The CSI Journal on
Computer Science and Engineering
Vol. #, No. #, 20##
Pages #-#
Regular Paper

Stochastic Spintronic Neuron for Hardware Implementation of Neural Networks

Abdolah Amirany[1], Mohammad Hossein Moaiyeri[1], Kian Jafari[1], and Masoud Meghdadi[1].

[1]Faculty of Electrical Engineering, Shahid Beheshti University, Tehran, Iran.

Abstract

The hardware implementation of neural network has always been of interest to the researchers as it can significantly increase the efficiency and application of neural networks due to the distributed nature of Artificial Neural Networks (ANNs) in both memory and computation. Direct implementation of ANNs also offer large gains when scaling the network sizes. Stochastic neurons are among the most significant aspects of machine learning algorithms and are very important in different neural networks. In this paper, a hardware model for the stochastic neuron based on the two-in-one magnetic tunnel junction (TiO-MTJ) in subcritical current switching regime is proposed. The use of TiO-MTJ has reduced the area of the proposed neuron and eliminated the risk of MTJ read disturbance. Functional evaluation of the proposed model demonstrates that the behavior of the proposed model is comparable to the mathematical description of the stochastic neuron, and it has a negligible error in comparison with the theoretical model. The simulation results of image binarization over 10,000 images indicate that the proposed hardware model has only 0.25% pack signal to noise ratio (PSNR) and 0.02% structural similarity (SSIM) variation compared to its software-based counterpart. The results of corners simulations also show the proper performance of the proposed neuron even in the presence of inevitable major process variations.

Keywords: Stochastic Neuron, Spintronic, Magnetic Tunnel Junction (MTJ), Neural Networks, Image Binarization.

1. Introduction

Conventional computers with Von-Neumann architecture, despite being able to execute billions of commands per second, have major challenges [1-6]. These systems generally perform operations serially and as a result, with the increasing complexity and number of computational operations, the efficiency of these systems is significantly reduced and in some cases they are not even able to perform the desired operations [1, 7].

Recently, the design of computers inspired from natural neural networks with features similar to the human brain, including the learning capability, has received considerable attention from researchers [1, 7-13]. In addition to the design, modeling and hardware implementation of these systems are also of great importance. Hardware implementation greatly increases system performance, as well as significantly reduces the power consumed to perform each operation [1, 6, 7, 12, 14].

One of the most important elements in the machine learning algorithms and hardware implementations of neural networks is stochastic neuron [15-17]. Stochastic neuron and in general stochastic neural network are useful tools for optimization problems since the random fluctuations help it escape from local minima [17, 18]. Application of the stochastic neural networks includes risk management, oncology, bioinformatics, and other similar fields [17].

Spintronic devices such as magnetic tunnel junction (MTJ) due to their attractive features such as compatibility and the ability to build alongside CMOS transistors, nonvolatility, the ability to integrate and stochastic behavior in subcritical current is one of the options considered by researchers for hardware implementation of the neural networks [8, 19-22].

In this paper, a hardware model for stochastic neuron is designed and evaluated using stochastic behavior of two-in-one MTJ (TiO-MTJ) in subcritical current regime. The proposed model has a similar behavior to the mathematical model of stochastic neuron and can be implemented in large numbers for hardware implementation of stochastic neural networks. It is noteworthy that, the use of TiO-MTJ has reduced the area of the proposed neuron and also eliminated the risk of MTJ read disturbance [23].

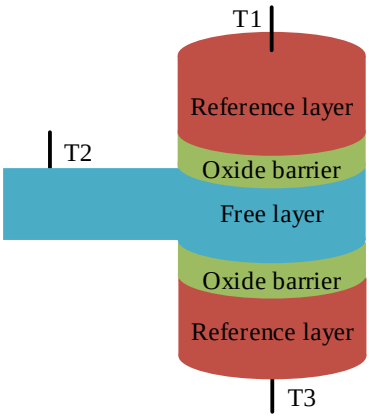


Figure 1. TiO-MTJ structure

The rest of this paper organized as follows: The backgrounds of the research including mathematical model of stochastic neuron, TiO-MTJ structure, and stochastic behavior of the TiO-MTJ in the subcritical current regime are presented in Section II. The proposed stochastic spintronic neuron is introduced in section III. In section IV simulation results are presented and analyzed to verify the functionality of the proposed design. Also in section IV, the results of corner simulations are presented to evaluate the performance of the proposed model in the presence major process variations. Finally based on the simulation results section V concludes the paper.

2. Backgrounds

2.1. Stochastic neuron

In a Threshold Logic Unit (TLU) neuron, if all the factors affecting the output are visible and included in the input, when the sum of the inputs is greater than the threshold value, the output will be '1' and if it is smaller, the output will be '0' [17, 24]. But in a stochastic neuron, not all the factors affecting the output are taken into account or it is not possible to consider all these factors. As a result, comparison with the threshold value is not the only determinant of output and only determines the probability of occurrence of each of the output states ('1' or '0') [17]. This probability function is displayed by $\phi(v)$ and is usually as expressed by (1).

$$\phi(v) = \frac{1}{1 + exp(\frac{-v+v_0}{T})} \quad (1)$$

In (1), v represents the effect of observable inputs (known factors) and T represents the effect of unobservable inputs (unknown factors). v_0 is also used to move the axis. According to (1), the v/T ratio determines the behavior of the neuron. If this ratio is too large or too small, the behavior of the neuron is completely deterministic and neuron behaves stochastically in the middle values of the ratio.

2.2. Magnetic tunnel junction

The TiO-MTJ consists of three ferromagnetic layers separated by two thin insulator layers (shown in Fig. 1). It is observed that if the thickness of the insulator layers is thin enough, the electrons in one ferromagnetic layer can tunnel through the insulator layer and go to the other ferromagnetic layer and create an electric current [20, 25, 26].

Due to the relative orientation of the magnetization vectors of the each of the two ferromagnetic layers two different working modes are possible for a MTJ:

- Parallel mode: In this mode, the magnetization vectors of the ferromagnetic layers are in the same direction.
- Antiparallel mode: In this mode, the magnetization vectors of the ferromagnetic layers are in opposite directions.

It has been observed that the electrons of the ferromagnetic layers in a MTJ in the parallel mode are more likely to tunnel through insulator layer than a MTJ in the antiparallel mode. Due to this phenomenon, a MTJ is like a switch that has two modes, parallel mode with low electrical resistance and antiparallel mode with high electrical resistance [20, 25, 26].

Based on the resistance of the MTJ in the parallel and the antiparallel mode the TMR ratio is defined as (2) [20, 25, 26]. The higher TMR ratio provided higher readability and read stability [21, 22, 27].

$$TMR = \frac{R_{AP} - R_P}{R_P} \times 100 \quad (2)$$

2.3. Stochastic behavior of the MTJ in the subcritical current

To change the state of the MTJ a bi-directional current must pass through the MTJ. According to the geometric structure of the MTJ, a parameter known as critical current (3) is defined for the MTJ [8, 28].

$$I_{Cri} = \alpha \frac{\gamma e}{\mu_b g}(\mu_0 M_s) H_K V_{sl} \quad (3)$$

In (3), α is the magnetic damping constant, γ is the gyromagnetic ratio, e is the electron charge, μ_0 is the vacuum magnetic permeability, M_s is the magnetic saturation field, H_K is the effective magnetic anisotropy field, V_{sl} is the free layer volume, μ_b is the bohr magnetic constant, and g is the spin polarization factor [28].

Behavior of the MTJ is different in the currents higher and lower than the critical current.

- At the currents higher than the critical current, MTJ switching is deterministic [28].
- At the currents lower than the critical current, MTJ switching is probabilistic and follows the probability distribution expressed by (4) [28].

$$P(I,t) = 1 - exp\left\{-\frac{t}{T_0} exp\left[-\Delta\left(1 - \frac{I}{I_{cri}}\right)^2\right]\right\} \quad (4)$$

In this distribution, t is the duration of the passing current, T0 is the attempt time, and Δ is the thermal stability factor.

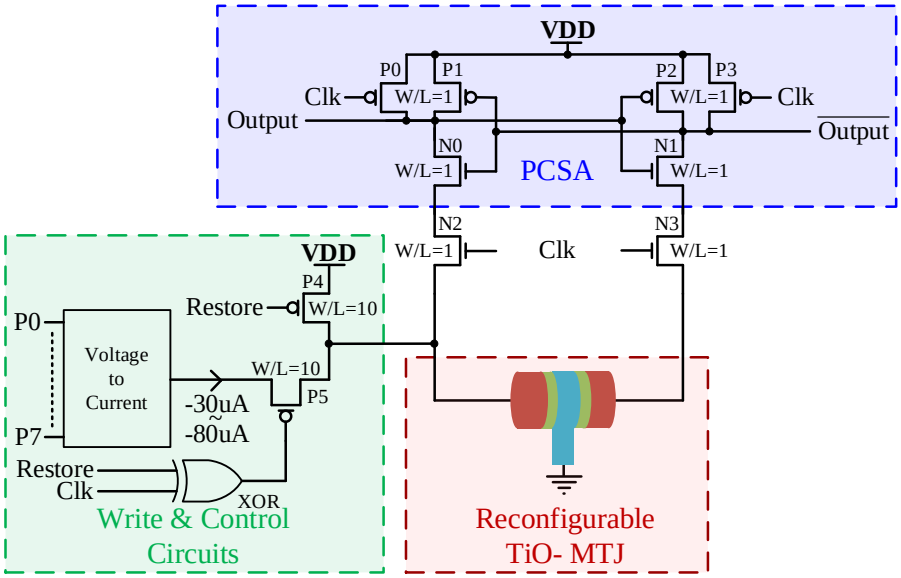


Figure 2. Proposed hardware model for the stochastic neuron

3. Proposed Hardware Model for Stochastic Neuron

Figure 2 shows the proposed hardware model for stochastic neuron. As this Figure shows the proposed model comprised of three main part.

3.1. Pre-charge sense amplifier

Sense amplifier (SA) is responsible for sensing the resistance difference between two terminals of the TiO-MTJ that connected to its tails. There are different types of SA, among them, pre-charge SA (PCSA) is the most used type thanks to its high sensing reliability, performance, and power efficiency than the other SA circuits [29].

As Fig. 2 shows PCSA is composed of six transistors, two PMOS transistors (P0 and P3) which act as the precharge paths, and four NMOS/PMOS transistors (N0, N1, P1, and P1) which act as two cross coupled inverters to construct a keeper. The sensing operation of PCSA includes two modes.

- Pre-charge phase (Clk='0'): During this phase, output nodes are both pre-charged to V_{DD} through P0 and P3. During this phase write and control circuits perform MTJ reconfiguration.
- Evaluation phase (Clk='1'): During this phase, N2, and N3 turn on to discharge one of the outputs to the ground. The MTJ with lower resistance will discharge its corresponding output to the ground faster and hence, its complementary output stays high.

3.2. Reconfigurable TiO-MTJ

This part of the circuit consists of a reconfigurable TiO-MTJ and is the part that creates the stochastic nature of the proposed model.

3.3. Write and control circuits

Write and control circuits is responsible for reconfiguration of the reconfigurable MTJ. The operation of this part of the proposed model can be divided into two phases:

- Deterministic restoring (Clk='0', and Restore='0'): At this time P4 turn on and pass 200µA current through reconfigurable MTJ for 10ns. This current, which is larger than the MTJ critical current (100µA), definitely changes the MTJ mode to antiparallel.
- Stochastic switching (Clk='0', and Restore='1'): During this phase a current between -80µA to -30µA pass through reconfigurable MTJ for 10ns. At this time, the configuration of the reconfigurable MTJ can change according to the probability distribution expressed by (4).

4. Simulation Results

4.1. Functional evaluations

Using 45nm CMOS technology model [30] and TiO MTJ model proposed in [23], HSPICE simulation carried out to validate the functionality of the proposed model. Figure 3 shown the output of the proposed model (Output in Fig. 2) for 1us. This Fig shows output of the proposed model for three different current and hence switching probability (25%, 50% and 75%). As Fig. 3 shown as the MTJ switching current increases, the MTJ switching probability and as well as the probability of the output switching also increases.

Figure 4 shows the performance of the proposed model compared to the mathematical model. The performance of the model can be divided into three regions as logical one saturation (synapse current between 70µA to 80µA), logical zero saturation (synapse current between 30µA to 40µA) and the probabilistic switching region (synapse current between 40µA to 70µA).

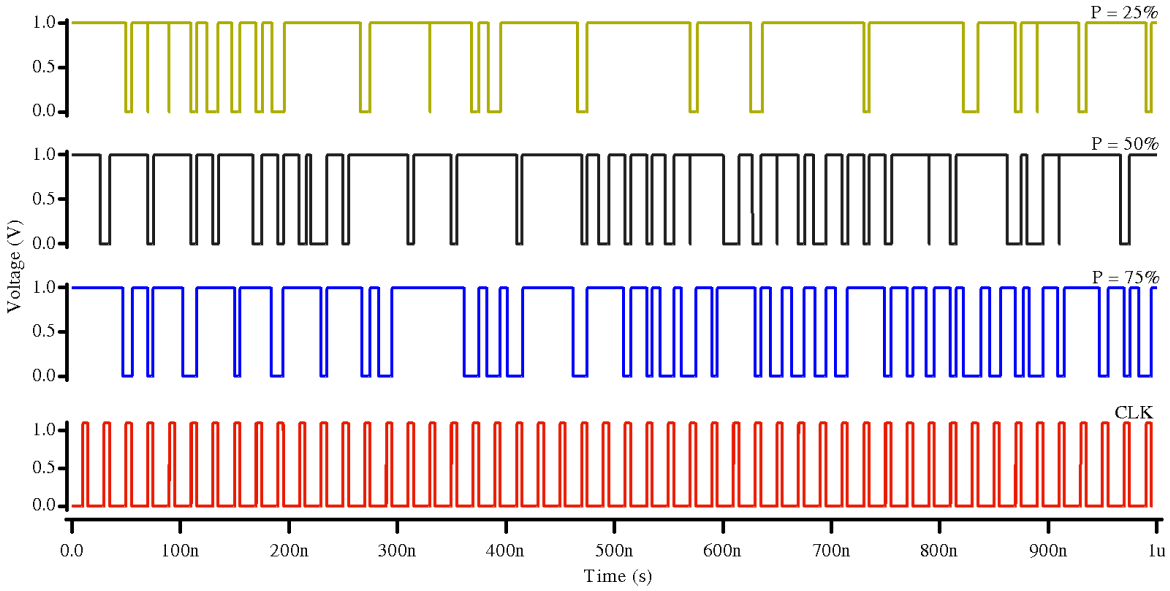


Figure 3. Transient output of the proposed model for three different current and hence switching probability (25%, 50% and 75%)

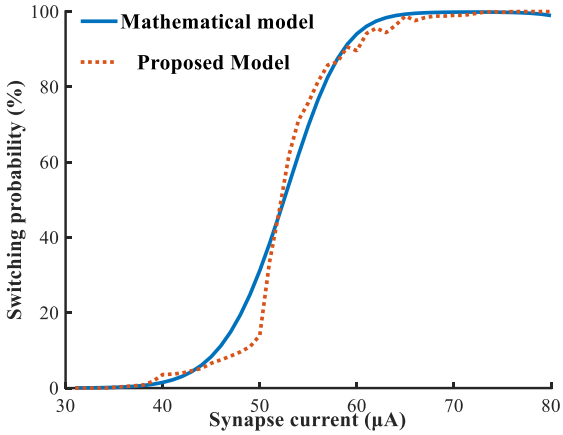


Figure 4. The performance of the proposed model compared to the mathematical model

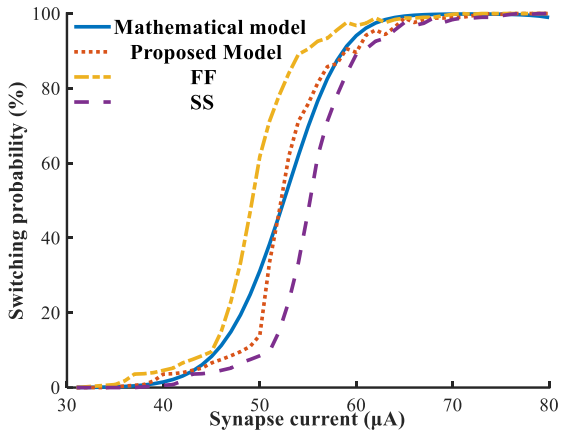


Figure 5. The performance of the proposed model in the fabrication process corners (SS and FF).

As Fig. 4 shows, in the logical one and zero saturation regions, the performance of the proposed model is quite similar to that of the mathematical model, and in the probabilistic switching region, it performs very close to the mathematical model.

4.2. Corners simulations

To investigate the effect of process variations on the performance of the proposed model, simulations were performed at two corners of the fabrication process: SS (lowest current) and FF (maximum current).

Figure 5 shows the probability of TiO-MTJ switching at the SS and FF corners. As this figure shows, the proposed model has an acceptable performance in the corners of the fabrication process.

4.3. Image binarization evaluations

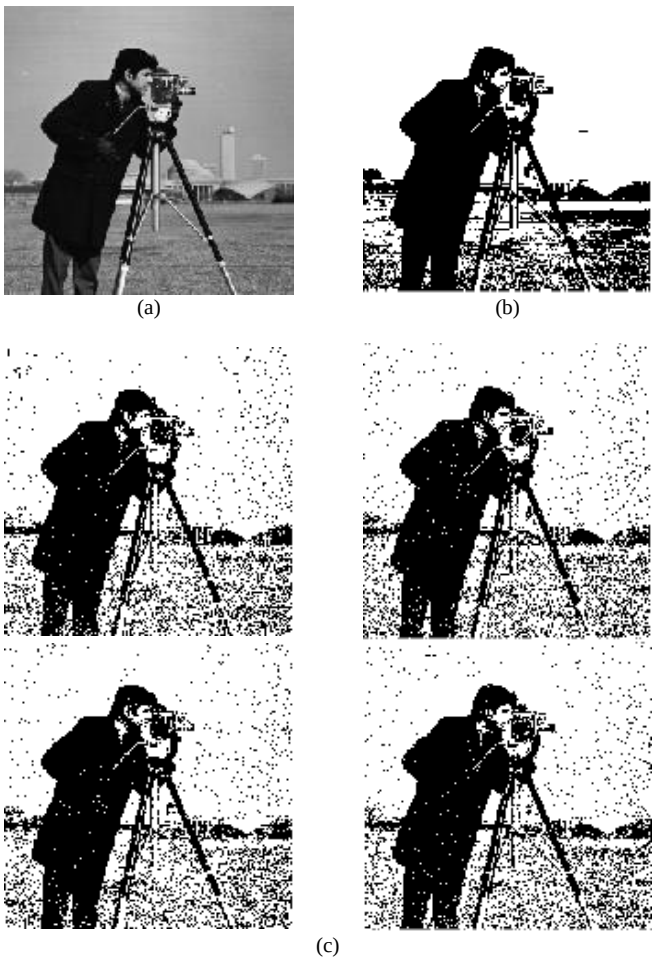
MATLAB simulation performed to validate the functionality of the proposed hardware model in image binarization. For this purpose, the following steps have been done for 10,000 times:

- The pixel values of the image were converted to a current between 30μA and 50μA.
- The corresponding current of each pixel was applied to the reconfigurable MTJ.
- The output of proposed model was considered as the binary equivalent of the image.

For each of the 10,000 performed simulations, peak signal to noise ratio (PSNR) [31] and structural similarity (SSIM) [32] was obtained relative to the ideal binary image.

Figure 6 shows the original grayscale image, the ideal binarized image generated using MATLAB ‘imbinarize’ function with a threshold value of 0.5 and four outputs of the proposed model. As this figure shows the output of the proposed hardware model for stochastic neuron in image binarization is very similar to that of the MATLAB ‘imbinarize’ function.

Figures 7 shows the boxcharts of the PSNR and SSIM of the binarized image using the proposed model. As these figures, show, the proposed model has 0.25% PSNR and 0.02% SSIM variation in comparison withs the software counterpart.



(a) (b) (c)

Figure 6. Simulation results of the proposed model a) Original grayscale image b) The ideal binarized image generated using MATLAB 'imbinarize' function with a threshold value of 0.5 c) four outputs of the proposed model

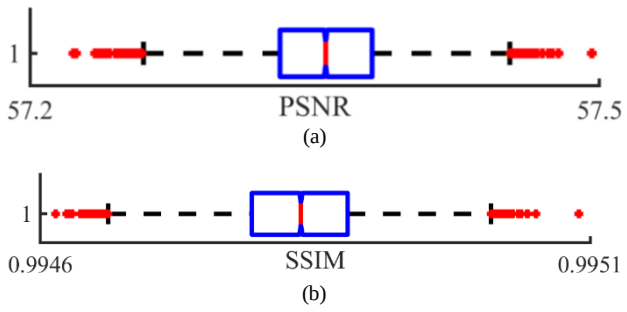


Figure 7. Boxcharts of the PSNR and SSIM of the binarized image using proposed model

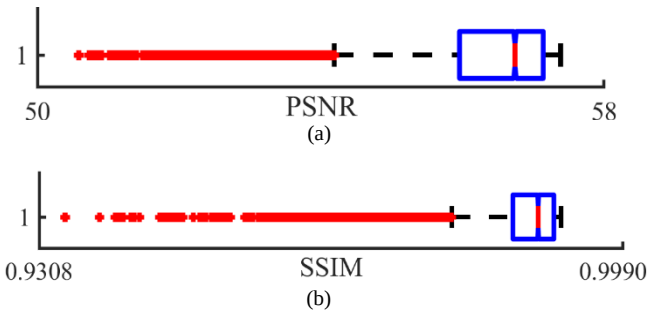


Figure 8. Boxcharts of the PSNR and SSIM of the binarized image using proposed model in the presence of the process variations

Figures 8 also shows the boxcharts of the PSNR and SSIM of the binarized image using the proposed model in the presence of the process variations. As these figures show, the proposed model has 10.24% PSNR and 5.28% SSIM variation as compared to the software counterpart in the presence of the process variations.

5. Conclusion

In this paper, a hardware model for the stochastic neuron based on the stochastic behavior of MTJ with the currents lower than the critical current is proposed. The stochastic neurons are one of the most important elements in the implementation of hardware neural networks, machine-learning algorithms, and in general, artificial intelligence. Due to the massive parallel nature of artificial neural network hardware the proposed hardware model can significantly increase the performance of the artificial neural network by taking the advantages of the parallelism of the neural networks and the possibility of implementing the model in high numbers and densities. Image binarization simulation results over 10,000 images indicate that the proposed hardware model has only 0.25% PSNR and 0.02% SSIM variation in comparison to the software counterpart and also 10.24% PSNR and 5.28% SSIM variation than the software counterpart in the presence of the process variations, which are completely acceptable.

References

[1] A. Amirany, M. H. Moaiyeri, and K. Jafari, "Nonvolatile Associative Memory Design Based on Spintronic Synapses and CNTFET Neurons," *IEEE Transactions on Emerging Topics in Computing*, pp. 1-1, 2020.

[2] D. Shin and H.-J. Yoo, "The Heterogeneous Deep Neural Network Processor With a Non-von Neumann Architecture," *Proceedings of the IEEE*, pp. 1-16, 2019.

[3] S. Basu, R. E. Bryant, G. De Micheli, T. Theis, and L. Whitman, "Nonsilicon, Non-von Neumann Computing—Part I [Scanning the Issue]," *Proceedings of the IEEE*, vol. 107, no. 1, pp. 11-18, 2019.

[4] A. Amirany and R. Rajaei, "Low Power, and Highly Reliable Single Event Upset Immune Latch for Nanoscale CMOS Technologies," presented at the Electrical Engineering (ICEE), Iranian Conference on, 2018.

[5] M. Krishna Gopi Krishna, A. Roohi, R. Zand, and R. F. DeMara, "Heterogeneous energy-sparing reconfigurable logic: spin-based storage and CNFET-based multiplexing," *IET Circuits, Devices & Systems*, vol. 11, no. 3, pp. 274-279, 2017.

[6] C. Pan and A. Naeemi, "Non-Boolean Computing Benchmarking for Beyond-CMOS Devices Based on Cellular Neural Network," *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits*, vol. 2, pp. 36-43, 2016.

[7] A. Amirany, M. H. Moaiyeri, and K. Jafari, "Process-in-Memory Using a Magnetic-Tunnel-Junction Synapse and a Neuron Based on a Carbon Nanotube Field-Effect Transistor," *IEEE Magnetics Letters*, vol. 10, pp. 1-5, 2019.

[8] A. Amirany, K. Jafari, and M. H. Moaiyeri, "True Random Number Generator for Reliable Hardware Security Modules Based on a Neuromorphic Variation-Tolerant Spintronic

Structure," *IEEE Transactions on Nanotechnology*, pp. 1-1, 2020.

[9] A. Sengupta, Y. Shim, and K. Roy, "Proposal for an All-Spin Artificial Neural Network: Emulating Neural and Synaptic Functionalities Through Domain Wall Motion in Ferromagnets," *IEEE Trans Biomed Circuits Syst*, vol. 10, no. 6, pp. 1152-1160, Dec 2016.

[10] Z. I. Mannan, S. P. Adhikari, C. Yang, R. K. Budhathoki, H. Kim, and L. Chua, "Memristive Imitation of Synaptic Transmission and Plasticity," *IEEE Trans Neural Netw Learn Syst*, Feb 11 2019.

[11] Y. Yan *et al.*, "Efficient Reward-Based Structural Plasticity on a SpiNNaker 2 Prototype," *IEEE Trans Biomed Circuits Syst*, vol. 13, no. 3, pp. 579-591, Jun 2019.

[12] A. Amirany, M. H. Moaiyeri, and K. Jafari, "Bio-Inspired Nonvolatile and Low-Cost Spin-Based 2-Bit per Cell Memory," presented at the 25th International Computer Conference, Computer Society Of Iran, Tehran, Iran, 2020.

[13] A. Amirany, K. Jafari, and M. H. Moaiyeri, "BVA-NQSL: A Bio-inspired Variation Aware Nonvolatile Quaternary Spintronic Latch," *IEEE Magnetics Letters*, vol. 11, pp. 1-5, 2020.

[14] C. Pan and A. Naeemi, "A Proposal for Energy-Efficient Cellular Neural Network Based on Spintronic Devices," *IEEE Transactions on Nanotechnology*, vol. 15, no. 5, pp. 820-827, 2016.

[15] K. Gurney, *An Introduction to Neural Networks*. Taylor \\& Francis, Inc., 1997, p. 288.

[16] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall PTR, 1994, p. 768.

[17] C. Turchetti, *Stochastic Models of Neural Networks*. IOS Press, 2004.

[18] A. Mondal and A. Srivastava, "Energy-efficient Design of MTJ-based Neural Networks with Stochastic Computing," *ACM Journal on Emerging Technologies in Computing Systems*, vol. 16, no. 1, pp. 1-27, 2020.

[19] A. Amirany, K. Jafari, and M. H. Moaiyeri, "High-Performance and Soft Error Immune Spintronic Retention Latch for Highly Reliable Processors," presented at the Electrical Engineering (ICEE), Iranian Conference on, 2020.

[20] Y. Wang, Y. Zhang, E. Y. Deng, J. O. Klein, L. A. B. Naviner, and W. S. Zhao, "Compact model of magnetic tunnel junction with stochastic spin transfer torque switching for reliability analyses," *Microelectronics Reliability*, vol. 54, no. 9-10, pp. 1774-1778, 2014.

[21] R. Rajaei and A. Amirany, "Nonvolatile Low-Cost Approximate Spintronic Full Adders for Computing in Memory Architectures," *IEEE Transactions on Magnetics*, vol. 56, no. 4, pp. 1-8, 2020.

[22] R. Rajaei and A. Amirany, "Reliable, High-Performance, and Nonvolatile Hybrid SRAM/MRAM-Based Structures for Reconfigurable Nanoscale Logic Devices," *Journal of Nanoelectronics and Optoelectronics*, vol. 13, no. 9, pp. 1271-1283, 2018.

[23] S. Huda and A. Sheikholeslami, "A Novel STT-MRAM Cell With Disturbance-Free Read Operation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 6, pp. 1534-1547, 2013.

[24] J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, "Neuromorphic spintronics," *Nature Electronics*, 2020.

[25] J. C. Slonczewski, "Conductance and exchange coupling of two ferromagnets separated by a tunneling barrier," *Phys Rev B Condens Matter*, vol. 39, no. 10, pp. 6995-7002, Apr 1 1989.

[26] Y. Zhang *et al.*, "Compact Modeling of Perpendicular-Anisotropy CoFeB/MgO Magnetic Tunnel Junctions," *IEEE Transactions on Electron Devices*, vol. 59, no. 3, pp. 819-826, 2012.

[27] A. Amirany, F. Marvi, K. Jafari, and R. Rajaei, "Nonvolatile Spin-Based Radiation Hardened Retention Latch and Flip-Flop," *IEEE Transactions on Nanotechnology*, vol. 18, pp. 1089-1096, 2019.

[28] Y. Qu, J. Han, B. F. Cockburn, W. Pedrycz, Y. Zhang, and W. Zhao, "A true random number generator based on parallel STT-MTJs," presented at the Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017, 2017.

[29] A. Amirany and R. Rajaei, "Fully Nonvolatile and Low Power Full Adder Based on Spin Transfer Torque Magnetic Tunnel Junction With Spin-Hall Effect Assistance," *IEEE Transactions on Magnetics*, vol. 54, no. 12, pp. 1-7, 2018.

[30] "Predictive Technology Model (PTM) Low Power 45nm Metal Gate / High-K / Strained-Si." http://ptm.asu.edu/modelcard/LP/45nm_LP.pm (accessed.

[31] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital image processing using MATLAB*. Pearson Education India, 2004.

[32] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans Image Process*, vol. 13, no. 4, pp. 600-12, Apr 2004.

Abdolah Amirany received the B.Sc. degree in electrical engineering from the University of Kashan, Kashan, Iran, in 2016, and the M.Sc. degree from Shahid Beheshti University, Tehran, Iran, in 2018, where he is currently pursuing the Ph.D. degree with the School of Electrical Engineering .

His research interests include VLSI design, emerging technologies, hard-ware security, neuromorphic computing, and approximate computing.

Mr. Amirany was awarded as the highest ranking student and top student researcher in 2018 and received the National Elite Foundation of Iran Scholarship Award in 2020.

Email: a_amirany@sbu.ac.ir

Mohammad Hossein Moaiyeri received the Ph.D. degree in computer architecture from Shahid Beheshti University, Tehran, Iran in 2012. He is currently an Assistant Professor in the Faculty of Electrical Engineering, Shahid Beheshti University. His research interests include VLSI design for beyond-CMOS emerging nanodevices, low-power VLSI design, and mixed-signal integrated circuits design.

Email:
h_moaiyeri@sbu.ac.ir

Kian Jafari received his Ph.D. in Electrical Engineering from SUPELEC (University Paris-Saclay) with a research scholarship of the French Ministry for Higher Education and Research. After successful completion of his Ph.D., he has served as the Scientific Project Manager in the French National Research Agency (ANR) in Paris since 2011 to 2014. He joined then the Electrical Engineering Department of Shahid Beheshti University (SBU), where he is currently an Assistant Professor and Head of Electronics Group. He has also served as vice-director of Technology Incubator Center of the University for 3 years up to 2019. During his professional activities at SBU, he carried out various research and industrial projects and published more than 50 research works in international journals and symposiums. His research interests include design, modeling and fabrication of M/NEMS devices for various applications (particularly biomedical applications) as well as neuromorphic computing and hardware implementation of neural networks.

Email: k_jafari@sbu.ac.ir

Masoud Meghdadi received the B.Sc. degree in electrical engineering from Amirkabir University of Technology, Tehran, Iran, in 2005, and the M.Sc. and the Ph.D. degrees in electrical engineering from Sharif University of Technology, Tehran, in 2007 and 2013, respectively. He is currently an Assistant Professor with the faculty of Electrical Engineering, Shahid Beheshti University, Tehran, Iran. His current research interests include high-frequency, analog, and mixed-mode circuits and systems.

Email: m_meghdadi@sbu.ac.ir

Paper Handling Data:

Submitted: -
Received in revised form: -
Accepted: -
Corresponding author: Kian Jafari, Faculty of Electrical Engineering, Shahid Beheshti University, Tehran, Iran.