



HAL
open science

Towards reliability of SNNs: defect implementation on STT-MRAM cell

Salah Daddinounou, Elena Ioana Vatajelu

► **To cite this version:**

Salah Daddinounou, Elena Ioana Vatajelu. Towards reliability of SNNs: defect implementation on STT-MRAM cell. ETS 2021, May 2021, Online (due to COVID, initially planned in Leuven), Belgium. hal-04358340

HAL Id: hal-04358340

<https://hal.science/hal-04358340>

Submitted on 28 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Towards reliability of SNNs: defect implementation on STT-MRAM cell

Salah DADDINOUNOU
TIMA laboratory
Université Grenoble Alpes
Grenoble, France

salah.daddinounou@univ-grenoble-alpes.fr

Elena Ioana VATAJELU
CNRS, TIMA laboratory
Université Grenoble Alpes
Grenoble, France

ioana.vatajelu@univ-grenoble-alpes.fr

Abstract—Spiking Neural Networks (SNNs) are being widely studied to mimic the intelligence of biological brains. In hardware level, multiple emerging non-volatile memories (NVM) are proposed to model the neural synapses, one of them being Spin-transfer Torque Magnetic Random-Access Memory (STT-MRAM). In this paper we describe a single STT-MRAM with the sense amplifier (SA) and study the effect of implemented defects on the performance of this memory cell.

Keywords—emerging NVM, reliability, STT-MRAM, spiking neural networks.

I. CONTEXT

In the era of big data and internet of things, the search for more energy efficient computing paradigms than classic Von Neuman architecture is in continuous growth. The use of neural networks for image recognition is known for decades now in the field of artificial intelligence (AI). To take the inspiration from nature one step further, spiking neural networks were proposed simultaneously with the rise of new devices, which will ultimately allow the emulation of the most powerful and yet energy efficient computing system: The Human brain.

STT-MRAM uses a magnetic tunnel junction (MTJ) which consists of a stack of magnetic layers separated by a thin insulating layer. This device having two stable resistances is used as a NV memory with high resistance (logic 1) and low resistance (logic 0) [1]. Furthermore, the MTJ device is stochastic, meaning that the state switching is not a deterministic process. Adding to that its accumulative behavior, which can be seen when consecutive pulses may add up causing the MTJ to switch state. The tow characteristics make the MTJ a suitable device to map for integrate-and-fire and probabilistic behaving neurons of the cortex.

II. MOTIVATIONS

Since the building blocks of our ‘synapses’ and ‘neurons’ are the MTJ devices, it is important to understand what may affect their performance. First by looking to the effect of intrinsic stochasticity due to thermal fluctuations, and then the variability of MTJ resistance resulting from the fabrication process. Third we intentionally include defects all around the MTJ device in order to simulate the resistive defects that can occur in the interconnections during fabrication. In our study, the MTJ is used as a memory cell with the classic read and write operations, with all the periphery that goes with. The results will benefit both uses of the MTJ: as a NV memory and in neuromorphic computing.

III. STATE OF THE ART

In this section, we will briefly report the main faults that can harm the MTJ performance -regardless of its use as an MRAM or as a synapse- as found in recent literature [2]. Faults which are caused during manufacturing process are referred to as defectiveness issues, whereas those which occur during the lifetime of the device are known as reliability issues.

A. defectiveness issues

Examples of these defects can be open, bridge resistances or parasitic coupling capacitance [3]. Inside the MTJ, the reasons that lead to these defects are multiple, such as dust particles deposition, void created while polishing the electrodes, breakdown of the thin oxide layer, or the metallic sideways deposition after an improper etching. Some defects can also damage the magnetic free layer of the MTJ making its magnetization blocked in a certain orientation. Depending on the size of the defect and its extent, we can classify them as strong ones, where the device is permanently damaged, or weak defects where the cell still works but only within a given range of defect size.

Parametric failures are another type of defects exclusive to MRAMs. The fabrication process includes two separated steps[4], namely the magnetic part where the MTJ layers are deposited, then comes the integration with CMOS to connect the MTJ to the access transistor. After integration, defects may occur due to three parameters: electrical (transistor V_{th}), material related (TMR, magnetic anisotropy), and lithographic (dimensions of the MTJ and transistor).

B. Reliability issues

The following reliability issues should be considered when scaling a product for a commercial success: *a) Process variability* can affect the MTJ conductance and switching threshold current and lead to fluctuations in MTJ resistance up to 4% from its nominal value [2] due to geometrical effects discussed earlier. *b) Stochastic switching* is an intrinsic parameter of the MTJ, although random switching is considered a fault in a synchronous system like MRAMs, it is a feature for neural networks. *c) Retention failure* happens when the MTJ state switches without accessing it, this is mainly due to thermal instability because retention time is exponentially proportional to thermal stability. *d) Read disturbance* is the result of accidentally switching the state while reading. A high reading current can be at the origin of this fault, but this can be mitigated using a longer pulse duration instead [2].

IV. PROPOSED METHOD

Our circuit design is shown in Fig.1. The cell consists of a single MTJ with the access transistor (the pinned magnetic layer towards the top). Switching between the two states of the MTJ requires a bidirectional current, the four transistors around the cell ensure the writing logic, they are controlled by the voltages program high (PH) and program-low (PL). Enabling PL allows the current to flow from the source line (SL) to the bit line (BL), letting it write a logic 0 in the cell ($R = 6.2k\Omega$). On the other hand, writing a logic 1 ($R = 18.5k\Omega$) is performed by enabling PH voltage which triggers the opposite process.

Reading operation uses a low unidirectional current (from BL to SL). The equivalent resistance of the four pre-written MTJs in the right-hand side stands as a reference cell whose resistance is the average between high and low resistances of

an MTJ ($R = 12.35k\Omega$). During reading operation, the SA adapted from [5] starts by equalizing the voltages, once the equalization is off, it detects the part having higher resistance according to the sensed voltages. For the sense amplifier transistors, we used the same dimensions reported at [5]. On the other hand, transistors that control read/write operations have dimensions $W/L = 7\mu\text{m}/30\text{nm}$. The oscilloscope at the bottom of the circuit stores the outputs of the SA and the actual state of the MTJ for further use. Finally, the resistive defects are implemented as shown in Fig.1.

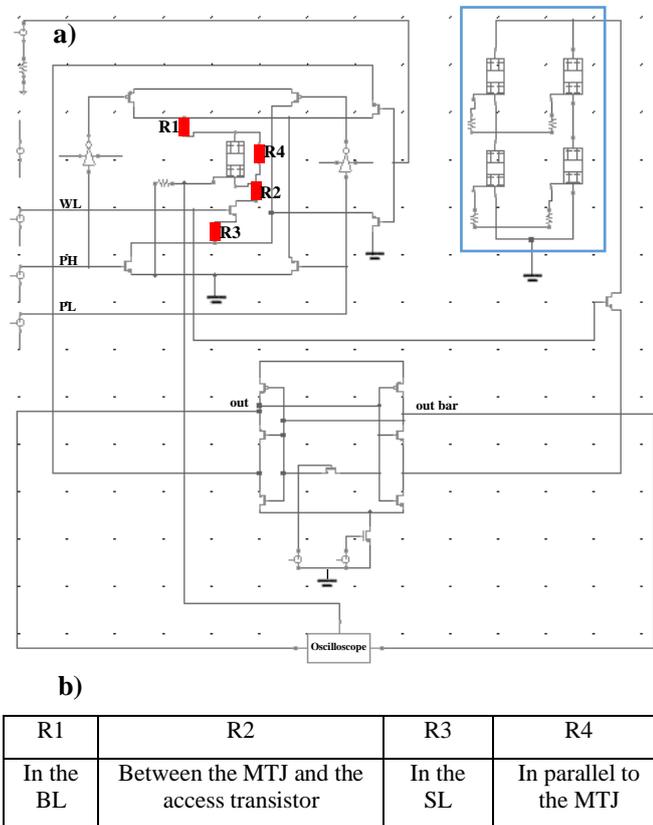


Fig. 1. a) Circuit of the MTJ cell with read and write peripheries, the defects are shown in red, the reference cell is outlined in blue, the voltage sources in the left-hand side enable the different operations. b) Resistive defects implemented around the STT-MRAM cell.

Furthermore, the MTJ model that we used in Cadence® [6] allows the implementation of MTJ-dependent failures. Stochastic switching is set to a gaussian distribution. Fluctuations in TMR ratio, in free layer and in oxide thickness are all combined in one variable called RV following a gaussian distribution too. In the next section we show the effect of each defect, these imperfections have different effects on writing and reading, for both logic values 0 and 1. This study aims to obtain the range of defect values that can be accepted.

V. RESULTS

Using Cadence® Virtuoso® System Design Platform, we run parametric analysis. For each defect parameter value, we run a Monte Carlo simulation with $N = 50$ points to get statistically relevant results. The stochastic behavior of the MTJ and its process variability are both taken into account, both in form of a gaussian distributions with a standard deviation of 3%. Obtained results are shown in the table at Fig.2. All R1, R2 and R3 are counted as the in-series defect. It is worth noting that the in-parallel (the in-series) defects have no effect on reading '0' (reading1) respectively, this is

because the equivalent resistance MTJ & defect is always smaller (greater) than that of the reference cell, respectively.

operation	defect	Totally correct	Partially faulty	Totally wrong
Write 0	In series	[0; 0.3]	[0.3; 2.1]	[2.1; +∞[
	In parallel	[24; +∞[[2.5; 24]	[0; 2.5]
Write 1	In series	[0; 13.8]	[13.8; 16.7]	[16.7; +∞[
	In parallel	[0.28; +∞[[0.18; 0.28]	[0; 0.18]
read 0	In series	[0; 5.6]	[5.6; 8]	[8; +∞[
	In parallel	[0; +∞[X	X
read 1	In series	[0; +∞[X	X
	In parallel	[60; +∞[[31; 60]	[0; 31]

Fig. 2. summary of the effect of defectiveness issues in read and write operations (all values in $k\Omega$). The 'X' refers to cases where the defect is not a problem for the specified parameter.

We can see from the table of Fig.2 that writing '1' process tolerates higher in-series defects, this is because less current is needed to write '1' (high resistance, ie: switch MTJ into anti-parallel state) as reported in [5].

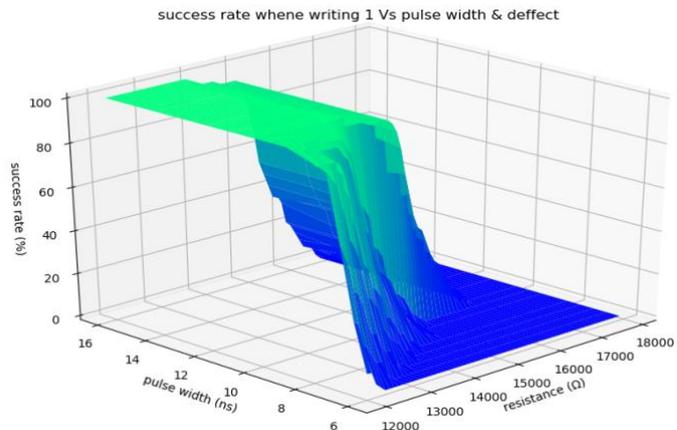


Fig. 3. success rate is given for the process of writing '1' with respect to the in-series defect values & writing pulse width.

The Fig.3 shows the effect of the writing time on the success rate of writing '1', indeed, we can achieve higher switching probability by increasing the width of the writing pulse. for the previously shown results, a pulse of 10ns width and 1V amplitude is used. since the MTJ has two switching modes: high current with short pulse or large pulse with low current, a compromise should be considered depending on the application, low current is rather preferred for the spiking neural networks since the energy consumption is a big concern for this specific application. Whereas the other mode is a good solution for the use of MTJ as a memory where we care more about the writing accuracy.

VI. CONCLUSION

We presented in this paper the possible effects that can occur in an MRAM cell based on spin-transfer torque MTJ, resistive defects were implemented in series (in SL and BL) and in parallel to the MTJ, the effect and the extent of each was analyzed, the results may be insightful when studying reliability of Spiking Neural Networks.

REFERENCES

- [1] Chun, K. C., Zhao, H., Harms, J. D., Kim, T.-H., Wang, J.-P., & Kim, C. H. (2013). 'A Scaling Roadmap and Performance Evaluation of In-Plane and Perpendicular MTJ Based STT-MRAMs for High-Density Cache Memory'. *IEEE Journal of Solid-State Circuits*, 48(2), 598–610. <https://doi.org/10.1109/jssc.2012.2224256>
- [2] Patrick Girard, Yuanqing Cheng, Arnaud Virazel, Weisheng Zhao, Rajendra Bishnoi, et al. 'A Survey of Test and Reliability Solutions for Magnetic Random Access Memories'. Proceedings of the IEEE, Institute of Electrical and Electronics Engineers, In press, pp.1-21. 10.1109/JPROC.2020.3029600. lirmm-03031646
- [3] Nowak, J. J., Robertazzi, R. P., Sun, J. Z., Hu, G., Abraham, D. W., Trouilloud, P. L., Brown, S., Gaidis, M. C., O'Sullivan, E. J., Gallagher, W. J., & Worledge, D. C. (2011). 'Demonstration of ultralow bit error rates for spin-torque magnetic random-access memory with perpendicular magnetic anisotropy'. *IEEE Magnetics Letters*, 2, 3000204. <https://doi.org/10.1109/lmag.2011.2155625>
- [4] Li, J., Ndai, P., Goel, A., Salahuddin, S., & Roy, K. (2010). Design Paradigm for Robust Spin-Torque Transfer Magnetic RAM (STT MRAM) From Circuit/Architecture Perspective. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 18(12), 1710–1723. <https://doi.org/10.1109/tvlsi.2009.2027907>
- [5] J.W. Kwak, A. Marshall, H. Stiegler, '28nm STT-MRAM Array and Sense Amplifier' 2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST), 978-1-7281-1184-1/19 IEEE2019
- [6] Y. Wang, Y. Zhang, W.S. Zhao, L. A. B. Naviner, H. Cai and J-O. Klein, 'Manual of compact models for Perpendicular Magnetic Anisotropy (PMA) Magnetic Tunnel Junction (MTJ)