Investigation of DW Spintronic Memristor performance in 2T1M Neuromorphic Synapse

Yasmin K. Abdelmagid
Nanotechnology and Nanoelectronics
Zewail City of Science and Technology
Cairo, Egypt
s-yasmin.khalil@zewailcity.edu.eg

Renad T. Nawar
Nanotechnology and Nanoelectronics
Zewail City of Science and Technology
Cairo, Egypt
s-renad.nawar@zewailcity.edu.eg

Mennatullah K. Rabie
Nanotechnology and Nanoelectronics
Zewail City of Science and Technology
Cairo, Egypt
s-mennah.khaled@zewailcity.edu.eg

Ahmed S. Tulan
Electronics and Communications Engineering Department, Cairo University, Giza 12613, Egypt.
Cairo, Egypt
Ahmedtulan@gmail.com

Ahmed H. Hassan
Electronics and Communications Engineering Department, Cairo University, Giza 12613, Egypt.
Giza, Egypt
ahhamed93@hotmail.com

Andoleet Saleh
Nanotechnology Department at Zewail City for Science and Technology,
Cairo, Egypt.
s-andoleet@zewailcity.edu.eg

Hassan Mostafa
Electronics and Communications Engineering Department, Cairo University, Giza 12613, Egypt
University of Science and Technology,
Nanotechnology and Nanoelectronics Program,
Zewail City of Science and Technology,
October Gardens, 6th of October, Giza 12578, Egypt
hmostafa@uwaterloo.ca

***Abstract*—Memristor, the two-terminal memory-resistance device discovered by Chua in 1971, is a promising solution for future processing problems. Its CMOS integration compatibility and large resistance in small size, makes it very successful candidate for large-scale systems like Neural Networks. In last decade, memristors were used in many Neuromorphic Synapses for its advantage of combining processing (dot-product) and memory in same device. There are different materials that can be used to fabricate memristors. In this paper, a comparison between spintronic and TiO2-resistive memristor in two-transistors-one memristor synapse, is introduced. The work was done on Cadence Virtuoso with using Verilog-A for memristor modeling. The comparison reveals that the synaptic implementation with a spintronic memristor is more efficient when high speed is needed. However, the resistive memristor is more adequate due to its lower power dissipation**.

***Keywords*—Spintronic Memristor, Neuromorphic Synapse, IoT Hardware Acceleration, Low-Power Circuits**

I. INTRODUCTION

Modern systems like Internet of Things (IoT) [1,2] need a large design effort for processing near sensor. This requires small scale and low power consumption, which hard to be found by von Neuman processors in small area near sensor [3, 4]. Neuromorphic computing is the solution for this as it is less in size and power consumption [5]. It is the main part of Neural Networks hardware. There are different types of Neural Networks (NNs), depending on the application they used for. In Image Recognition, Convolutional Neural Networks (CNN) are the most popular [6,7]. Spiking NNs (SNN), are the most type that mimic the human neural activity in terms of time difference between neurons [8]. Also, Deep NNs (DNN), which has many layers of neurons and synapses to increase accuracy (maybe CNN or SNN) [9]. Memristor is widely used in neuromorphic synapse. Its main advantage is that it makes the dot-product of machine learning algorithms (between inputs and weights) and stores the weights of the synapse in one device [10]. There are different configurations for synapses with memristors [3], and one should choose between them depending on his system function and scale required. The one memristor
crossbar array was firstly used [11]. Its disadvantage is that it suffers from sneak path problem and don't have negative weights capability. Another configuration is with two memristor crossbar [12], which achieves negative weights but with double the size of 1M. The third configuration is 2M with one resistor (2M1R), which is useful in Spiking Neural Networks [13]. Another configuration is with adding two transistors to control the memristor synapse, called 2T1M Synapse [10]. This configuration allows on-chip control rather than memristor only synapses, which require additional control circuit and disconnection of memristor during control from the crossbar array [3]. However, it is not suitable for large-scale neural networks and requires that the memristor characteristic to be in linear region of operation. This paper will focus only on 2T1M synapse, and future work will be on the other types.

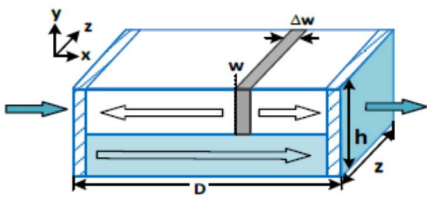

Fig. 1. Structure of DW Spintronic Memristor in [22] with domain wall width, Δw and position, w. White layer is free, while blue is fixed layer.

Fig. 2. Equivalent circuit model of DW spintronic memristor

There are two types of memristors that can be used in 2T1M synapse: Domain-Wall (DW) Spintronic Memristor [14] and Resistive Memristor [15]. Many materials can be the active layer of resistive memristors, TiO2, for example, is used in [10, 16]. A comparison between the proposed work with spintronic and resistive memristor in [10, 16] is illustrated in section V.

The spintronic memristor [17, 18-21], an alternate physical realization of HP's original memristor, relies on changing the position of its domain wall via a moving current to modify its memristance state. Unlike the resistive memristor, system resistivity is controlled by the magnetization direction of the layers. Despite having a lower Ron/Roff [Rp/Rap] ratio, which is further explored in section IV, the spintronic memristor possesses a higher data density, faster response and can be readily integrated into CMOS technology. In this paper, we incorporate the spintronic memristor model in [14] into the synaptic circuit and compare its performance with the results yielded in [10, 16].

The rest of the paper is organized as follows, the domain wall (DW) spintronic memristor background, is introduced in section II. The 2T1M synapse used in [10] for applying Online gradient descent algorithm in MNN, is introduced in section III. The simulation results of using spintronic memristor in the synapse are illustrated in section IV. Finally, A comparison between using spintronic memristor and the resistive alternative in [16] in 2T1M synapse is illustrated in Section V.

II. DW Spintronic Memristor

Domain Wall (DW) Spintronic Memristor [14, 18, 22] is composed of two ferromagnetic layers; the pinned (or reference) layer, with a fixed magnetization direction, and the free layer, where the magnetization changes according to the spin-transfer torque effect. The free layer is divided into two segments: the first with a magnetization direction parallel to the reference layer while the second is antiparallel. These segments are separated by a domain wall whose position determines the memristance state of the memristor through the relative proportion between the lengths of both parallel and antiparallel segments, given that the parallel part is of lower resistance, see Fig. 1. In case all magnetization is parallel (P) to the fixed layer, the memristor has Ron (Rp), while if antiparallel (AP), the memristor in Roff (Rap). In the middle of that range, the memristance of the device will be:

$$M(\alpha) = R_p\left(\beta - \frac{1}{2N}\right) + R_{DW} + R_{AP}\left(1 - \beta - \frac{1}{2N}\right) \quad (1)$$

Where R_{DW} is the resistance of the domain wall (which is constant but shifted in position based on the external bias direction). β is displacement factor of domain wall, see Fig. 2. From Fig. 1, β=w/D, 0≤β≤1.

Where

$$N = \frac{D}{\Delta w} \quad (2)$$

D is the length of the device, and Δw is the domain wall width [14], see Fig. 2.

III. 2T1M Neuromorphic Synapse

The 2T1M synapse schematic is declared in Fig. 3. The synapse contains memristor for processing (dot product implementation) between input attribute (X) of the machine learning data and the weight stored (WT) in the same memristor. So, memristor acts as memory and processing element, which declares the advantage of using it in neuromorphic computing (small size) [10]. Two MOS Transistors (P-Channel and N-Channel MOSFETs) are used to control whether the synapse in reading or updating of weight stored. The PMOS connect negative voltage to the memristor while the NMOS connects positive voltage to the memristor.

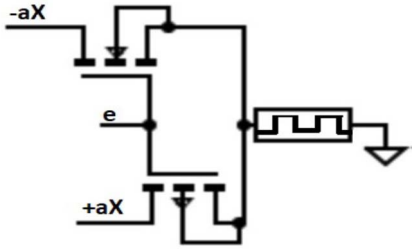

Fig. 3. 2T1M Synapse schematic in [10]

Both transistors conductance should be very higher than the memristor one to ensure that the voltage signal at input of transistors (aX or -aX) reaches memristor without loss in amplitude due to transistor resistance [10]. To achieve this, the input attribute (X) is scaled down to 1mV by scaling factor (a) and gate of transistors are connected high voltage (1.8V) compared to (aX) (deep triode region of operation) [10].

The read and write control signal (gate voltage of transistors) used in [10] is illustrated in Fig. 4 (a). The signal is containing two consecutive phases, read and write, with overall 100ms period. Read phase is symmetric voltage signals in time with opposite voltage polarity, so no destructive reading of the state stored happens. The write phase voltage polarity and duration are determined by the error signal feedback (y) in the machine learning algorithm [10] between desired (d) output and actual output (r) of the dot-product operation mentioned earlier.

$$y = d - r \quad (3)$$

$$r = WT.X \quad (4)$$

The corresponding change in memristance due to the original signal is shown in Fig. 4(b). The read phase make no change to the memristance state, then write phase updates the weight value to new one based on Eq. 3.

The read phase in Fig. 4 (a) is based on ideal pulse signal, positive and negative halves are equal in time. However, this can not be achieved practically with any signal generator [16]. Accordingly, a mismatch, α, between the two halves will make the reading operation destructive, see Fig. 5. Due

to this destructive reading, authors in [16] made two new control signals (A and B) that are less destructive than the one in Fig. 4(a). Both signals have same read and write phases duration as in original signal of Fig. 4(a). However, the read phase is not divided to equal signal duration with opposite polarities, instead configured with different durations and voltage polarities, see Fig. 6. The two signals were applied to Resistive Memristor (TiO2) in [16] and achieved less destructive reading compared to original signal in [10].

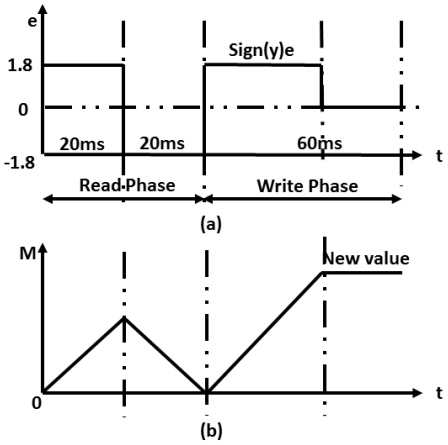

Fig. 4. (a) The original control signal applied to the synapse in [10, 16]. The read phase is in first 40ms and write phase is the rest 60ms. (b) The change in the memristance state (M) due to signal in (a).

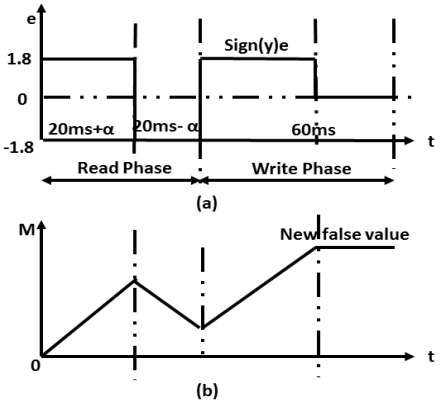

Fig. 5. (a) The original control signal with mismatch in reading phase applied to the synapse in [10, 16]. The read phase is in first 40ms and write phase is the rest 60ms. (b) The change in the memristance state (M) due to signal in (a). Note that weight update is at false value compared to Fig. 4(b).

In [16], the failure of the synapse in the 2T1M synapse was defined as writing weight value in memristor that should be written in next cycle of control signal, see Fig. 7. The authors in [16] increased the number of cycles to fail (n, see Fig. 7) to very high value, 310T, with pulse width mismatch α=10% (which cannot even happen practically), and this was achievement. It should be noted that this high number of cycles to fail is also due to large Ron/Roff of resistive memristors, while in spintronic memristor this number is much less, as Rp/Rap is lower than in resistive memristor.

I. Simulation Verification

In this paper, the resistive memristor model in [10, 16] that used in 2T1M synapse is replaced by the DW spintronic memristor model in [14]. The investigation of the new memristor in the synapse performance is done in two steps. First step is to study the I-V characteristics of the DW spintronic memristor and extract the conditions in which the memristor will be in linear operation (as required for 2T1M synapse, see Section III). The second step is to investigate the behavior of applying different control signals, that were mentioned in Section III, on the memristance change.

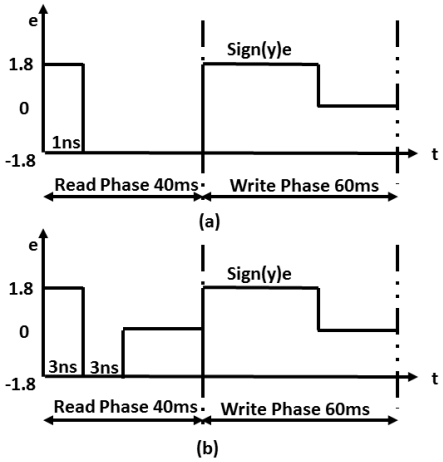

Fig. 6. The two control signals in [16]. (a) is signal A and (b) is signal B.

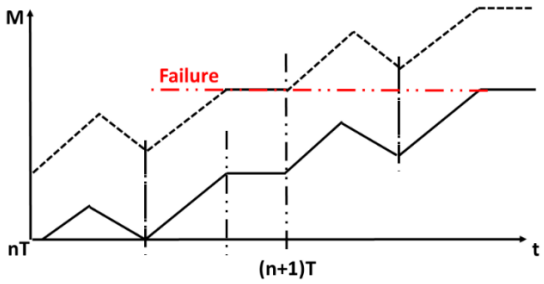

Fig. 7. Failure of writing weight in memristor due to destructive reading signal. nT means after n full periods (read and write phases) of operation. Solid line for memristance change due to original signal and dash line for memristance with destructive reading signal.

A. I-V Characteristics of DW Spintronic Memristor

In [14], the I-V characteristic curves showed that by increasing the applied signal's frequency, the characteristics goes to be more linear. Frequencies above 20 MHz gives good linearity that suitable for the 2T1M synapse. The authors also stated that changing the state variable (position of domain wall) from R_{AP} to Rp is much easier than changing from Rp to R_{AP} as the critical current needed to make memristor state variable change is smaller in R_{AP} to Rp change. This will greatly affect the results in this paper, as the voltage applied to the memristor (accordingly, the current) is very small (10mV), as shown in Table II. So, in this paper, the state variable is made to change from R_{AP} to Rp.

B. Control Signals Study

According to the results obtained from previous subsection, the original signal in Fig. 4(a) [10] will not work well with DW spintronic memristor, as its frequency is less than 20 MHz (1/40ms or 25Hz).

Both signals (A and B) in [16] worked well as their highest frequency (read phase) is 300MHz, much larger than 20MHz, however, with smalling the write phase period to 50ns instead of 60ms to be high frequency as well (write phase

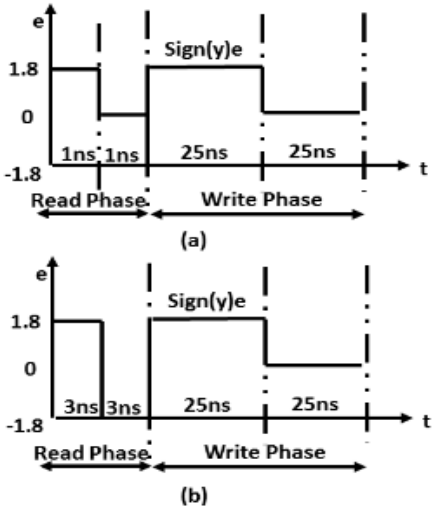

Fig. 8. The modified two control signals in [16]. (a) is signal A and (b) is signal B.

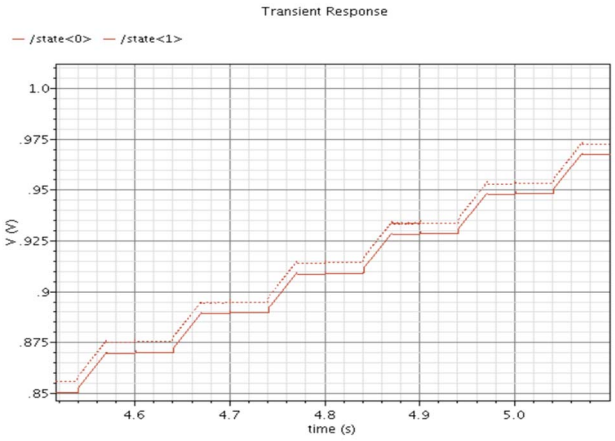

Fig. 9. The state variable change after applying modified signal A. Note the slight change of state variable with no failure happens. State<0> and State<1> are for signal without and with mismatch, respectively.

frequency=50MHz). The new modified A and B signals are shown in Fig. 8.

Figures 9 and 10 show the result of applying the new control signals in Fig. 8 on the state variable (or memductance) with and without read signal mismatch (10% mismatch [16]) stated in Section III. Signal A has much destructive effect than signal B, as the read signal is not symmetric (only one positive pulse). However, no failure happened because the small pulse duration makes slight change in the state variable, which will make failure after large number of cycles. In spintronic memristor, large number of cycles can not happen because of small R_{AP}/Rp ratio (a few cycles change state variable from R_{AP} to Rp), and this is advantage over resistive memristor [14]. This means no failure will happen at all, however, a RESET procedure should be applied on the spintronic memristor after state variable changed from R_{AP} to Rp to reuse it in the synapse.

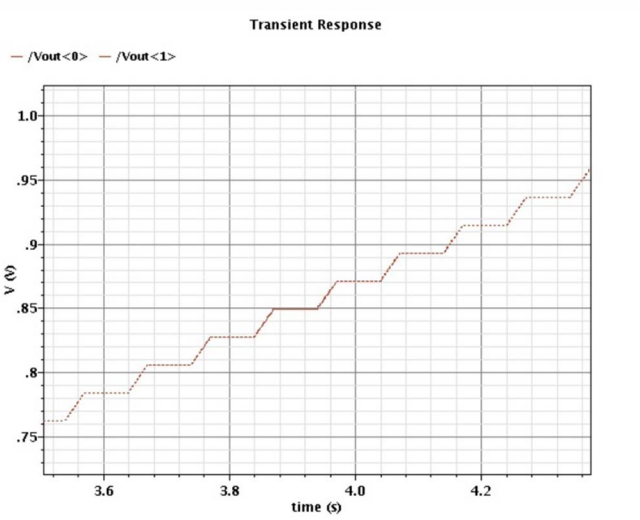

Fig. 10. The state variable change after applying modified signal B. Note that both signal with and without mismatch have almost same effect on state variable.

TABLE II: CIRCUIT PARAMETERS

	Parameter	Value	Description
Memristor	R_P	1.5kΩ	Parallel Resistance
	R_{AP}	6kΩ	Anti-Parallel Resistance
	N	8000	No. of steps of state variable changes
	D	500nm	Length
	z	60nm	Width
Transistors	W	2um	Width
	L	130nm	Length
Input	a	10mV	Scale
	X	10	ML Input

All circuit parameters used are stated in Table II. The other parameters, like R_{DW} and Δw, are calculated by simulator. The memristor dimensions are taken from [14] as well as all other device physical parameters.

II. COMPARISON BETWEEN SPINTRONIC AND RESISTIVE MEMRISTOR

Both spintronic and resistive memristors can be used in 2T1M synapse, however, with different control signals. This is due to the difference in nature and physics of both devices. Spintronic memristor, for example, can be used when high speed is requires, however, with higher power consumption than resistive memristor. On the other hand, resistive memristor does not need RESET operation due to large Ron/Roff ratio. Table III give a summary of advantages and disadvantages of both memristors in the synapse.

TABLE III: COMPARISON BETWEEN SPINTRONIC AND RESISTIVE MEMRISTORS IN 2T1M SYNAPSE

	Spintronic	Resistive
Speed	high	low
Power consumption	high	low
RESET	required	Not required
Range of Weight values	small	large
Size	Small (Nanoscale)	Small (Nanoscale)

ACKNOWLEDGMENT

This work was partially funded by ONE Lab at Zewail City of Science and Technology and Cairo University, NTRA, ITIDA, ASRT, Mentor Graphics, NSERC.

REFERENCES

[1] I. Lee and K. Lee, "The Internet of Things (IoT): Applications, investments, and challenges for enterprises," Business Horizons, vol. 58, no. 4, pp. 431–440.

[2] S. Madakam, R. Ramaswamy, and S. Tripathi, "Internet of Things (IoT): A Literature Review," Journal of Computer and Communications, vol. 03, no. 05, pp. 164–173.

[3] O. Krestinskaya, A. P. James and L. O. Chua, "Neuromemristive Circuits for Edge Computing: A Review," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 1, pp. 4-23, Jan. 2020.

[4] C. Frenkel, M. Lefebvre, J. Legat and D. Bol, "A 0.086-mm2 12.7-pJ/SOP 64k-Synapse 256-Neuron Online-Learning Digital Spiking Neuromorphic Processor in 28-nm CMOS," in IEEE Transactions on Biomedical Circuits and Systems, vol. 13, no. 1, pp. 145-68, Feb. 2019.

[5] Y. Jiang, C. Yang, J. Na, G. Li, Y. Li, and J. Zhong, "A Brief Review of Neural Networks Based Learning and Control and Their Applications for Robots," Complexity, vol. 2017, pp. 1–14, 2017.

[6] N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," 2017 International Conference on Communication and Signal Processing (ICCSP), 2017.

[7] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X.Wang, G. Wang, J. Cai, and T. Chen, "Recent advances in convolutional neural networks," Pattern Recognition, vol. 77, pp. 354–377, 2018.

[8] M. Pfeiffer and T. Pfeil, "Deep Learning With Spiking Neurons: Opportunities and Challenges," Frontiers in Neuroscience, vol. 12, 2018.

[9] A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," IEEE Access, vol. 7, pp. 53040–53065, 2019.

[10] D. Soudry, D. Di Castro, A. Gal, A. Kolodny and S. Kvatinsky,"Memristor-Based Multilayer Neural Networks with online gradient descent training," in IEEE Transactions on Neural Networks and Learning Systems, vol. 26, no. 10, pp. 2408-2421, Oct. 2006.

[11] Y. Zhang, X. Wang, and E. G. Friedman, "Memristor-based circuit design for multilayer neural networks," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 65, no. 2, pp. 677–686, 2018.

[12] M. Prezioso, F. Merrikh-Bayat, B. Hoskins, G. Adam, K. K. Likharev, and D. B. Strukov, "Training and operation of an integrated neuromorphic network based on metal-oxide memristors," Nature, vol. 521, no. 7550, p. 61, 206.

[13] M. Hu, Y. Chen, J. J. Yang, Y. Wang, and H. H. Li, "A compact memristor-based dynamic synapse for spiking neural networks," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 36, no. 8, pp. 1353–1366, Aug 2017.

[14] S. F. Nafea, A. A. Dessouki, S. El-Rabaie, B. E. Elnaghi, Y. Ismail, and H. Mostafa, "An accurate model of domain-wall-based spintronic memristor," Integration, vol. 65, pp. 149–162, 2019.

[15] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," Nature, vol. 453, no. 7191, pp. 80–83, Mar. 2008.

[16] A. H. Hassan and H. Mostafa, "Nondestructive Reading and Refreshment Circuit for Memristor-based Neuromorphic Synapse," *2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS)*, Dallas, TX, USA, 2019, pp. 642-645.

[17] M. Elshamy, H. Mostafa, and M. S. Said, "Comparative review of the TiO2 and the spintronic memristor devices," 2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE), 2014.

[18] S. F. Nafea, , A. A. S. Dessouki, S. El-Rabaie, B. E. Elnaghi, Y. Ismail, and H. Mostafa, "An Accurate Model of Domain-Wall-Based Spintronic Memristor", Elsevier Integration VLSI Journal, vol. 65, pp. 149-162, 2019.

[19] H. Mostafa, and Y. Ismail, "Process Variation Aware Design of Multi-Valued Spintronic Memristor-Based Memory Arrays", IEEE Transactions on Semiconductor Manufacturing (TSM), vol. 29, issue 2, pp. 145-152, 2016.

[20] S. F. Nafea, A. A. S. Dessouki, S. El-Rabaie, Kh.El-Barbary, and H. Mostafa, "Read Disturbance and Temperature Variation Aware Spintronic Memristor Model", IEEE Canadian Conference on Electrical and Computer Engineering (CCECE 2016), Vancouver, Canada, pp. 1-4, 2016.

[21] M. Abdallah, H. Mostafa, and M. Fathy, "Yield Optimization of Spintronic Memristor Based Memory Arrays", IEEE International Conference on Electronics, Circuits, and Systems (ICECS 2015), Cairo Egypt, pp. 523-526, 2015.

[22] Nafea, Sherif F., et al. "Area-Efficient Read/Write Circuit for Spintronic Memristor Based Memories." 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), 2017.