

Chapter 4

Modeling and Simulation of Spiking Neural Networks with Resistive Switching Synapses



Valerio Milo

Abstract Artificial intelligence (AI) has recently reached excellent achievements in the implementation of human brain cognitive functions such as learning, recognition and inference by running intensively neural networks with deep learning on high-performance computing platforms. However, excessive computational time and power consumption required for achieving such performance make AI inefficient compared with human brain. To replicate the efficient operation of human brain in hardware, novel nanoscale memory devices such as resistive switching random access memory (RRAM) have attracted strong interest thanks to their ability to mimic biological learning in silico. In this chapter, design, modeling and simulation of RRAM-based electronic synapses capable of emulating biological learning rules are first presented. Then, the application of RRAM synapses in spiking neural networks to achieve neuromorphic tasks such as on-line learning of images and associative learning is addressed.

4.1 Introduction

In recent years, artificial intelligence (AI) has achieved outstanding performance in a wide range of machine learning tasks including recognition of faces [1] and speech [2] which now play a crucial role in many fields such as transportation and security. To obtain such achievements, AI has first exploited the availability of very large datasets for training deep neural networks (DNNs) in software according to deep learning [3]. Moreover, the maturity of high-performance computing hardware such as the graphics processing unit (GPU) [4] and the tensor processing unit (TPU) [5] has further contributed to accelerate DNN training, enabling to outperform human ability in certain tasks such as image classification [6] and playing board game of Go [7]. However, efficient implementation of AI tasks on modern digital computers based on von Neumann architecture and complementary metal-oxide-semiconductor

V. Milo (✉)

Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano and Italian Universities Nanoelectronics Team (IU.NET), Piazza Leonardo da Vinci 32, Milano 20133, Italy
e-mail: valerio.milo@polimi.it

© The Author(s) 2020

B. Pernici (ed.), *Special Topics in Information Technology*, PoliMI SpringerBriefs,
https://doi.org/10.1007/978-3-030-32094-2_4

49

(CMOS) technology has been recently challenged by fundamental issues such as the excessive power consumption and latency due to the looming end of Moore's law [8] and physical separation between memory and processing units [9].

To overcome the so-called von Neumann bottleneck of conventional hardware, novel non-von Neumann computing paradigms have been intensively explored with a view of bringing data processing closer to where data are stored [10]. In this wide range, neuromorphic computing has emerged as one of the most promising approaches since it aims at improving dramatically computation mainly in terms of energy efficiency taking inspiration from how the human brain processes information via biological neural networks [11].

In the last decade, strong research efforts in the field of neuromorphic engineering have led to build medium/large-scale analog/digital neuromorphic systems using CMOS technology [9, 12, 13]. However, the use of bulky CMOS circuits to closely reproduce synaptic dynamics has proved to be a major issue toward hardware integration of massive synaptic connectivity featuring human brain. This limitation has thus led to the exploration of novel memory device concepts, such as resistive switching random access memory (RRAM) and phase change memory (PCM), which display features suitable for synaptic application such as nanoscale size, fast switching behavior, low power operation, and tunable resistance by application of electrical pulses enabling to emulate synaptic plasticity at device level [14].

This chapter covers the application of hafnium-oxide (HfO_2) RRAM devices as plastic synapses in spiking neural networks (SNNs) to implement brain-inspired neuromorphic computing. After reviewing the main features of brain-inspired computation, physical mechanisms and operation principle of RRAM devices are described. Then, the scheme and operation of two hybrid CMOS/RRAM synapse circuits capable of implementing bio-realistic learning rules such as spike-timing dependent plasticity (STDP) and spike-rate dependent plasticity (SRDP) are presented. Finally, SNNs with resistive synapses are explored in simulation demonstrating their ability to achieve fundamental cognitive primitives such as on-line learning of visual patterns and associative memory.

4.2 Brain-Inspired Computing

Brain-inspired computing is considered a promising approach capable of tackling issues challenging today's digital processors thanks to its ability to replicate the massive parallelism and high energy efficiency of the human brain. To achieve such performance, human brain first relies on a high-density layered architecture consisting of large networks of biological processing units referred to as neurons where each neuron is connected with other neurons via 10^4 synapses on average [14]. In addition to the architecture, another feature playing a key role in brain computation is the spike-driven information processing. As illustrated in Fig. 4.1a, in biological neural networks the neurons interact with the next neuron by propagation of voltage spikes along the axon and their transmission through the synapses, namely the nanoscale gaps between the axon terminals and the dendrites. As a result of spike transmis-

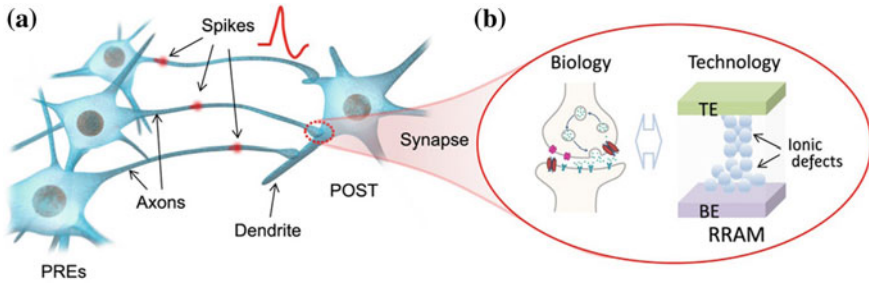


Fig. 4.1 **a** Illustration of some PREs connected to a single POST by synapses in a bio-realistic neural network. PREs emit spikes that are sent to the POST causing an increase of its membrane potential. **b** Sketch evidencing the strong analogy between a biological synapse where conductivity is tuned by voltage-induced ion migration and a RRAM device where conductivity is controlled by ionic defect migration resulting in formation/retraction of conductive paths. Reprinted from [18]

sion, calcium ions diffuse in the neuron activating the release of neurotransmitters within the synaptic gap where they diffuse eventually binding to sites of sodium ion channels of the post-synaptic neuron. This induces the opening of sodium ion channels and consequently the diffusion of sodium ions in the cell, which results in an increase of membrane potential leading to the generation of a spike by post-synaptic neuron as it crosses an internal threshold [14]. These biological processes at synaptic level thus suggest that information is processed by generation and transmission of spikes whose short duration of the order of 1 ms combined with typical neuron spiking rate of 10 Hz leads to a power consumption of only 20 W that is dramatically lower than power dissipated by modern computers [9, 14]. Therefore, brain computing relies on neurons integrating input signals sent by other neurons and firing spikes after reaching the threshold, and synapses changing their weight depending on spiking activity of pre-synaptic neuron (PRE) and post-synaptic neuron (POST). In particular, two biological rules like STDP [15] and SRDP [16] are considered two fundamental schemes controlling synaptic weight modulation, which is in turn believed to underlie learning ability in the brain.

To faithfully replicate the brain-inspired computing paradigm in hardware, in recent years neuromorphic community has intensively investigated novel material-based devices such as RRAM which, as shown in Fig. 4.1b, can mimic biological processes governing synaptic plasticity by exploiting its ability to change resistance via creation/rupture of conductive filaments in response to application of voltage pulses thanks to the resistive switching phenomenon [17, 18].

4.3 Resistive Switching in RRAM Devices

RRAM is a two-terminal nanoscale memory device displaying resistive switching, namely the ability to change the device resistance via the creation and disruption of filamentary conductive paths under the application of voltage pulses. The RRAM

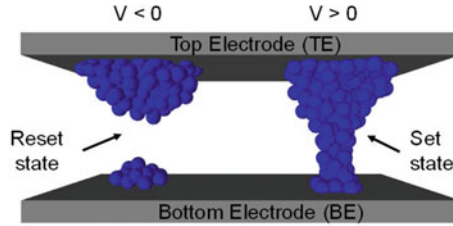


Fig. 4.2 Sketch of RRAM states, namely the reset state (left) and the set state (right). The application of a positive voltage to the TE causes the migration of defects toward the depleted gap, thus recreating the continuous filamentary path. The application of a negative voltage to the TE causes the retraction of defects back to the top reservoir, and the formation of a depleted gap

structure relies on a metal-insulator-metal (MIM) stack where a transition metal oxide layer, such as HfO_x , TiO_x , TaO_x and WO_x , is sandwiched between two metal electrodes referred to as top electrode (TE) and bottom electrode (BE), respectively [19, 20]. To initiate the filamentary path across the MIM stack, a soft breakdown process called forming is first operated, which locally creates a high concentration of defects, e.g., oxygen vacancies or metallic impurities, enhancing conduction. After forming, RRAM can exhibit bipolar resistive switching, where the application of a positive voltage can induce an increase of conductance, or set transition, whereas the application of a negative voltage can lead to reset transition, or a decrease of conductance [20].

Figure 4.2 schematically shows the 2 states of a RRAM device, namely the reset state (left), or high resistance state (HRS), where the filamentary path of defects is disconnected because of a previous reset transition. Under a positive voltage applied to the TE, defects from the top reservoir migrate toward the depleted gap, thus leading to the set state (right), or low resistance state (LRS), via a set transition. Applying a negative voltage to the TE causes the retraction of defects from the filament toward the top reservoir, and the formation of the depleted gap [20]. Note that the TE and BE generally differ by material and/or structure, with the TE being generally chemically active, e.g., Ti, Hf, or Ta, which enables the creation of an oxygen exchange layer with a relatively high concentration of oxygen vacancies [21]. On the other hand, the BE is chemically inert to prevent set transition when a positive voltage is applied to the BE [22].

4.4 Synapse Circuits with RRAM Devices

RRAM devices exhibit features as nanoscale size and low current operation making them attractive for realization of hardware synapses. The major challenge thus consists of implementing the synaptic plasticity schemes considered at the basis of biological learning such as STDP [15] and SRDP [16] at device level. To achieve synaptic

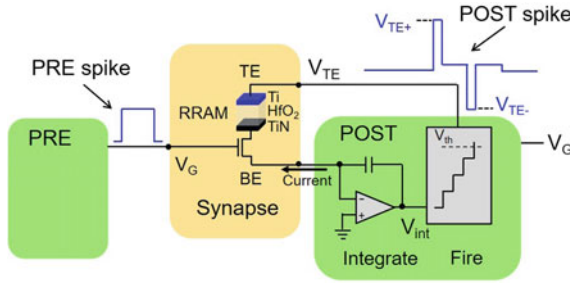


Fig. 4.3 Schematic of a 1T1R RRAM structure serving as synaptic connection between a PRE circuit and a POST circuit with I&F architecture. Application of a PRE spike at FET gate induces a synaptic current in 1T1R cell which is integrated by POST leading its membrane potential V_{int} to become more positive. As V_{int} hits the threshold V_{th} , a POST spike is backward sent at TE to update synaptic weight according to STDP rule. Adapted with permission from [24]. Copyright 2016 IEEE

plasticity in hardware, the combination of RRAM devices and field-effect transistors (FETs) serving as both cell selectors and current limiters has been widely used leading to the design of hybrid synaptic structures such as the one-transistor/one-resistor (1T1R) structure [23, 24] and the four-transistors/one-resistor (4T1R) structure [24, 25].

4.4.1 1T1R Synapse

Figure 4.3 shows a circuit schematic where a hybrid structure based on serial connection of a Ti/HfO₂/TiN RRAM cell and a FET, referred to as 1T1R structure, works as electronic synapse connecting a PRE to a POST with an integrate-and-fire (I&F) architecture. This building block was designed to achieve STDP rule which is considered one of the key mechanisms regulating learning in mammals. According to STDP rule, synaptic weight can change depending on the relative time delay Δt between spikes emitted by PRE and POST. If the PRE spike precedes the POST spike, Δt is positive resulting in an increase of synaptic weight or long-term potentiation (LTP) of the synapse. Otherwise, if PRE spike generation takes place slightly after the POST spike, Δt has negative value leading to a decrease of synaptic weight, or long-term depression (LTD) of the synapse [15]. The STDP implementation in 1T1R synapse was achieved as follows. When the PRE sends a 10-ms-long voltage pulse at FET gate, a current proportional to RRAM conductance flows across the synapse since its TE is biased by a continuous voltage with low amplitude used for communication phase. This current thus enters POST where it is integrated, causing an increase of POST membrane/internal potential V_{int} . As this integral signal crosses a threshold V_{th} , the POST sends both a forward spike toward next neuron layer and a suitably-designed spike including a 1-ms-long pulse with positive amplitude

followed, after 9 ms, by a 1-ms-long negative pulse, which is backward delivered at TE to activate a synaptic weight update according to STDP rule. Here, as the PRE spike precedes the POST spike ($0 < \Delta t < 10$ ms), PRE voltage overlaps only with the positive pulse of the POST spike, thus causing a set transition within RRAM device resulting in LTP of 1T1R synapse. Otherwise, if the PRE spike follows the POST spike ($-10 \text{ ms} < \Delta t < 0$), overlap occurs between PRE spike and the negative pulse in the POST spike, thus activating a reset transition in RRAM device resulting in LTD of 1T1R synapse [23, 24]. Note that STDP implementation in 1T1R RRAM synapse was demonstrated in both simulation using a stochastic Monte-Carlo model of HfO₂ RRAM [23] and hardware as reported in [26].

4.4.2 4T1R Synapse

In addition to STDP implementation via 1T1R synapse, a novel hybrid CMOS/RRAM synapse was designed to replicate another fundamental biological learning rule called SRDP which states that high-frequency stimulation of PREs leads to synaptic LTP whereas a low-frequency stimulation of PREs leads to synaptic LTD [16]. As shown in Fig. 4.4a, PRE and POST are connected by a synapse circuit based on a 4T1R structure including a HfO₂ RRAM device and 4 FETs shared into M₁–M₂ and M₃–M₄ branches. PRE circuit includes both a signal channel transmitting a Poisson-distributed spike train with average frequency f_{PRE} at the M₁ gate and its copy delayed by a time Δt_D at the M₂ gate, and a noise channel driving M₃ by application of noise spikes at low frequency f_3 , whereas POST circuit relies on an I&F stage with fire and random noise outputs which alternatively control RRAM TE and M₄ gate by a multiplexer. If $f_{PRE} > \Delta t_D^{-1}$, the high probability that M₁ and M₂ are simultaneously enabled by overlapping PRE spikes at gate terminals leads to the activation of synaptic currents within M₁/M₂ branch causing, after integration by

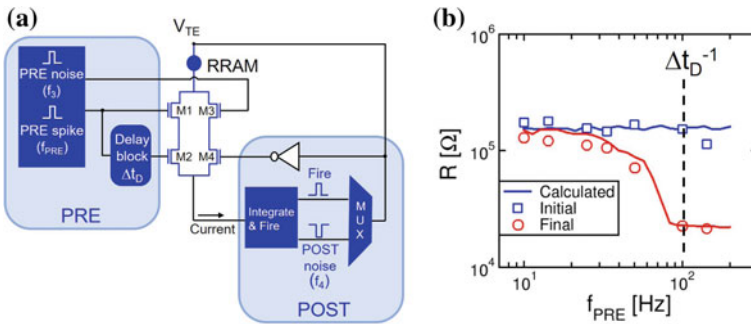


Fig. 4.4 **a** Schematic of 4T1R RRAM circuit operating as synaptic connection between a PRE circuit and a POST circuit with I&F architecture. **b** Measured and calculated resistance of 4T1R synapse for increasing f_{PRE} supporting high-frequency synaptic potentiation. Adapted from [25]

I&F stage, the generation of a fire pulse which is delivered at TE. As a result, overlap between spikes applied to M_1 gate, M_2 gate, and TE induces a set transition within RRAM device, namely LTP. Note that M_1/M_2 branch is the only active branch during LTP operation since fire pulse disables M_4 because of the inversion by NOT gate in the POST. Otherwise, if $f_{PRE} < \Delta t_D^{-1}$, no spike overlap between channels driving M_1 and M_2 takes place, thus making LTP branch disabled. To achieve LTD at low f_{PRE} , stochastic overlap among random noise voltage spikes provided at gate of M_3 and M_4 and a negative pulse applied at TE is exploited. In fact, these overlapping spikes cause a reset transition within RRAM device resulting in LTD at the synapse. Therefore, the operation principle of 4T1R synapse supports its ability to replicate SRDP rule implementing LTP at high f_{PRE} and a noise-induced stochastic LTD at low f_{PRE} [24, 25].

The ability to reproduce SRDP in 4T1R synapse was validated testing separately LTP and LTD by use of 2T1R integrated structures. Figure 4.4b shows measured and calculated RRAM resistance as a function of f_{PRE} , evidencing that a resistance change from initial HRS to LRS can be activated in a 2T1R structure serving as M_1/M_2 branch only when f_{PRE} is greater or equal than the reciprocal of $\Delta t_D = 10$ ms used in the experiment, namely $f_{PRE} \geq 100$ Hz. Similar to LTP, LTD was also experimentally demonstrated in [24, 25] evidencing that RRAM resistance initialized in LRS can reach HRS provided that PRE noise frequency f_3 is higher than POST noise frequency f_4 to achieve a sufficiently high overlap probability among random noise spikes controlling LTD branch.

4.5 Spiking Neural Networks with RRAM Synapses

4.5.1 Unsupervised Pattern Learning by SRDP

The fundamental ability of biological brain consists of learning by adaptation to environment with no supervision. This process, which is referred to as unsupervised learning, relies on schemes such as STDP and SRDP that adjust synaptic weights of neural networks according to timing or rate of spikes encoding information such as images and sounds. In particular, unsupervised learning of visual patterns has recently attracted increasing interest leading to many simulation/hardware demonstrations of SNNs with RRAM synapses capable of STDP [18, 23, 24, 26–30] and SRDP [25, 31].

Figure 4.5a illustrates a SNN inspired to perceptron network model developed by Rosenblatt in the late 1950s [32] consisting of 64 PREs fully connected to a single POST. This network was simulated using 4T1R RRAM structures as synapses to demonstrate unsupervised on-line learning of images with SRDP by conversion of image pattern (the ‘X’) and its surrounding background in high-frequency spiking activity of PREs and low-frequency spiking activity of PREs, respectively. To achieve this goal, after random initialization of synaptic weights between HRS and LRS

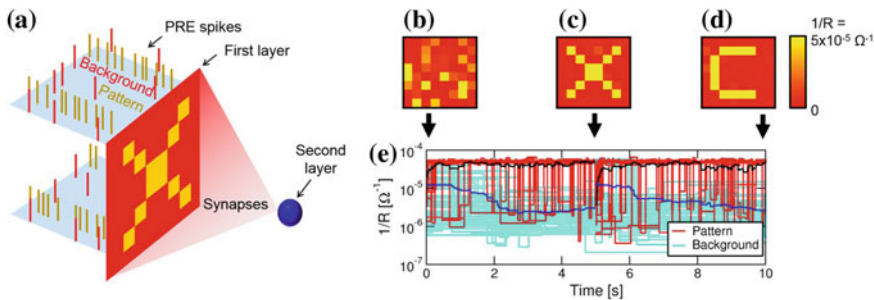


Fig. 4.5 **a** Sketch of a perceptron neural network where PREs are fully connected to a single POST. PREs receive spike sequences at high/low frequency upon pattern/background presentation which results in potentiation/depression of corresponding synaptic weights according to SRDP rule. Color plots of synaptic weights at **b** $t = 0$, **c** $t = 5$ s and **d** $t = 10$ s during a calculated on-line learning of two 8×8 sequential images. **e** Time evolution of calculated conductance supporting the ability of perceptron with 4T1R SRDP synapses to learn a new pattern after forgetting the previously stored pattern. Adapted from [25]

(Fig. 4.5b), an image including an ‘X’ pattern was submitted for 5 s to the input layer leading to high-frequency stimulation of PREs ($f_{PRE} = 100$ Hz) within the ‘X’ and low-frequency stimulation of PREs ($f_{PRE} = 5$ Hz) outside the ‘X’. As described in Sect. 4.4.2, this resulted in a selective LTP of ‘X’ synapses and stochastic LTD of background synapses, thus leading synaptic weights to adapt to submitted image within 5 s (Fig. 4.5c). After $t = 5$ s, ‘X’ image was replaced by a ‘C’ image with no overlap with ‘X’ for other 5 s resulting in a high/low frequency stimulation of PREs within/outside ‘C’. As shown by color map in Fig. 4.5d, external stimulation causes potentiation of ‘C’ synapses and depression of all the other synapses, which evidences network ability to learn a new pattern erasing the previously stored pattern [25]. This result is also supported by calculated time evolution of synaptic conductance during learning shown in Fig. 4.5e, which displays a fast conductance increase (LTP) of pattern synapses and a slower conductance decrease (LTD) of background synapses achieved using PRE and POST noise frequencies equal to $f_3 = 50$ Hz and $f_4 = 20$ Hz, respectively [25]. This simulation study corroborated the ability of perceptron SNNs to learn on-line sequential images thanks to 4T1R RRAM synapses capable of SRDP, and provided a solid basis for its experimental demonstration presented in [25].

4.5.2 Associative Memory with 1T1R RRAM Synapses

In addition to unsupervised image learning, another brain-inspired function receiving strong interest is the associative memory, namely the ability to retrieve past memories by their partial stimulation. To achieve this cognitive primitive, pioneering works by Hopfield [33] and Amit [34] focused on a type of neural network, called recurrent

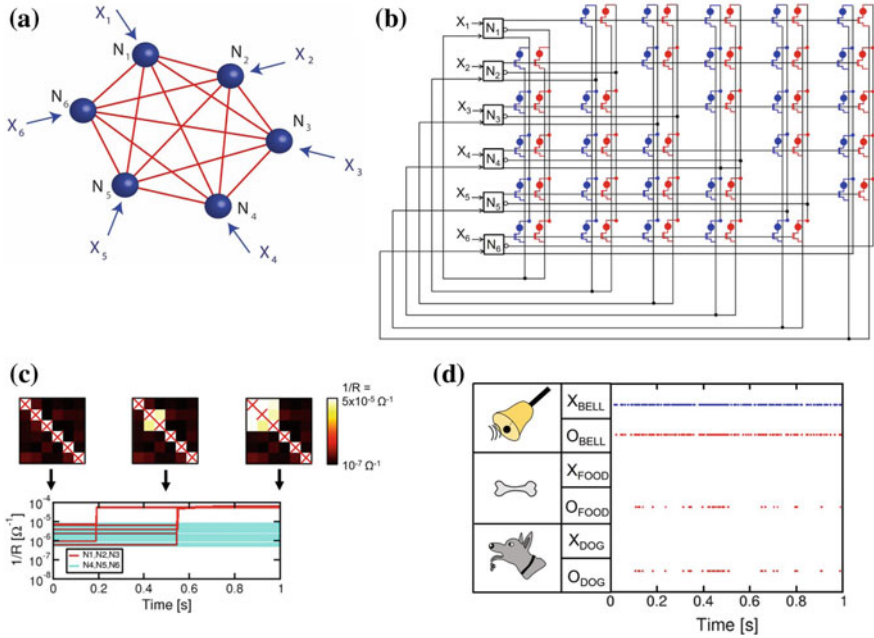


Fig. 4.6 **a** Sketch of a Hopfield recurrent neural network with 6 fully connected neurons receiving external stimulation X_i , $i = 1:6$. **b** Circuit schematic of a Hopfield network with 6 neurons fully connected via bidirectional excitatory (blue) and inhibitory (red) 1T1R RRAM synapses. **c** Calculated attractor learning process evidencing storage of the attractor state linking N_1 , N_2 and N_3 . **d** Illustration of associative memory implementation by recall process in an attractor network with RRAM synapses taking inspiration from Pavlov’s dog experiments. Adapted with permission from [35]. Copyright 2017 IEEE

neural network or attractor neural network, where the neurons are fully connected with each other by bidirectional excitatory/inhibitory synapses.

Figure 4.6a shows a sketch of a Hopfield recurrent neural network with 6 neurons evidencing all-to-all synaptic connections, external stimulation X_i provided to each neuron, and the absence of self-feedback to prevent divergence of network dynamics. Based on this configuration, the circuit schematic of Hopfield neural network with bidirectional excitatory/inhibitory 1T1R RRAM synapses shown in Fig. 4.6b was designed and tested in simulation [35]. In this network, each neuron is implemented by an I&F block that provides signal to other neurons as a PRE but also receives signal by other neurons as a POST. Specifically, each I&F neuron has two current inputs, namely the external current spikes X_i and the sum of weighted currents activated by other neurons, and three outputs driving the gate of row 1T1R synapses, the TE of column excitatory synapses (blue 1T1R cells) and the TE of column inhibitory synapses (red 1T1R cells) by voltage pulses, respectively [35].

To demonstrate associative memory in this Hopfield network implementation, the network was first operated in learning mode stimulating a subset of neurons (N_1 , N_2 ,

and N_3) with high-frequency Poisson distributed spike trains to store their attractor state. This process led each attractor neuron to fire three output spikes, namely (i) a positive voltage pulse at gate of all its row 1T1R synapses, (ii) a positive voltage pulse at TE of its column excitatory 1T1R synapses and (iii) a negative voltage pulse at TE of its column inhibitory 1T1R synapses, by causing a stochastic co-activation of attractor neurons at certain times. As this event occurs, voltage overlap at gate/TE of synapses shared by pairs of attractor neurons leads mutual excitatory 1T1R cells to undergo a set transition, thus LTP, whereas the corresponding inhibitory 1T1R cells undergo a reset transition, thus LTD. It means that learning phase in this Hopfield SNN with 1T1R synapses consists of the storage of attractor state associated with externally stimulated neurons via potentiation/depression of mutual excitatory/inhibitory synapses by a STDP-based learning scheme inspired to well-known Hebb's postulate stating that neurons that fire together, wire together [36].

Figure 4.6c shows a calculated learning process with N_1 , N_2 , and N_3 as attractor neurons evidencing the gradual storage of corresponding attractor state via potentiation of mutual excitatory synapses (top) due to the convergence of corresponding RRAM devices to high conductance values (bottom). After implementing attractor learning, the network was operated in another mode referred to as recall mode. During recall process, a high-frequency stimulation of a part of stored attractor state is applied, e.g., only one out of 3 attractor neurons, leading to activation of high currents across high conductance synapses shared with other attractor neurons which are transmitted at their inputs. As a result, attractor neurons with no external stimulation start spiking, thus retrieving the whole attractor state via a sustained spiking activity able to persist even after removing external stimulation [35]. Importantly, the ability of Hopfield networks to recall an attractor state by its incomplete stimulation was exploited to replicate in simulation a fundamental mammalian primitive referred to as associative memory taking inspiration from the Pavlov's dog experiments [37]. Indeed, as illustrated in Fig. 4.6d, after repeatedly stimulating a dog combining the ring of a bell with the presentation of food leading it to salivate, an external stimulation only with the ring of bell is able to reactivate the bell-food-salivation attractor in the dog [35]. Finally, Hopfield neural network with 1T1R synapses was also successfully used to explore in simulation pattern completion task, namely the reconstruction of an image stimulating its small features [38], which supports the strong potential of Hopfield neural networks with resistive synapses in computational tasks.

4.6 Conclusions

This chapter covers design, modeling and simulation of SNNs with CMOS/RRAM synapses capable of implementing brain-inspired neuromorphic computing. First, unsupervised learning at synaptic level has been addressed by development of 1T1R synapse and 4T1R synapse capable of replicating STDP and SRDP, respectively. Then, applications of these resistive synapses in SNNs have been investigated. A perceptron network with 4T1R synapses has been simulated demonstrating its ability to

achieve on-line learning of sequential images via SRDP-based adaptation of synaptic weights. In addition, attractor learning and recall processes have been achieved in a Hopfield recurrent neural network with excitatory/inhibitory 1T1R synapses by simulations supporting its ability to implement associative memory. These results support RRAM as promising technology for future development of large-scale neuromorphic systems capable of emulating unrivaled energy and computational efficiency of biological brain.

References

1. Taigman Y, Yang M, Ranzato M, Wolf L (2014) DeepFace: Closing the gap to human-level performance in face verification. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 1701–1708
2. Xiong W et al (2017) The Microsoft 2017 conversational speech recognition system. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 5934–5938
3. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444
4. Coates A et al (2013) Deep learning with COTS HPC systems. In: Proceedings of the 30th International Conference on Machine Learning vol 28(3), pp 1337–1345
5. Jouppi NP et al (2017) In-datacenter performance analysis of a Tensor Processing UnitTM. In: Proceedings of the 44th Annual International Symposium on Computer Architecture (ISCA), pp 1–12
6. He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In: IEEE ICCV, pp 1026–1034
7. Silver D et al (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* 529:484–489
8. Theis TN, Wong H-SP (2017) The end of Moore’s law: a new beginning for information technology. *Comput Sci Eng* 19(2):41–50
9. Merolla PA et al (2014) A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* 345(6197):668–673
10. Wong H-SP, Salahuddin S (2015) Memory leads the way to better computing. *Nat Nanotechnol* 10(3):191–194
11. Mead C (1990) Neuromorphic electronic systems. *Proc IEEE* 78(10):1629–1636
12. Furber SB, Galluppi F, Temple S, Plana LA (2014) The SpiNNaker project. *Proc IEEE* 102(5):652–665
13. Qiao N et al (2015) A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128K synapses. *Front Neurosci* 9:141
14. Kuzum D, Yu S, Wong H-SP (2013) Synaptic electronics: materials, devices and applications. *Nanotechnology* 24:382001
15. Bi G-Q, Poo M-M (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and post synaptic cell type. *J Neurosci* 18(24):10464–10472
16. Bear MF (1996) A synaptic basis for memory storage in the cerebral cortex. *Proc Natl Acad Sci USA* 93(24):13453–13459
17. Yu S et al (2011) An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation. *IEEE Trans Electron Devices* 58(8):2729–2737
18. Wang W et al (2018) Learning of spatiotemporal patterns in a spiking neural network with resistive switching synapses. *Sci Adv* 4:eaat4752
19. Wong H-SP et al (2012) Metal-oxide RRAM. *Proc IEEE* 100(6):1951–1970

20. Ielmini D (2016) Resistive switching memories based on metal oxides: mechanisms, reliability and scaling. *Semicond. Sci Technol* 31(6):063002
21. Lee HY et al (2008) Low power and high speed bipolar switching with a thin reactive Ti buffer layer in robust HfO₂ based RRAM. In: *IEEE IEDM Tech Dig* 297–300
22. Balatti S et al (2015) Voltage-controlled cycling endurance of HfO_x-based resistive-switching memory (RRAM). *IEEE Trans Electron Devices* 62(10):3365–3372
23. Ambrogio S et al (2016) Neuromorphic learning and recognition with one-transistor-one-resistor synapses and bistable metal oxide RRAM. *IEEE Trans Electron Dev* 63(4):1508–1515
24. Milo V et al (2016) Demonstration of hybrid CMOS/RRAM neural networks with spike time/rate-dependent plasticity. In: *IEEE IEDM Tech Dig* 440–443
25. Milo V et al (2018) A 4-transistors/one-resistor hybrid synapse based on resistive switching memory (RRAM) capable of spike-rate dependent plasticity (SRDP). *IEEE Trans Very Large Scale Integration (VLSI) Syst* 26(12):2806–2815
26. Pedretti G et al (2017) Memristive neural network for on-line learning and tracking with brain-inspired spike timing dependent plasticity. *Sci Rep* 7(1):5288
27. Yu S et al (2012) A neuromorphic visual system using RRAM synaptic devices with sub-pJ energy and tolerance to variability: experimental characterization and large-scale modeling. In: *IEEE IEDM Tech Dig* 239–242
28. Suri M et al (2012) CBRAM devices as binary synapses for low-power stochastic neuromorphic systems: auditory (cochlea) and visual (retina) cognitive processing applications. In: *IEEE IEDM Tech Dig* 235–238
29. Serb A et al (2016) Unsupervised learning in probabilistic neural networks with multi-state metal-oxide memristive synapses. *Nat Commun* 7:12611
30. Prezioso M et al (2018) Spike-timing-dependent plasticity learning of coincidence detection with passively integrated memristive circuits. *Nat Commun* 9:5311
31. Ohno T et al (2011) Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. *Nat Mater* 10(8):591–595
32. Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 65(6):386–408
33. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA* 79:2554–2558
34. Amit DJ (1989) *Modeling brain function: the world of attractor neural networks*. Cambridge University Press, Cambridge
35. Milo V, Ielmini D, Chicca E (2017) Attractor networks and associative memories with STDP learning in RRAM synapses. In: *IEEE IEDM Tech Dig* 263–266
36. Hebb DO (1949) *The organization of behavior: a neurophysiological theory*. Wiley, New York
37. Pavlov IP (1927) *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. Oxford University Press, London
38. Milo V, Chicca E, Ielmini D (2018) Brain-inspired recurrent neural network with plastic RRAM synapses. In: *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp 1–5

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

