Practical Implementation of 2T1M Neuromorphic synapse with spintronic memristor and refreshment circuit

AHMED H. HASSAN, HASSAN MOSTAFA, AMIN M. NASSAR, MOHAMED M. RIAD

Electronics and Communication Engineering, Cairo University, Egypt

email: ahmed.hassan.j@eng-st.cu.edu.eg

Abstract

This work is a practical implementation of the 2Transistors and 1Memristor (2T1M) Neuromorphic synapse by using spintronic as a practical device of memristor rather than theoretical linear models. Due to the destructive nature of the reading signals on the memductance of memristor, a new reading signal that reduces the destruction and suitable for spintronic linearity of operation is proposed. The destruction on memductance still happens but after 1600 cycles of proposed reading signal, which is very high number compared to traditional reading signals in literature. Implementation of refreshment circuit to adjust errors in memductance based on destructive nature of reading signal is applied. All subcircuits of the refreshment circuit are practical components instead of ideal components used before in literature. That practical implementation of the refreshment circuit showed high speed and accuracy not far from the one with ideal components.

Keywords: Neuromorphic computing, In-memory computing IMC, Spintronic memristor, Memristive Synapse, Refreshment Circuit.

1. Introduction

Neuromorphic computing is considered the future of computing mechanism for Artificial Intelligent (AI) and Neural networks [1]. The more advanced features the AI implements, the more speed and low power the computing system needs to be [2]. The In-Memory Computing (IMC) is the most attractive computer structure that achieves minimum size and low power demands of AI applications [3,4] compared to von-Neuman architectures.

Memristor is the most common device used for IMC due to its ability of processing and storage of Machine Learning (ML) weights in Neuromorphic Synapses [5] at same place. One of the popular memristor devices used in Neuromorphic computing is the Spintronic memristor. The most advantage of the spintronic memristor is the wide resistance range up to Mega ohms compared to other memristors [1, 6].

Many architectural schemes introduced in literature to present the Synapse part of the AI Neuromorphic system with use of IMC feature [7-12]. Some used 1 Memristor (1M) crossbar arrays [10], others used resistors with the memristor like (1M1R) [11,12] and some used transistors with memristors for controlling weight polarity like 2T1M [7-9].

In this work, the 2 Transistors 1 Memristor (2T1M) synapse in [7-9] is used with a spintronic memristor model in [14] instead of theoretical linear models. Spintronic memristor is practical and widely used in many Neuromorphic systems [15-19]. As discussed in [13] and

in section 4 in this work, the reading signal that contains equal consecutive positive and negative pulses always has mismatch in real implementation and a refreshment circuit is highly needed for the 2T1M synapse [8]. So, a practical implementation of the refreshment circuit in [8] is introduced. Finally, application of the practical refreshment circuit on the 2T1M spintronic synapse is illustrated by simulations and showed good practical approach of refresh after high number of cycles of operation.

2. Theoretical Background:

2.1 2T1M Neuromorphic synapse

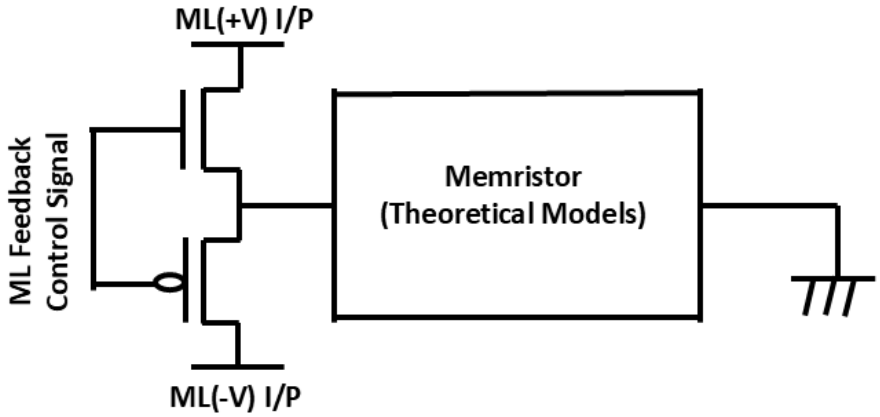


Figure 1. Neuromorphic synapse using memristor used in literature [7-9,13]. ML(+/-V)I/P are the positive and negative machine learning inputs to the synapse. ML Feedback Control Signal is the error signal feedback from machine learning algorithm to adapt (or keep) the weight value stored in memristor conductance (memductance/state variable)

The 2T1M Synapse structure used in [7] is shown in Fig. 1. The memristor is the element responsible for dot product computing of both weight and machine learning value. It also stores the weight value to be read in reading phase. The two transistors are used for controlling positive and negative weights. Their conductance should be very high compared to the memductance to avoid loss of accurate weight value by voltage drop on them. Compared to other memristor structures like 1M [20], 4T1M [21] and 2M [22] synapses, it achieves the polarity of weight functionality with suitable area size and not suffering from severe sneak path issues. The other main advantage of 2T1M structure is that it enables on-chip control rather than other structures [9]. It is used in Multi-Input-Multi-Output (MIMO) modern systems as in [23] and Magnetic-Random-Access-Memory (MRAM) [24].

The memristor used in [7] is linear model along with empirical model for nonlinear model. Those models are not practical, as they are theoretical and early-stage assumptions of the memristor behavior. In this work, a physical model [14] for spintronic memristor (practical device) is used [9] alongside applying refreshment circuit [8] for destructive reading nature of any function generator in lab [13], see section IV.

2.2 Spintronic Memristor

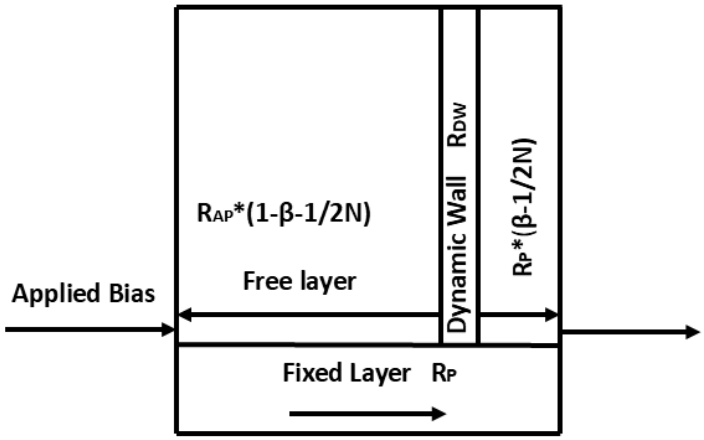


Figure 2. Spintronic Memristor structure [14]. β is the displacement factor of the Dynamic Wall (DW) and N is the No. of steps of state variable changes

The spintronic memristor structure [14] and electrical model are shown in Fig. 2. The device consists of two layers of ferromagnetic materials, fixed and free layers. The fixed layer is constant in resistance while the free one is variable in resistance based on the bias applied to the device. The applied bias changes the dynamic wall (DW) position, so the total resistance changes of the free layer. This is because the free layer is divided into three parts: Parallel, Anti-parallel and DW resistances. The parallel term means its spin direction is parallel to the fixed layer one. The dynamic wall resistance RDW is constant but moved based on applied bias.

The model used to simulate the real spintronic memristor is illustrated in [14]. Table 1 declares all model parameters' value and their description. Other physical parameters for material are used from [14].

The following is a brief for the important equations used from model in [14] to define the resistances of the electrical model R_{AP} and R_{DW}. The Tunneling Magnetoresistance Ratio TMR for the spintronic memristor is function of applied bias and given by [14]

$$TMR = TMR_0 \frac{1}{1+(V/V_h)^2} \quad (1)$$

Where TMR_0 the TMR at 0 Volt and 25°C, V_h is a voltage dependance parameter for TMR [14] and V is the applied bias voltage. The fixed layer resistance R_P is defined in Table 1 to be 1.5 kΩ and used in the following equations [14]. The dynamic wall resistance has two equations, one for parallel direction (R_{P_DW}) and another for anti-parallel (R_{AP_DW}) based on polarity of bias. Note that the latter is a function of bias based on Eq. 1, which makes the model in [14] more accurate than others in literature.

$$R_{P\_DW} = R_P/N \quad (2)$$

$$R_{AP_{DW}}(V) = R_P(1 + TMR) \quad (3)$$

The previous equation is general for Anti-parallel (AP) resistance, so used to calculate AP resistance part R_{AP}.

The total resistance of the memristor (memristance M) [14] as a function of bias and displacement β will be

$$M(\beta, V) = R_{AP}\left(1 - \beta - \frac{1}{2N}\right) + R_{DW} + R_P(\beta - \frac{1}{2N}) \quad (4)$$

Where R_{AP} is taken from Eq. 3 and R_{DW} taken from one of Eq. 2 or 3 based on bias direction. The equations are calculated internally in simulator by model in [14].

Table I: Spintronic Memristor Model Parameters

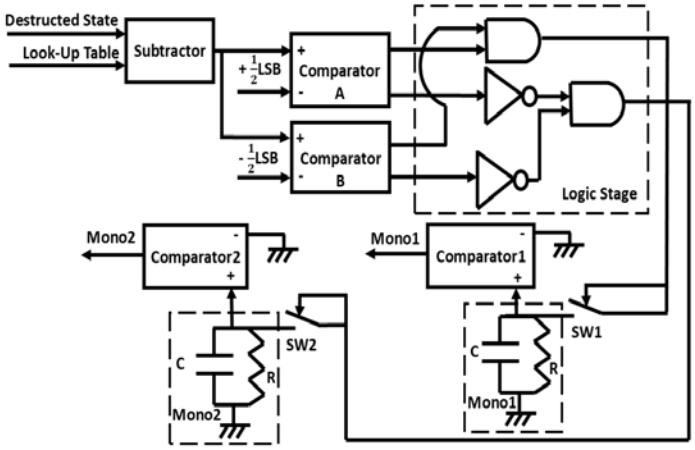
Parameter	Value	Unit	Description
β	0.002		Damping Parameter
D	500	nm	Length
z	60	nm	Width
h	1.5	nm	Thickness
N	8000		No. of steps of state variable changes= D/Δw
R_P	1.5	kΩ	Parallel Resistance
Δw	0.6	nm	Dynamic wall width
TMR_0			TMR at 0V and 25℃
V_h	0.5	V	TMR voltage dependance

2.3 Destructive reading and need of refresh circuit

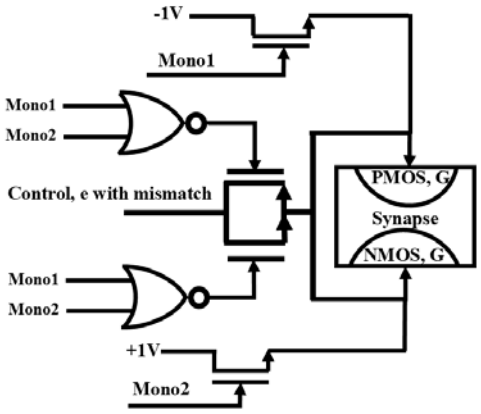
The method used in [7] to read the memrestance of the memristor is to apply consecutive positive and negative pulses that are equal in time duration. In this way, there will be no change on weight stored in the synapse. However, in any function generator in the world, there is a mismatch between positive and negative pulses in their time [13]. So, the authors in [13] proposed signals that are narrow in time width (in nano seconds) instead of milli-seconds like in [7]. However, the need for a refreshment circuit is highly needed in practical world, as to be free to apply any signal. In [8] the authors represented a refreshment circuit and applied it to the synapse in [7], but they used linear model of memristor which is not practical. The main advantage of the refreshment circuit in [8] is that it can operate online on a synapse. This means no need to disconnect the synapse or neuromorphic system from the application to make refresh for synapse weights or SET and RESET long processes like in [25,26]. In this work, the use of the refreshment circuit in [8] on the 2T1M synapse having spintronic memristor is investigated.

The refreshment circuit proposed in [8] is shown on Fig. 3. The core part in (a) contains subtractor as first stage to determine how much is the difference between the destructed memductance state from the Look-Up Table. The look-Up Table is simply the memductance states if no mismatch in the reading synapse occurs (ideal case). Based on the output of the subtractor, the comparator A or B will go high if the destructed state is higher or lower than the ideal case, respectively. The following logic stage is needed to control the two Mono circuits Mono1 and Mono2 while comparators 1 and 2 are used to convert the Mono circuit output to binary. The Mono circuits are used to give suitable time for the refreshment circuit to interact on the synapse, so their R and C are computed wisely to achieve suitable time constant. The control part in (b) contains 2 NOR gates and 4 Transistors to apply the core output from (a) on the affected synapse from destructive reading. The original synapse

structure in Fig. 1. Is the same, but with a new control signal rather than the ML feedback control signal.



(a)

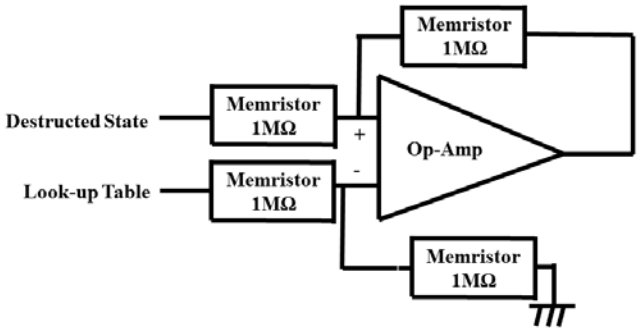


(b)

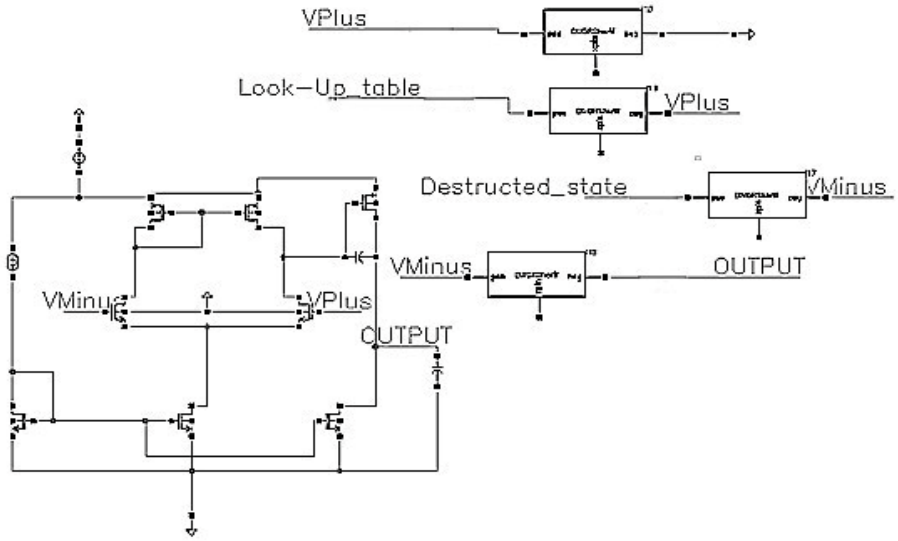
Figure 3. The refreshment circuit proposed in [8]. (a) the core part of refreshment. (b) the control of the core part on the 2T1M Synapse

3. The proposed practical implementation of the refreshment circuit

The subtractor and comparator used in [8] are ideal components in simulator. In this work, practical implementation of both subblocks is introduced. The subtractor shown in Fig. 4 (a) is simple op-amp with equal resistors in feedback connections. The op-amp implementation in Fig. 4(b) is simple 5T-OTA and Common-Source CS output stage with compensation capacitors [27]. Instead of using large resistance with 1MΩ, spintronic memristor is a good replacement for them. This is due its small size in CMOS fabrication process compared to resistors and large resistance values in MΩ is achieved [28]. The technology used is 45nm for transistors [27], and all their sizes are shown in Table. II. The design proposed for here is simpler than other ones in literature for analog op-amp designs, like folded cascode amplifiers [29]. This is because of privilege of memristors large resistance and small size.



(a)



(b)

Figure 4. Practical implementation of subtractor in refreshment circuit of [8]. (a) The top level of the subtractor, and (b) The implementation in cadence virtuoso using 5T-OTA and CS stage for Op-Amp implementation

Table II: Parameters used in Subtractor

Name	W	L	Function
M0 & M1	3um	500nm	Input pairs
M3	1um	1um	Tail current source
M2	12um	1um	Current mirror transistor
M5 & M6	7um	500nm	Load transistors for differential stage
M7	7um	500nm	PMOS CS input transistor
M4	75um	1um	NMOS current source for CS stage
C1	500fF		Compensation cap
C0	2pF		Output cap
VDD	1V		Supply voltage

The comparator used here as a practical implementation for [8] is StrongARM Latch [30,31] with using enable signal instead of running clk. The StrongARM latch is used instead of simple Asynchronous comparators, because of its speed and low power consumption [30]. However, it is used here as Asynchronous comparator, to prevent wrong output (VDD or high output) of pre-charge phase [30,31]. This is achieved by applying enable signal with low voltage (0V) for 1ns and (1V) for rest running time. In this way, the internal nodes of the latch are precharged for first 1ns, then enter amplification phase for time afterwards. This 1ns will not affect the operation of the synapse at all, as shown in section VI. The comparator is shown in Fig. 5 and full transistor sizes are illustrated in Table III.

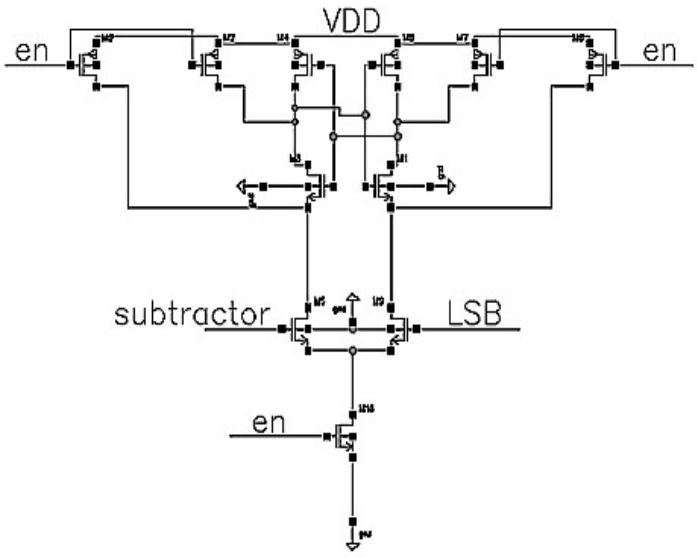


Figure 5. The practical implementation of comparators in [8] presented in this work using strongARM Latch architecture in [30,31]

From [8], the failure in writing state happens when the mismatch in the function generator writes value that should be in next writing phase. Fig. 6 shows the mismatch effect on state. For the refresh circuit to gain control on synapse and correct the value, the difference between the destructed state and the look-up table should be equal or greater than the value happens in Fig. 6. This value is called the Least Significant Bit (LSB).

Table III: Transistors parameters used in Comparator from [30]

Name	W	L	Unit	Function
M0 & M6	120	45	nm	Two PMOS switches outside
M2 & M7	270			Two PMOS switches inside
M4 & M8	120			Two PMOS of inverter
M10	180			Tail current source and NMOS switch
M3 & M4	210			Two NMOS of inverters
M5 & M9	180			Input Pairs
VDD	1			Supply Voltage
en	1 (delayed by 1ns)		V	Enable signal
Subtractor	From simulation			The subtractor stage output
LSB	±150		uV	The difference between destructed state and look-up table causes failure.

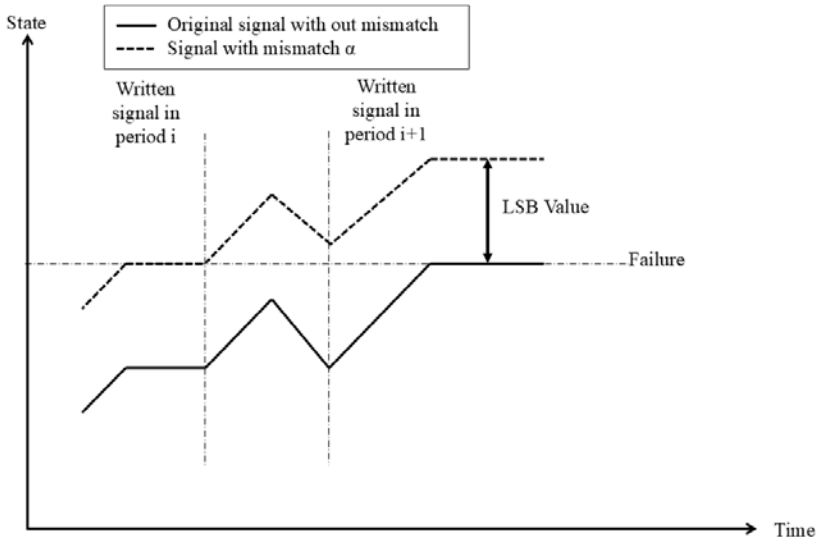


Figure 6. The failure in the destructed state that happens due to mismatch in function generator for positive and negative parts of any pulse's duration

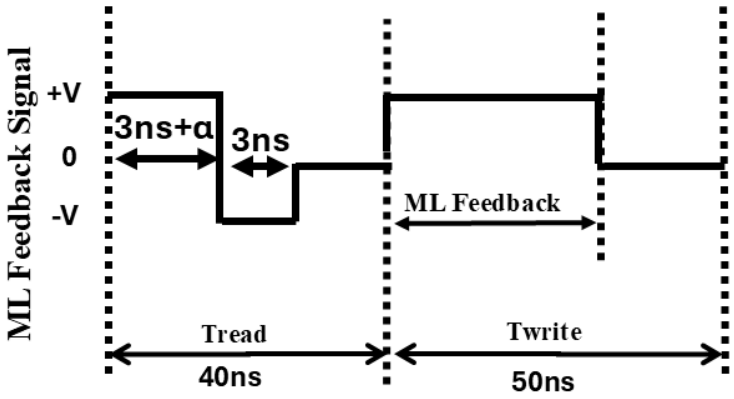


Figure 7. The ML control signal applied to the 2T1M spintronic memristor synapse. Tread is the read phase period and equals 40ns which corresponds to 20MHz that is suitable for spintronic linear region of operation [9]. α is the mismatch in read signal and equals 5% of the 50ns (2.5ns)

4. Application of refresh circuit on the synapse and simulation results

The refreshment circuit in [8] was applied on the 2T1M synapse in section II with spintronic memristor instead of linear model in [9] and with subtractor and comparator implementation in section IV. The applied control signal on the spintronic memristor synapse is shown in Fig. 7. From [9], the frequency of read signal should be equal or higher than 20MHz to let the spintronic memristor in linear region as required in 2T1M in [7]. The mismatch in read signal is declared in Fig. 7. By symbol (α) and has value 5% of read period (Tread).

The failure happened after 1600 cycles of the mentioned control signal in Fig. 7 and the refreshment circuit acted well after failure as shown in Fig. 8. It is better than other reading methods that require refresh after 60 cycles [32]. The state window in [14] model ranges between 0 and 1, it is measured in the model in volts, however it indicates the position of dynamic wall in the device. The PVT variations on the number of cycles to get failure is shown in Table IV. The power consumption of the refreshment circuit is 16.52mW after using

practical components for subtractor and comparator compared to 14.42mW in [8] with ideal components.

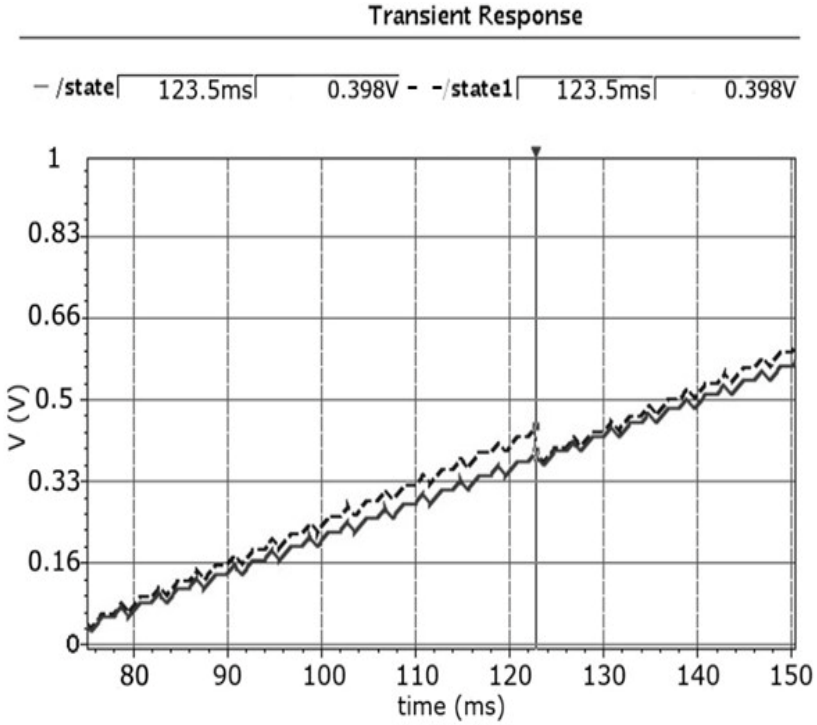


Figure 8. The refreshment circuit action on the destructed state in blue after failure happens as compared with the look-up table in red

Table IV: PVT variation effect on No. of cycles to get failure

Name	Variation	No. of cycles to get failure
Process	TT	1600
	SF	1611
	FS	1609
	SS	1612
	FF	1608
Voltage	0.9	1611
	1	1600
	1.1	1607
Temperature	-40°C	1610
	25°C	1600
	85°C	1607

5. Conclusion and future work

This work discussed the practical implementation of 2T1M neuromorphic synapse by using spintronic as memristor [14] instead of linear theoretical models [7,8]. A novel refreshment circuit in [8] is used to refresh the destructed memductance due to mismatch in reading signals, but with all subcircuits in transistor level implementation instead of ideal components used in [8]. The results agree with the ideal method in [8] in terms of functionality and speed response, however, with practical device used here instead of theoretical one in [8].

Future work will be a series of investigations of the proposed practical refreshment circuit in this work on other practical memristor models rather than spintronic ones. Also, the possibility of the refreshment circuit application on other synapse architectures will be discussed.

Data Availability

The datasets analyzed during the current study are available in the [Iris Flower] repository, [https://archive.ics.uci.edu/dataset/53/iris]. The methodology of applying the dataset in our work is presented in the manuscript by parameter (ML Input Signal) and as in [7].

For demonstration purposes and simplicity, the dataset was applied as PWL Signal source in simulator. The data used as column by column to test the hardware synapse. As the main purpose of our work is the hardware implementation of practical subcircuits and refreshment.

Future work will focus on different neural networks with large synapse grid, not single one as in this work.

Acknowledgement

The author would like to thank Cairo University facilities for their great contribution tools and labs to get this work done.

Conflict of Interest

The author declares no conflict of interest in the paper content or method used to generate results.

References

[1] Roy, Kaushik, et al. "Spintronic neural systems." *Nature Reviews Electrical Engineering* (2024): 1-16.

[2] Cui, Can, et al. "Spintronic Artificial Neurons Showing Integrate-and-Fire Behavior with Reliable Cycling Operation." *Nano Letters* 25.1 (2024): 361-367.

[3] Lone, Aijaz H., et al. "Multilayer ferromagnetic spintronic devices for neuromorphic computing applications." *Nanoscale* 16.26 (2024): 12431-12444.

[4] Liu, Long, et al. "Domain wall magnetic tunnel junction-based artificial synapses and neurons for all-spin neuromorphic hardware." *Nature Communications* 15.1 (2024): 4534.

[5] Duan, Xuegang, et al. "Memristor-based neuromorphic chips." *Advanced materials* 36.14 (2024): 2310704.

[6] Jaroš, Adam, et al. "Spinristor: A Spin-Filtering Memristor." *Advanced Electronic Materials* 9.8 (2023): 2300360.

[7] 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.

[8] Goda, Mai M., et al. "A novel refreshment circuit for 2T1M neuromorphic synapse." *Journal of Circuits, Systems and Computers* 31.03 (2022): 2250047.

[9] Y. K. Abdelmagid *et al*., "Investigation of DW Spintronic Memristor performance in 2T1M Neuromorphic Synapse," *2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, Giza, Egypt, 2020, pp. 573-577, doi: 10.1109/NILES50944.2020.9257896.

[10] Krestinskaya, Olga, et al. "Neural architecture search for in-memory computing-based deep learning accelerators." *Nature Reviews Electrical Engineering* 1.6 (2024): 374-390.

[11] Bae, Jihong, Jongbum Won, and Wooyoung Shim. "The rise of memtransistors for neuromorphic hardware and In-memory computing." *Nano Energy* (2024): 109646.

[12] Gökgöz, Baki, Fatih Gül, and Tolga Aydın. "An overview memristor based hardware accelerators for deep neural network." *Concurrency and Computation: Practice and Experience* 36.9 (2024): e7997.

[13] Hassan, Ali H., and Hassan Mostafa. "Nondestructive reading and refreshment circuit for memristor-based neuromorphic synapse." *2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS)*. IEEE, 2019.

[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] Dhull, Seema, et al. "Quantized magnetic domain wall synapse for efficient deep neural networks." *IEEE Transactions on Neural Networks and Learning Systems* (2024).

[16] Haragopal, Vemuru, et al. "Field-Induced Multistate Magnetization Switching in Ferromagnetic Nanowire with Parallel Anti-dots for Memristor Applications." *Journal of Superconductivity and Novel Magnetism* 37.11 (2024): 1793-1800.

[17] Cruz-Camacho, Elkin, et al. "Performance evaluation of spintronic-based spiking neural networks using parallel discrete-event simulation." *ACM Transactions on Modeling and Computer Simulation* 35.1 (2024): 1-30.

[18] Fu, Siqing, et al. "Spin-NeuroMem: A Low-Power Neuromorphic Associative Memory Design Based on Spintronic Devices." *arXiv preprint arXiv:2404.02463* (2024).

[19] Cui, Can. *Emulating advanced neural functions with spin textures for neuromorphic computing*. Diss. 2024.

[20] 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.

[21] Yang, Le, et al. "Memristive crossbar-based circuit design of back-propagation neural network with synchronous memristance adjustment." *Complex & Intelligent Systems* 10.4 (2024): 4883-4896.

[22] 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.

[23] Lee, Byung Moo. "Massive MIMO for Massive Industrial Internet of Things Networks: Operation, Performance, and Challenges." *IEEE Transactions on Cognitive Communications and Networking* (2024).

[24] Zhang, Fan, et al. "On-device continual learning with STT-assisted-SOT MRAM based in-memory computing." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* (2024).

[25] Ito, Megumi, et al. "Lightweight refresh method for PCM-based neuromorphic circuits." *2018 IEEE 18th International Conference on Nanotechnology (IEEE-NANO)*. IEEE, 2018.

[26] Nomura, Akiyo, et al. "Analysis of effect of weight variation on SNN chip with PCM-refresh method." *Neural Processing Letters* 53 (2021): 1741-1751.

[27] "Design and Implementation of 45nm Operational Amplifier", International Journal of Science & Engineering Development Research (www.ijrti.org), ISSN:2455-2631, Vol.7, Issue 9, page no.452 - 458, September-2022, Available :http://www.ijrti.org/papers/IJRTI2209058.pdf.

[28] Chen, Jing, et al. "Spintronic memristor synapse and its RWC learning algorithm." *IET Circuits, Devices & Systems* 12.5 (2018): 579-588.

[29] Kaushik, Roohie, and Jasdeep Kaur. "Design of folded cascode op amp and its application–bandgap reference circuit." *Circuit World* 50.1 (2024): 9-17.

[30] Dutta, Sounak. "Design of a Strong-Arm Dynamic-Latch based comparator with high speed, low power and low offset for SAR-ADC." *arXiv preprint arXiv:2209.07259* (2022).

[31] Razavi, Behzad. "The StrongARM latch [a circuit for all seasons]." *IEEE Solid-State Circuits Magazine* 7.2 (2015): 12-17.

[32] S. F. Nafea, A. A. S. Dessouki, S. El-Rabaie, B. E. Elnaghi, Y. Ismail and H. Mostafa, Area-e±cient read/write circuit for spintronic memristor based memories, IEEE 60th Int. Midwest Symp. Circuits and Systems (MWSCAS), Boston, MA, USA, August 2017, pp. 1544–1547.