Circuit Techniques for Online Learning of Memristive Synapses in CMOS-Memristor Neuromorphic Systems

Sagarvarma Sayyaparaju, Gangotree Chakma, Sherif Amer, and Garrett S. Rose

ACM Great Lakes Symposium on VLSI (GLSVLSI), Banff, Alberta, Canada, May 2017.

DOI: https://doi.org/10.1145/3060403.3060418

Circuit Techniques for Online Learning of Memristive Synapses in CMOS-Memristor Neuromorphic Systems

Sagarvarma Sayyaparaju, Gangotree Chakma, Sherif Amer and Garrett S. Rose
Department of Electrical Engineering & Computer Science
University of Tennessee, Knoxville
1520 Middle Drive, Knoxville, Tennessee
{ssayyapa, gchakma, samer1, garose}@utk.edu

ABSTRACT

Memristors are widely leveraged in neuromorphic systems for constructing synapses. Resistance switching characteristics of memristors enable online learning in synapses. This paper addresses a fundamental issue associated with the design of synapses with memristors whose switching rates in either direction differ up to two orders of magnitude. A twin-memristor synapse that uses memristors with identical switching rates is first presented. It is shown this design fails in the case of disproportionate switching times. To circumvent this issue, a quad-memristor synapse is also considered. The scheme used for online learning of the synapse circuit implementation, and simulation results are also presented. To compare the two synapses, their area, clock frequency, dynamic power and energy per spike values are provided.

Keywords

Memristors, Online learning, Switching rate, Synapse

1. INTRODUCTION

Neuromorphic systems are brain-inspired computing systems, wherein synapses are utilized as the medium of communication. Synapses are responsible for the adaptability of any system which mimics the brain. Learning is that characteristic of the synapse which enables it to update itself depending upon the relative timing of the occurrence of input and output spikes. The learning mechanism is mainly classified into two types: online learning and offline learning. Online learning is often based on unsupervised learning such as Spike Time Dependent Plasticity (STDP), Long Term Potentiation (LTP), and Long Term Depression (LTD) [9, 10]. Memristors are widely utilized to construct synapses in neuromorphic computing systems [5] owing to the analog programmability of their resistance, which represents the synaptic weight. In addition, memristors are nano-scale devices, hence paving the way for low power designs. During the online learning of these memristive synapses, the weight of the synapse is altered by the application of voltage pulses across the memristors.

Memristors can be made to switch between several resistance states. For a given voltage across the memristor, the specific switching mechanism in the device determines the time it takes to switch from the Low Resistance State (LRS) to the High Resistance State (HRS) or vice-versa. However, due to the disparate nature of the switching mechanism for opposite polarities, the switching rates could differ significantly [1–3]. Failure to acknowledge this difference in the switching rates leads to functional errors in the operation of memristive systems. Specifically, with same voltage applied in both the directions, the memristor switches in a single step in one of the directions (to attain LRS or HRS), but attains an intermediate resistance state (due to slower switching rate) in the other direction. In this paper, we present two different mechanisms for online learning, to cater to this dichotomy of the symmetric and asymmetric switching rates possible in memristive devices. We also describe the advantages and disadvantages associated with each learning technique depending on the above-mentioned switching mechanism and the available memristor models.

The remainder of the paper is organized as follows. Section 2 details the construct of the twin-memristor synapse that uses memristors with symmetric switching rates. Section 3 delineates the quad-memristor approach for designing with memristors with asymmetric switching rates. Section 4 illustrates some results to compare the two synapses and details some thoughts that emerge from this discussion from the perspective of both circuit designers and device personnel. Section 5 concludes the paper.

2. TWIN-MEMRISTOR SYNAPSE

The twin-memristor synapse structure is comprised of two memristors connected with opposite polarity. The synaptic weights are represented using memristors, where the memristance is proportional to the desired weight. This approach follows that of several other memristor-based neural network designs where voltage inputs across memristive weights yields a weighted sum in the form of a current [4, 6–8].

Negative weights are achieved through the use of a pair of memristors (Figure 1) for every synaptic weight. This approach is similar to that presented in [4] where a pair of memristive crossbar arrays are used to represent the negative and positive components of the weights. In the memristor pair for each synapse, one memristor is used to drive positive current while the other drives negative current (pulls current from the integrator). If the memristance of both

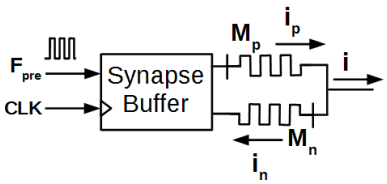

Figure 1: Schematic of the twin-memristor synapse structure showing two memristors used to provide both positive and negative weights.

memristors in the pair are equal then the currents will cancel each other for any given input spike and the effective weight is zero. From this point, where the memristance values are equal, one memristance can be increased (decreased) to change the weight while the other memristance decreases (increases) by a complementary amount. The weight of the synapse is proportionate to the effective conductivity of the pair of memristors, given by:

$$G_{eff} = \frac{1}{M_p} - \frac{1}{M_n}. \quad (1)$$

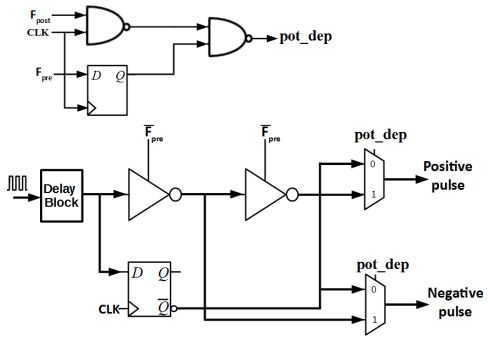

Figure 2: Detailed view of the synaptic buffer.

The synaptic buffer (Figure 2) is used to generate driving currents for both memristors in a synapse. It contains a delay block (chain of shift registers) which represents the delay of a synapse and it also handles the online learning mechanism. The synaptic buffer generates a signal "*pot_dep*" in itself which determines the potentiation or depression of a synapse. This signal becomes low whenever there is a post-synaptic fire caused by the pre-synaptic spike, otherwise it remains high such that online learning is inactive during that period. The condition for potentiation or depression is determined by the occurrence of both pre and post synaptic fires. If the pre-synaptic fire causes a post-synaptic firing spike, the synapse associated with the firing is potentiated or the weight of the synapse is increased (Figure 3). Likewise, if the pre-synaptic fire arrives simultaneously with the post-synaptic signal, the synapse is depressed.

3. QUAD-MEMRISTOR SYNAPSE

During the learning process in the twin-memristor synapse, both memristors switch in opposite directions with voltage pulses of opposite polarity and of the same duration. However, switching rates are rarely equal for both increasing and decreasing memristance. In this paper, we consider the

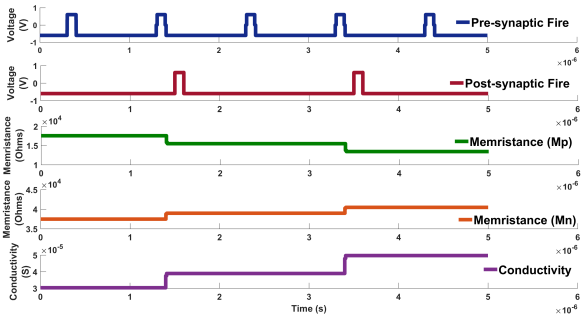

Figure 3: Basic LTP operation on the twin-memristor synapse.

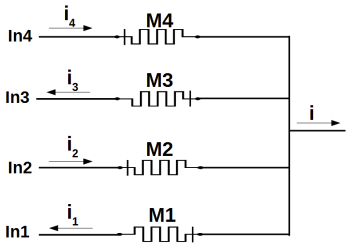

Figure 4: Simplified view of the quad-memristor synapse.

HfO_2 device described in [1]. Based on experimental results reported therein, we conservatively assume that the time taken for switching from LRS to HRS is 1μs whereas switching from HRS to LRS takes only 10ns. In the case where a voltage pulse greater than the switching threshold and of duration larger than or equal to 10ns is applied, the memristor switches from HRS to LRS in a single step. Thus, for such voltage pulses, the memristor is unable to display analog programmability, making LTP/LTD infeasible.

To be able to perform LTP/LTD in the face of asymmetric switching rates of memristors, we propose a quad-memristor synapse structure as follows.

3.1 Working Scheme

A simplified view of this synapse is shown in Figure 4. Currents through M1 and M3 flow in the negative direction, whereas those through M2 and M4 flow in the positive direction. The total current flowing out of the synapse is proportional to its weight and its direction determines if the weight is positive or negative. The current through the synapse is proportional to the effective conductivity of the memristor cluster, given by:

$$G_{eff} = \frac{1}{M2} + \frac{1}{M4} - \frac{1}{M1} - \frac{1}{M3}. \quad (2)$$

.

The learning mechanism for this synapse abides by the following set of rules:

1) A memristor switching during LTP/LTD is always from LRS to HRS.

2) Only one of the memristors switches during LTP/LTD. Either of M1 and M3 is selected for potentiation while one of M2 and M4 is used for depression.

3) Any pair of memristors, whose memristances are equal and both higher than LRS, are both reset to LRS. The possible combinations being (M1, M2), (M2, M3) and (M1, M4).

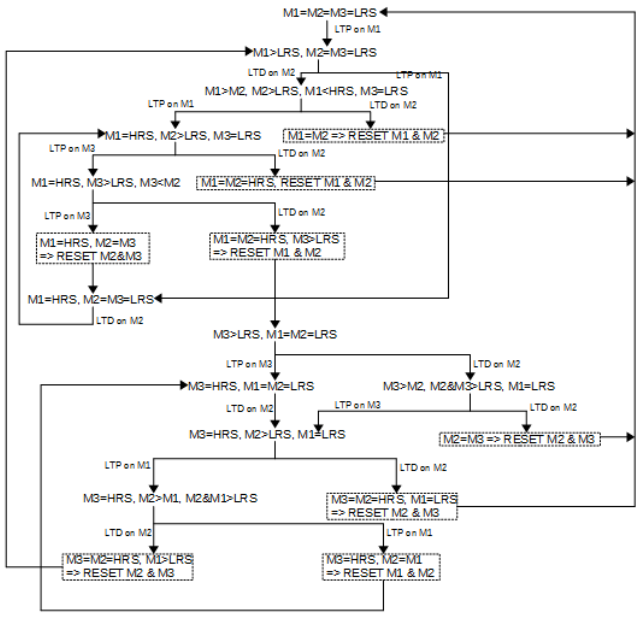

Figure 5: State change diagram for the quad-memristor synapse showing all the operational and transitional states.

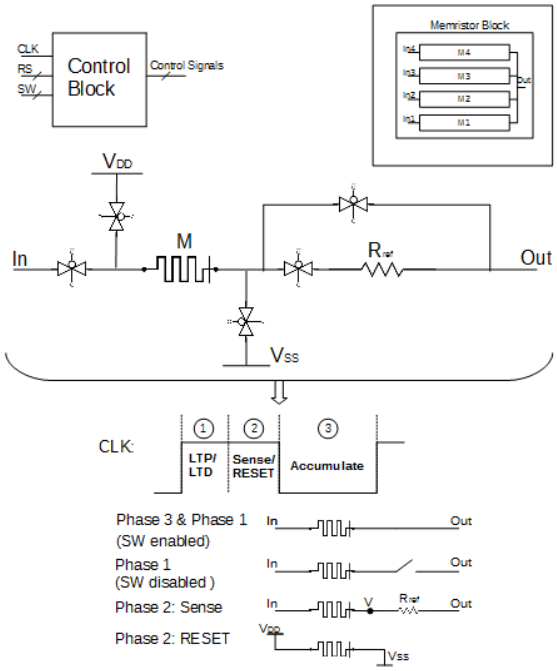

Figure 6: A single memristor circuit in the memristor block.

Starting with an initial weight, when an LTP/LTD is performed on the synapse, it can have several states as shown in Figure 5. Here, starting with a positive weight, the state diagram shows which of the memristors is to be selected for an LTP/LTD event. After a weight change operation, a new state results and a new pair of memristors are to be selected for LTP/LTD. In this process, whenever the memristances of two devices (*not* in their LRS state) are equal, they are $RESET$ to LRS. This leads the synapse into one of the other "Operational States." The states on which a $RESET$ operation is performed are named the "Transitional States." Figure 5 shows the transitions for positive weights. An analogous diagram can be drawn for negative weights.

3.2 Circuit for the quad-memristor synapse

Figure 6 shows a single memristor circuit inside the memristor block. Four such memristor circuits make up the memristor block as shown in the inset to the figure. As shown in Figure 6, the operation of the synapse is split into three phases. The control block for the memristor is used to enable the appropriate transmission gates of the memristor circuit so as to form the various circuit configurations as needed for the three phases of operation.

LTP/LTD phase: In this phase, the synapse's weight changes. Each memristor is assigned a signal SW, to determine if its state should change during this phase. Only those devices whose SW signal is enabled have their state changed.

Sense and RESET: This phase is used to ascertain the state of the synapse and determine which pair of devices to use in case of a weight change operation. In this phase, a voltage divider network is formed as shown in Figure 6. The sensed voltage V, from the divider of all the four memristors drives the main control block. The control block compares these voltages against reference voltages V_{ref0} and V_{ref1}, which represent LRS and HRS states for the memristor, respectively. This comparison is accomplished using a differential voltage-mode comparator.

Using the digital outputs of the analog comparators, the control block ascertains the state of the synapse and outputs SW, to assert the memristors to be used for the next LTP/LTD phase. If the main control block senses the current state of the synapse to be a transitional state, it enables the appropriate RS signal to $RESET$ the pertinent pair of memristors to LRS. A $RESET$ operation is again followed by a sensing phase to sense the new operational state that the synapse has attained (after the $RESET$ operation).

Accumulation: In this phase no state change occurs. Current flows through the synapse proportionate to its weight and accumulates charge on the post-synaptic neuron.

3.3 Simulated Synaptic Operation

In this section, we present some simulations of the synapse's weight change scheme described thus far.

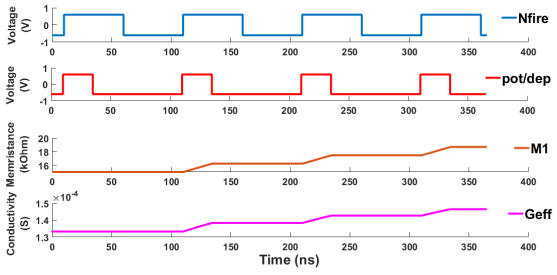

Figure 7: Basic LTP operation on the quad-memristor synapse.

Figure 7 shows the basic LTP operation being performed on the synapse. Here, the signal $Nfire$ indicates the firing

event of the post-synaptic neuron and the signal pot/dep indicates LTP/LTD. When pot/dep is asserted HIGH, it indicates potentiation and when it is LOW the weight is depressed. The synapse here is in the initial state of $M1 > LRS$ and $M2 = M3 = M4 = LRS$. With each $Nfire$ event, the pot/dep is asserted HIGH, indicating potentiation. Therefore, M1 is used for weight change.

Figure 8 shows the case when the synapse is initially in the state $M1, M2 > LRS$, $M1 > M2$ and $M3 = M4 = LRS$. In this case, when repeated LTD is performed on the synapse, at a certain stage, M2 attains the same memristance as that of M1. This is a transitional state, hence both M1 and M2 go to LRS, as shown in Figure 8. Further LTD operations result in the switching of M2.

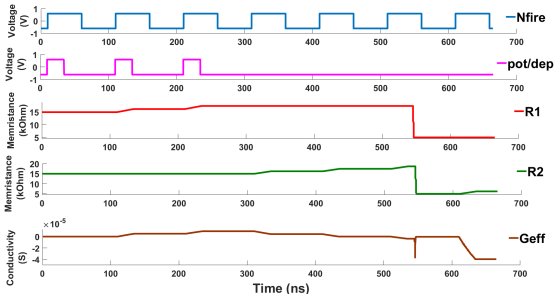

Figure 8: Reset operation performed when a transitional state is encountered.

Table 1: Metrics for comparison of the two synapses.

Metric	Twin-memristor	Quad-memristor
No. of MOS devices	80	1481
No. of memristors	2	4
Clock frequency used	10 MHz	10 MHz
Phases of operation	2	3
Average Power	45.35 μW	306.2 μW
Energy per Spike	4.558 pJ	30.6 pJ

4. COMPARISON AND DISCUSSION

Table 1 shows some metrics for the two synapses. Obviously, the quad-memristor synapse has a higher silicon footprint, consumes more power and has an extra phase of operation (the sense/$RESET$ phase). However, it may be noted that the quad-memristor design addresses a major (often neglected) issue with respect to the switching of some memristors: asymmetric switching rates. The analysis of these synapses provokes the following conclusions:

1) Better memristive device stacks need to be engineered to reduce the mismatch in the switching times.

2) In the wake of having to design synapses with contemporary devices having disproportionate switching rates, circuit designers could leverage our scheme of operation for the quad-memristor synapse to realize such synapses.

5. CONCLUSION

In this paper, we have considered an issue that is fundamental to the physics of switching in the memristor, the mismatch in the switching rates in opposite directions. Starting with the design of a twin-memristor synapse suited for online learning using memristors with identical switching rates, we have presented a quad-memristor synapse structure for use with devices whose switching rates differ by a margin as large as two orders of magnitude. We have also shown and compared the metrics of their circuit implementations.

6. ACKNOWLEDGMENTS

This material is based in part on research sponsored by Air Force Research Laboratory under agreement number FA8750-16-1-0065. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory or the U.S. Government.

7. REFERENCES

[1] K. Beckmann, J. Holt, H. Manem, J. Van Nostrand, and N. C. Cady. Nanoscale hafnium oxide rram devices exhibit pulse dependent behavior and multi-level resistance capability. *MRS Advances*, pages 1–6, 2016.

[2] Y. H. Do, J. S. Kwak, Y. C. Bae, K. Jung, H. Im, and J. P. Hong. Hysteretic bipolar resistive switching characteristics in tio2/tio2-x multilayer homojunctions. *Applied Physics Letters*, 95(9):093507–093507, 2009.

[3] N. Hashem and S. Das. Switching-time analysis of binary-oxide memristors via a nonlinear model. *Applied Physics Letters*, 100(26):262106, 2012.

[4] M. Hu, H. Li, Y. Chen, Q. Wu, G. S. Rose, and R. W. Linderman. Memristor crossbar-based neuromorphic computing system: A case study. *IEEE Trans. Neural Netw. Learning Syst.*, 25(10):1864–1878, October 2014.

[5] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu. Nanoscale memristor device as synapse in neuromorphic systems. *Nano letters*, 10(4):1297–1301, 2010.

[6] I. Kataeva, F. Merrikh-Bayat, E. Zamanidoost, and D. Strukov. Efficient training algorithms for neural networks based on memristive crossbar circuits. In *2015 International Joint Conference on Neural Networks, IJCNN*, pages 1–8, July 2015.

[7] C. E. Merkel and D. Kudithipudi. A current-mode cmos/memristor hybrid implementation of an extreme learning machine. In *Great Lakes Symposium on VLSI 2014, GLSVLSI '14, Houston, TX, USA - May 21 - 23, 2014*, pages 241–242, 2014.

[8] G. S. Rose, H. Manem, J. Rajendran, R. Karri, and R. Pino. Leveraging memristive systems in the construction of digital logic circuits. *Proceedings of the IEEE*, 100(6):2033–2049, June 2012.

[9] T. Serrano-Gotarredona, T. Masquelier, T. Prodromakis, G. Indiveri, and B. Linares-Barranco. Stdp and stdp variations with memristors for spiking neuromorphic learning systems. *Frontiers in neuroscience*, 7:2, 2013.

[10] G. S. Snider. Spike-timing-dependent learning in memristive nanodevices. In *2008 IEEE International Symposium on Nanoscale Architectures*, pages 85–92. IEEE, 2008.