M, Wu Y, Jeyasingh R, et al. An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation. IEEE Trans Electron Dev, 2011, 58: 2729–2737
67 Gao B, Bi Y, Chen H Y, et al. Ultra-low-energy three-dimensional oxide-based electronic synapses for implementation of robust high-accuracy neuromorphic computation systems. ACS Nano, 2014, 8: 6998–7004
68 Choi H, Jung H, Lee J, et al. An electrically modifiable synapse array of resistive switching memory. Nanotechnology, 2009, 20: 345201
69 Panwar N, Kumar D, Upadhyay N K, et al. Memristive synaptic plasticity in Pr0.7Ca0.3MnO3 RRAM by bio-mimetic programming. In: Proceedings of 72nd Device Research Conference, Santa Barbara, 2014. 135–136
70 Pershin Y V, Di Ventra M. Neuromorphic, digital, and quantum computation with memory circuit elements. Proc IEEE, 2012, 100: 2071–2080
71 Kim S, Du C, Sheridan P, et al. Experimental demonstration of a second-order memristor and its ability to biorealistically implement synaptic plasticity. Nano Lett, 2015, 15: 2203–2211
72 Valov I, Waser R, Jameson J R, et al. Electrochemical metallization memories—fundamentals, applications, prospects. Nanotechnology, 2011, 22: 254003
73 Kozicki M N, Gopalan C, Balakrishnan M, et al. Nonvolatile memory based on solid electrolytes. In: Proceedings of Symposium on Non-Volatile Memory Technology, Orlando, 2004. 10–17
74 Kund M, Beitel G, Pinnow C-U, et al. Conductive bridging RAM (CBRAM): an emerging non-volatile memory technology scalable to sub 20nm. In: Technical Digest of IEEE International Electron Devices Meeting, Washington

DC, 2005. 754–757
75 Hirose Y, Hirose H. Polarity-dependent memory switching and behavior of Ag dendrite in Ag-photodoped amorphous As2S3 films. J Appl Phys, 1976, 47: 2767–2772
76 Gopalan C, Ma Y, Gallo T, et al. Demonstration of conductive bridging random access memory (CBRAM) in logic CMOS process. In: Proceedings of 2010 IEEE International Memory Workshop, Seoul, 2010. 1–4
77 Lu W, Jeong D S, Kozicki M, et al. Electrochemical metallization cellsblending nanoionics into nanoelectronics? MRS Bull, 2012, 37: 124–130
78 Liu Q, Sun J, Lv H, et al. Resistive switching: real-time observation on dynamic growth/dissolution of conductive filaments in oxide-electrolyte-based ReRAM. Adv Mater, 2012, 24: 1774
79 Ohno T, Hasegawa T, Tsuruoka T, et al. Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. Nat Mater, 2011, 10: 591–595
80 Atkinson R, Shiffrin R. Human Memory: a Proposed System and its Control Processes. 2nd ed. Psych Learn Motiv, 1968, 2: 89–195
81 Yu S M, Wong H S P. Modeling the switching dynamics of programmable-metallization-cell (PMC) memory and its application as synapse device for a neuromorphic computation system. In: Proceedings of 2010 IEEE International Electron Devices Meeting (IEDM), San Francisco, 2010. 520–523
82 Yu S M, Wong H S P. compact modeling of conducting-bridge random-access memory (CBRAM). IEEE Trans Electron Dev, 2011, 58: 1352–1360
83 Suri M, Querlioz D, Bichler O, et al. Bio-inspired stochastic computing using binary CBRAM synapses. IEEE Trans Electron Dev, 2013, 60: 2402–2409
84 Mahalanabis D, Barnaby H J, Gonzalez-Velo Y, et al. Incremental resistance programming of programmable metallization cells for use as electronic synapses. Solid State Electron, 2014, 100: 39–44
85 Jo S H, Chang T, Ebong I, et al. Nanoscale memristor device as synapse in neuromorphic systems. Nano Lett, 2010, 10: 1297–1301
86 Kim K H, Gaba S, Wheeler D, et al. A functional hybrid memristor crossbar-array/CMOS system for data storage and neuromorphic applications. Nano Lett, 2012, 12: 389–395
87 Petersen C C, Malenka R C, Nicoll R A, et al. All-or-none potentiation of CA3-CA1 synapses. Proc Nat Acad Sci USA, 1998, 95: 4732–4737
88 O'Connor D H, Wittenberg G M, Wang S S-H. Graded bidirectional synaptic plasticity is composed of switch-like unitary events. Proc Nat Acad Sci USA, 2005, 102: 9679–9684
89 Suri M, Bichler O, Querlioz D, et al. Bio-inspired computing with binary stochastic CBRAM synapses. IEEE Trans Electron Dev, 2013, 60: 2402–2409
90 Li S Z, Zeng F, Chen C, et al. Synaptic plasticity and learning behaviours mimicked through Ag interface movement in an Ag/conducting polymer/Ta memristive system. J Mater Chem C, 2013, 1: 5292–5298
91 Yang Y, Chen B, Lu W D. Memristive physically evolving networks enabling the emulation of heterosynaptic plasticity. Adv Mater, 2015, 27: 7720–7727
92 Ielmini D. Filamentary-switching model in RRAM for time, energy and scaling projections. In: Proceedings of 2011 IEEE International Electron Devices Meeting (IEDM), Washington DC, 2011. 17.2.1–17.2.4
93 Belmonte A, Kim W, Chan B T, et al. A thermally stable and high-performance 90-nm-based 1T1R CBRAM cell. IEEE Trans Electron Dev, 2013, 60: 3690–3695
94 Russo U, Kamalanathan D, Ielmini D, et al. Study of multilevel programming in programmable metallization cell (PMC) memory. IEEE Trans Electron Dev, 2009, 56: 1040–1047
95 Åkerman J. Toward a universal memory. Science, 2005, 308: 508–510
96 Wang K L, Alzate J G, Amiri P K. Low-power non-volatile spintronic memory: STT-RAM and beyond. J Phys D Appl Phys, 2013, 46: 074003
97 Augustine C, Mojumder N N, Fong X, et al. Spin-transfer torque MRAMs for low power memories: perspective and prospective. IEEE Sens J, 2012, 12: 756–766
98 Roy K, Fan D, Fong X, et al. Exploring spin transfer torque devices for unconventional computing. IEEE J Emerg Sel Top Circuits Syst, 2015, 5: 5–16
99 Devolder T, Hayakawa J, Ito K, et al. Single-shot time-resolved measurements of nanosecond-scale spin-transfer induced switching: stochastic versus deterministic aspects. Phys Rev Lett, 2008, 100: 057206
100 Zhang Y, Zhao W, Prenat G, et al. Electrical modeling of stochastic spin transfer torque writing in magnetic tunnel junctions for memory and logic applications. IEEE Trans Magn, 2013, 49: 4375–4378
101 Vincent A F, Larroque J, Locatelli N, et al. Spin-transfer torque magnetic memory as a stochastic memristive synapse for neuromorphic systems. IEEE Trans Biomed Circuits Syst, 2015, 9: 166–174
102 Zeng Z M, Amiri P K, Rowlands G, et al. Effect of resistance-area product on spin-transfer switching in MgO-based magnetic tunnel junction memory cells. Appl Phys Lett, 2011, 98: 072512
103 Zhou P, Zhao B, Yang J, et al. Energy reduction for STT-RAM using early write termination. In: Digest of Technical Papers of 2009 IEEE/ACM International Conference on Computer-Aided Design, San Jose, 2009. 264–268
104 Daughton J M. Advanced MRAM Concepts. 2001. http://www.nve.com/Downloads/mram2.pdf
105 Ovshinsky S R. Reversible electrical switching phenomena in disordered structures. Phys Rev Lett, 1968, 21: 1450–

1453
106 Wong H S P, Raoux S, Kim S, et al. Phase change memory. Proc IEEE, 2010, 98: 2201–2227
107 Lai S. Current status of the phase change memory and its future. In: Technical Digest of IEEE International Electron Devices Meeting, Washington DC, 2003. 10.1.1–10.1.4
108 Lankhorst M H R, Ketelaars B W S M M, Wolters R A M. Low-cost and nanoscale non-volatile memory concept for future silicon chips. Nat Mater, 2005, 4: 347–352
109 Park J-B, Park G-S, Baik H-S, et al. Phase-change behavior of stoichiometric $Ge_2Sb_2Te_5$ in phase-change random access memory. J Electrochem Soc, 2007, 154: H139–H141
110 Loke D, Lee T H, Wang W J, et al. Breaking the speed limits of phase-change memory. Science, 2012, 336: 1566–1569
111 Ovshinsky S R, Pashmakov B. Innovation providing new multiple functions in phase-change materials to achieve cognitive computing. MRS Proc, 2003, 803
112 Suri M, Bichler O, Querlioz D, et al. Phase change memory as synapse for ultra-dense neuromorphic systems: application to complex visual pattern extraction. In: Proceedings of 2011 IEEE International Electron Devices Meeting (IEDM), Washington DC, 2011. 4.4.1–4.4.4
113 Jackson B L, Rajendran B, Corrado G S, et al. Nanoscale electronic synapses using phase change devices. J Emerg Technol Comput Syst, 2013, 9: 12:1–12:20
114 Eryilmaz S B, Kuzum D, Jeyasingh R, et al. Brain-like associative learning using a nanoscale non-volatile phase change synaptic device array. Front Neurosci, 2014, 8: 1–11
115 Schaller R R. Moore's law: past, present and future. IEEE Spectr, 1997, 34: 52–59
116 Mead C. Neuromorphic electronic systems. Proc IEEE, 1990, 78: 1629–1636
117 Hasler P, Diorio C, Minch B A, et al. Single transistor learning synapse with long term storage. In: Proceedings of 1995 IEEE International Symposium on Circuits and Systems, Seattle, 1995. 3: 1660–1663
118 Merolla P, Arthur J, Akopyan F, et al. A digital neurosynaptic core using embedded crossbar memory with 45pJ per spike in 45nm. In: Proceedings of 2011 IEEE Custom Integrated Circuits Conference (CICC), San Jose, 2011. 1–4
119 Seo J, Brezzo B, Liu Y, et al. A 45nm CMOS neuromorphic chip with a scalable architecture for learning in networks of spiking neurons. In: Proceedings of 2011 IEEE Custom Integrated Circuits Conference (CICC), San Jose, 2011. 1–4
120 Bartolozzi C, Indiveri G. Synaptic dynamics in analog VLSI. Neural Comput, 2007, 19: 2581–2603
121 Mack C A. Fifty years of Moore's law. IEEE Trans Semicond Manuf, 2011, 24: 202–207
122 Likharev K K. Neuromorphic CMOL circuits. In: Proceedings of 2003 3rd IEEE Conference on Nanotechnology, San Francisco, 2003. 2: 339–342
123 Likharev K, Mayr A, Muckra I, et al. CrossNets: high-performance neuromorphic architectures for CMOL circuits. Ann N Y Acad Sci, 2003, 1006: 146–163
124 Likharev K K, Strukov D. B. CMOL: Devices, Circuits, and Architectures. In: Cuniberti G, Richter K, Fagas G, eds. Introducing Molecular Electronics. Berlin/Heidelberg: Springer, 2006. 447–477
125 Feldheim D L, Keating C D. Self-assembly of single electron transistors and related devices. Chem Soc Rev, 1998, 27: 1–12
126 Ma X, Strukov D B, Lee J H, et al. Afterlife for silicon: CMOL circuit architectures. In: Proceedings of 2005 5th IEEE Conference on Nanotechnology, Nagoya, 2005. 175–178
127 Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors. Nature, 1986, 323: 533–536
128 Maass W. Networks of spiking neurons: the third generation of neural network models. Neural Netw, 1997, 10: 1659–1671
129 Hodgkin A, Huxley A. A quantitative description of membrane current and its application to conduction and excitation in nerve. Bull Math Biol, 1990, 52: 25–71
130 Izhikevich E M. Hybrid spiking models. Philos Trans A Math Phys Eng Sci, 2010, 368: 5061–5070
131 O'Reilly R C. Biologically based computational models of high-level cognition. Science, 2006, 314: 91–94
132 Herz A V M, Gollisch T, Machens C K, et al. Modeling single-neuron dynamics and computations: a balance of detail and abstraction. Science, 2006, 314: 80–85
133 Brüderle D, Petrovici M A, Vogginger B, et al. A comprehensive workflow for general-purpose neural modeling with highly configurable neuromorphic hardware systems. Biol Cybern, 2011, 104: 263–296
134 Arthur J V, Boahen K. Silicon-neuron design: a dynamical systems approach. IEEE Trans Circuits Syst I-Regul Pap, 2011, 58: 1034–1043
135 Rachmuth G, Poon C-S. Transistor analogs of emergent iono-neuronal dynamics. HFSP J, 2008, 2: 156–166
136 Mead C. Analog VLSI and Neural Systems. Boston: Addison-Wesley Longman Publishing Co., Inc., 1989. 179–186
137 Pickett M D, Medeiros-Ribeiro G, Williams R S. A scalable neuristor built with Mott memristors. Nat Mater, 2013, 12: 114–117
138 Park S, Noh J, Choo M-L, et al. Nanoscale RRAM-based synaptic electronics: toward a neuromorphic computing device. Nanotechnology, 2013, 24: 384009
139 Serrano-Gotarredona T, Prodromakis T, Linares-Barranco B. A proposal for hybrid memristor-CMOS spiking neuromorphic learning systems. IEEE Circuits Syst Mag, 2013, 13: 74–88