Fault Tolerant Non-Volatile Spintronic Flip-Flop

Rajendra Bishnoi, Fabian Oboril and Mehdi B. Tahoori
Chair of Dependable Nano Computing (CDNC), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
Email: {rajendra.bishnoi, fabian.oboril, mehdi.tahoori}@kit.edu

***Abstract*—With technology down scaling, static power has become one of the biggest challenges in a *System-On-Chip*. Normally-off computing using non-volatile sequential elements is a promising solution to address this challenge. Recently, many non-volatile shadow flip-flop architectures have been introduced in which *Magnetic Tunnel Junction* (MTJ) cells are employed as backup storing elements. Due to the emerging fabrication processes of magnetic layers, MTJs are more susceptible to manufacturing defects than their CMOS counterparts. Moreover, unlike memory arrays that can effectively be repaired with well-established memory repair and coding schemes, flip-flops scattered in the layout are more difficult to repair. So, without effective defect and fault tolerance for non-volatile flip-flops, the manufacturing yield will be severely affected. In this paper, we propose a *Fault Tolerant Non-Volatile Latch* (FTNV-L) design, in which several MTJ cells are arranged in such a way that it is resilient to various MTJ faults. Simulation results show that our proposed FTNV-L can effectively tolerate all single MTJ faults with considerably lower overhead than traditional approaches.**

I. Introduction

With the advancements in technology scaling, the excessive leakage power of CMOS devices becomes a major design issue. Therefore, non-volatile magnetic memories using spintronic technologies such as *Spin Transfer Torque (STT)* or *Spin Orbit Torque* (SOT) are gaining popularity. This is due to their various advantageous features such as high endurance, scalability, high density, low access latency, soft-error immunity and CMOS compatibility [1, 2]. These technologies use a *Magnetic Tunnel Junction* (MTJ) cell as a storing device, which stores the logic value as resistance state. These storing devices can also be employed for flip-flops design in a low power *System-On-Chip* (SoC). Therefore, many *non volatile* (NV) shadow flip-flop architectures have recently been introduced which exploit the normally-off and instantly-on attributes of MTJs [3, 4]. However, a single MTJ failure in such designs can lead to a complete breakdown of the normally-off capabilities.

The fabrication of the magnetic layers to implement MTJ cells is more complex than that of conventional CMOS, since it is a new process based on new materials. Therefore, it is expected that magnetic layers are more prune to manufacturing defects than CMOS layers [5]. For instance, there is a possibility of the *barrier short* during the ion beam etching process [6]. As a consequence, the affected MTJs have a very low resistance value [7]. On the contrary, MTJ cells exhibit an extremely high resistance for an *open defect*. In addition, the magnetic orientation of the MTJ cell can be fixed to a specific magnetization configuration, meaning their magnetic orientation cannot be changed [8–10]. This may happen permanently because of manufacturing defects in the magnetic layer or due to loss of margin in the CMOS support circuitry, such as reduced switching current or duration [8]. All these defects can severely hurt the manufacturing yield of these emerging technologies to prevent their widespread adoption.

In order to render manufacturing defects, memories are usually equipped with redundancies and error detection/correction mechanisms [9]. However, these techniques are inapplicable to flip-flop designs, because flip-flops are scattered widely in the SoC layout as individual cells. Nevertheless, in flip-flop designs, these faults can be addressed using traditional *Triple Modular Redundancy* (3MR), in which the shadow latch component is triplicated, and the final output is generated based on a voting mechanism. In fact, it incurs huge area, energy and latency costs. Therefore, it is a decisive need to have a cost effective solution to deal with MTJ faults for overall yield and energy efficiency.

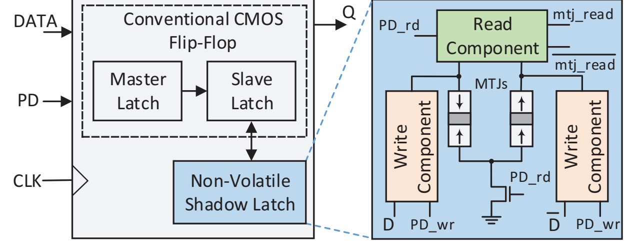

Fig. 1. Overview of shadow non-volatile flip-flop architecture

In this paper, we propose a novel shadow flip-flop architecture, in which we design a generic *Fault Tolerant Non-Volatile Latch* (FTNV-L) applicable to all NV flip-flops, to address the aforementioned faults in MTJ cells. In our proposed FTNV-L design, several MTJ cells are structured in such a way that it can easily tolerate all single MTJ faults within a flip-flop. Simulation results demonstrate that our proposed FTNV-L design delivers required resistance differences in the presence of any single faults to guarantee a fault-free functionality. Moreover, it has almost the same performance and energy for both backup and restore processes as a standard NV flip-flop. In addition, adding MTJ cells has a minimal impact on the overall flip-flop area, as they are fabricated in different layers.

II. Background

A. Spin Transfer Torque

The Magnetic Tunnel Junction (MTJ) cell, which is a storing device, consists of two ferromagnetic layers separated by a thin barrier oxide. One of the ferromagnetic layers, whose magnetic orientation is always fixed, is known as *Referenced Layer* (RL). The magnetic orientation of the other layer can be freely rotated, which is termed as *Free Layer* (FL). An MTJ cell stores data as a resistance state. When the magnetic orientation of the two ferromagnetic layers are parallel to each other ('P' configuration), it exhibits a low resistance value. Otherwise, it has a high resistance value, when the magnetic orientation of those two layers are anti-parallel to each other ('AP' configuration).

B. Overview of Non-Volatile flip-flop

A shadow flip-flop architecture using MTJ-based non-volatile storing devices is very effective for leakage power reduction. This is due to the fact that by adopting these designs, the entire logic core can be power gated, unlike for conventional CMOS-based flip-flops. As illustrated in Fig 1, it consists of three components, namely, *master latch, slave latch,* and *NV shadow latch*. Here, the shadow latch consists of two MTJs as well as read and write components. During power-down, the data is stored in the shadow latch before

978-3-9815370-6-2/DATE16/©2016 EDAA

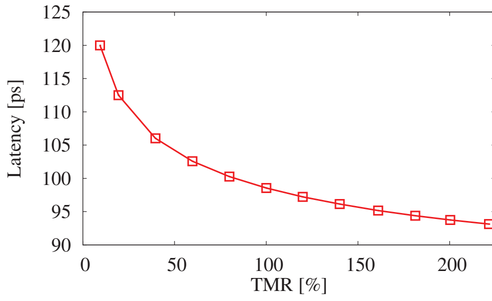

Fig. 2. Read latency for various TMR values (for typical process corner and see Table I for setup information)

going into the sleep-mode, and it is read and restored into the slave latch during wakeup. The *PD_rd* and *PD_wr* signals, which are generated using the PD pin, are employed to activate the read and store operations, respectively. The two MTJs should always store the opposite magnetizations which assists the read process by sensing the resistance differences. If any one of the MTJs is faulty, the entire shadow latch component cannot be used in the given architecture. Hence, our proposed shadow latch component is designed in such a way that it is capable of delivering the correct output in the presence of a single faulty MTJ within each flip-flop.

C. MTJ read and Tunnel Magneto-Resistance (TMR)

The MTJ read is performed by passing a small amount of current through the MTJ layer stacks. It is carried out using a differential amplifier, in which the resistance of the MTJ is sensed and compared with a reference value to generate an output. The relative difference between the two resistance states of the MTJ cell plays a vital role for the output generation, which rely upon the TMR value. The TMR of an MTJ cell is defined as [11]:

$$TMR(\%) = \frac{R_{AP} - R_P}{R_P} \times 100 \tag{1}$$

where R_{AP} and R_P are the resistances in the 'AP' and 'P' magnetization states, respectively. The TMR value is highly dependent on the property of the barrier oxide layer of the MTJ cell [11]. High TMR values are always desirable for fast and reliable read. Furthermore, the read latency exponentially decreases with an increase in TMR, as shown in Figure 2. Please note that a high TMR value for an MTJ cell also requires a high energy to switch its magnetization.

D. Defects in MTJs and fault modeling

The MTJ device uses different materials and processes for manufacturing compared to CMOS. Due to the complexity of these fabrication processes and the interdependency of magnetic materials, MTJ cells are subject to various and new failure mechanisms [5–9]. For instance, during the ion beam etching process, due to sputtering effects, the sputtered atoms re-deposit at the MTJ sidewall leading to a barrier short [6]. As a consequence, the isolation of the ferromagnetic layers is damaged, resulting in a very low resistance value [7]. In this case, although current can flow through the device, the cell no longer behaves as an MTJ cell. On the other hand, an MTJ cell can also have an open connection due to internal damage, which delivers a very high resistance value. As a result, current cannot flow through the device because of a discontinuity in the design.

There are some cases, where the MTJ cell resistance values are not so severely affected as in open and short faults, however, their values can easily influence the sense amplifier so that an incorrect output can be generated. For instance, during fabrication,

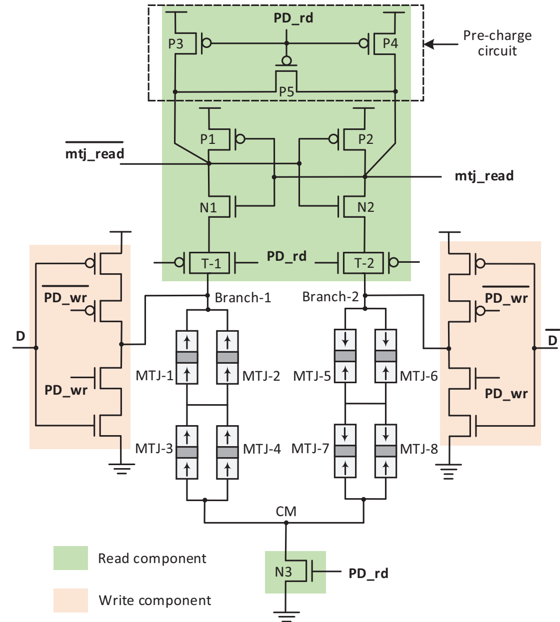

Fig. 3. Schematic diagram of proposed FTNV-L design

the magnetization of the FL can be permanently fixed to either 'P' or 'AP' configuration [10]. Another possibility is that the switching margin and/or the current value are not sufficient enough to flip the magnetization of an MTJ cell [8]. This is due to defects or the impact of process variation. In addition to manufacturing defects, MTJs are also vulnerable to runtime failures such as *read disturb*, *retention failures* and *back-hopping* [9, 12, 13].

We can broadly classify all these MTJ defects into four groups:

- *Short fault*: FL and RL are connected.
- *Open fault*: Discontinuity in the device.
- *Stuck-at-P fault*: MTJ magnetization is permanently or temporally locked to the 'P' state.
- *Stuck-at-AP fault*: MTJ magnetization is permanently or temporally locked to the 'AP' state.

III. Proposed Fault Tolerant Non-Volatile Latch

Here we propose a low cost solution using a novel fault tolerant MTJ-based latch design that can withstand various defects, and deliver a correct output. The implementation details of our proposed latch design, along with its functionality in the presence of all possible faults, are discussed next.

The circuit diagram for our proposed FTNV-L design is shown in Fig 3. It primarily consists of three components, namely, *write*, *read* and *MTJ cell arrangements*. The purpose of the write component is to store the content of the conventional CMOS flip-flop in the MTJ cells during power-down. This can be achieved by establishing a bi-directional current path such that the switching current flows through each MTJ cell. To assure the magnetic switching, the write component has to be designed in such a way that a sufficient amount of switching current for a required duration can flow through each MTJ. This current value is adjusted with the transistor widths in the write components, whereas its duration is synchronized with the 'PD_wr' period. Note that, the main requirement of this write process is that the two branches (i.e., *Branch-1 and Branch-2*) should always have a set of MTJs with opposite magnetizations. This design creates a self-referenced structure which is necessary for a proper read operation.

The read component of the design is composed of a pre-charge circuit, a pair of back-to-back connected inverters and a tail transistor. The purpose of the pre-charge circuit is to provide an equipotential at the output nodes (read_mtj and $\overline{\text{read_mtj}}$) before the actual read is started. In our implementation, read is performed with the activation of the 'PD_rd'. During the read process, the pre-charge circuit is deactivated, and the two back-to-back connected inverters are coupled with the two branches of the MTJ sets, since the transmission gates T1 and T2 are ON. Additionally, the tail transistor 'N3' is also ON at the same time. Therefore, a current path is established, and the sensing process begins. During this sensing process, one of the output nodes goes to a low steady state, while the other remains at a high state. The two back-to-back connected inverters develop a positive feedback loop that accelerates the process of stabilizing the two output nodes.

The arrangement of the MTJ cells is one of the key components in our design implementation. All MTJs in each branch have the same magnetization, and as mentioned previously, the MTJs in those two branches always have the opposite magnetization. The branch in which all MTJs are in 'P' and 'AP' states are referred as *branch-P* and *branch-AP*, respectively. Each branch has a serial connection of the two parallel connected MTJs. This type of arrangement serves two purposes in FTNV-L design: (1) The parallel connection addresses short and open faults. (2) The serial connections are to increase the *ratio of the effective resistance difference* between the two branches, which we named *equivalent TMR* (TMR_{eq}). In other words, the flip-flop design has to meet the minimum TMR_{eq} requirement during the read operation to generate the correct output. Thus, the equivalent resistance for the branch-P is given by the following equation:

$$R_{eq-P} = \frac{R_{P1} \times R_{P2}}{R_{P1} + R_{P2}} + \frac{R_{P3} \times R_{P4}}{R_{P3} + R_{P4}} \quad (2)$$

where R_P is the resistance value of the corresponding MTJ that has 'P' magnetization. Similarly, the equivalent resistance for AP is:

$$R_{eq-AP} = \frac{R_{AP1} \times R_{AP2}}{R_{AP1} + R_{AP2}} + \frac{R_{AP3} \times R_{AP4}}{R_{AP3} + R_{AP4}} \quad (3)$$

where R_{AP} is the resistance of the corresponding MTJ that has 'AP' magnetization. Using the above two equations, TMR_{eq} is defined as:

$$TMR_{eq}(\%) = \frac{R_{eq-AP} - R_{eq-P}}{R_{eq-P}} \times 100 \quad (4)$$

If one MTJ cell has a permanent or temporal defect, the equivalent resistance changes based on the fault type, as discussed next.

Short fault: When one of the MTJs has a short fault, a relatively high current flows though that defective MTJ. Consequently, the MTJ which is in parallel to the shorted one, is bypassed for both read and write operations. Hence, the equivalent resistance for both 'P' and 'AP' is :

$$R_{eq-short\{P,AP\}} = \frac{R_{eq\{P,AP\}}}{2} \quad (5)$$

where $R_{eq\{P,AP\}}$ is the equivalent resistance of either branch-P or branch-AP.

Open fault: When one of the MTJs is open, no current flows through that MTJ. Unlike for shorts, the MTJ which is in parallel to the defective MTJ is usable and it becomes in series with the other two parallel connected MTJs. In this case, the equivalent resistance for both 'P' and 'AP' configuration is:

$$R_{eq-open\{P,AP\}} = \frac{R_{eq\{P,AP\}}}{2} + R_{\{P,AP\}} \quad (6)$$

where $R_{P,AP}$ is the resistance of a single MTJ in either 'P' or 'AP'.

TABLE I. CIRCUIT-LEVEL SETUP

Parameters	Value
VDD and Temperature	1.2 V and 27 °C
CMOS Technology	TSMC 65 nm GP
Thermal stability factor	60
Free/Oxide layer thickness	1.84/1.48 nm
RA	6.145 $\Omega \mu m^2$
TMR @ 0 V	200 %
'AP'/'P' resistance	3.6 KΩ/1.2 KΩ

Stuck-at-P and Stuck-at-AP faults: When one of the MTJ cells is stuck at the 'P' (or 'AP') configuration, then only the 'AP' ('P') branch is affected. Therefore, the equivalent resistance in 'P' and 'AP' branches are as follows:

$$R_{eq-stuck-at-P} = \frac{R_{AP}}{2} + \frac{R_P \times R_{AP}}{R_P + R_{AP}} \quad (7)$$

$$R_{eq-stuck-at-AP} = \frac{R_P}{2} + \frac{R_{AP} \times R_P}{R_{AP} + R_P} \quad (8)$$

IV. EXPERIMENTAL SETUP AND RESULTS

We performed a circuit-level analysis in order to evaluate the efficiency of our proposed FTNV-L design. The simulation setup is discussed first, followed by the circuit-level results. In the end, a comparison of our proposed technique with 3MR is performed.

A. Simulation setup

For the circuit design implementation, we employed the MTJ model presented in [14], and the other design parameters for the simulations are depicted in Table I. Here, our MTJ model is tuned for the TMR and resistance values specified in the table, which are determined using Eq (5)-(8), with the assumption of an acceptable TMR_{eq} of 50 %.

The resistance value associated with each MTJ is obtained by measuring the current value and voltage across its terminals. Furthermore, to obtain a setup for the defective MTJ cell, we employed a resistance device to replace the MTJ in the design. For instance, a low (around 5 Ω) and a high (around 5 MΩ) resistances are connected to demonstrate the short and open faults, receptively. Similarly, to show the stuck-at-P and stuck-at-AP behavior in the design, a resistance value equivalent of R_P and R_{AP} is connected, respectively. Please note that only one resistance at a time is connected, as our design targets a single fault per latch.

B. Circuit-level design analysis

The TMR value is very sensitive to MTJ defects, which in turn influences the functionality of the design. In our proposed design, we have performed a detailed TMR_{eq} analysis and the results for both branch-P and branch-AP are demonstrated in Fig 4. As shown in the figure, the worst TMR_{eq} value is obtained for open and short faults in branch-P and branch-AP, respectively. The functionality of the FTNV-L design in the presence of a short fault which is the worst among all faults, is demonstrated in Fig 6. In this figure, the read outputs and effective resistances for both branches are shown. This figure clearly shows that the TMR_{eq} is relatively low when '0' value is read, but good enough to read the output correctly. The reason for low TMR_{eq} is that the effective resistance of branch-AP becomes low due to an MTJ short. On the other hand, in the presence of short-P and open-AP, TMR_{eq} is high, even more than that of a fault-free MTJ cell. This is because the faulty MTJs in these two cases are additive to the resistance differences which further increases the overall effective resistance. Moreover, the TMR_{eq} value for both stuck-at faults is slightly less than that of the TMR value of an MTJ. These TMR_{eq} values influence the delay of the restore (read latency)

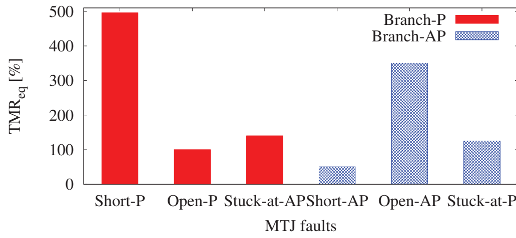

Fig. 4. TMR values in the presence of various faults

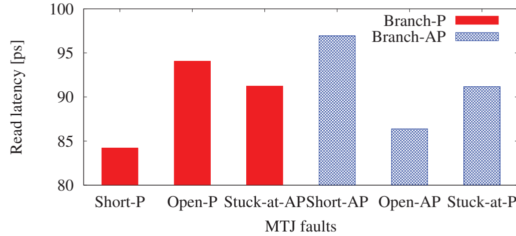

Fig. 5. Read latency values in the presence of various faults

operation of the FTNV-L as demonstrated in Fig 5. Since, the read latency is inversely proportional to the TMR_{eq} value, short-P has the lowest and short-AP has the highest delay.

In addition to the design parameters, we conducted an area analysis for our proposed FTNV-L design. Compared to a standard latch design, the area of FTNV-L is impacted by two factors: (1) transistor width of the driver circuitry is increased, and (2) six additional MTJs are employed. In the first case, the drive strength of the write component is increased by 3.4 X to pass more current (2 X) through the MTJ branches. However, in the second case, MTJs are fabricated in another layer, and additionally flip-flops are widely distributed all over the logic core unlike memory bit-cells. Hence, adding extra MTJs introduces negligible impact on the overall area of the SoC design [15].

C. Comparison with triple modular redundancy

To illustrate the advantages of our proposed FTNV-L design, we compare it with a standard latch as well as 3MR. For the standard NV latch implementation, we use only two MTJs, one per each branch. For the 3MR implementation, we employ three standard NV latch designs with a voting circuit. The results of comparison for the three designs for a normal operation are summarized in Table II.

As specified in the table, TMR_{eq} value for each design is the same for the normal operation when fault-free MTJs are considered.

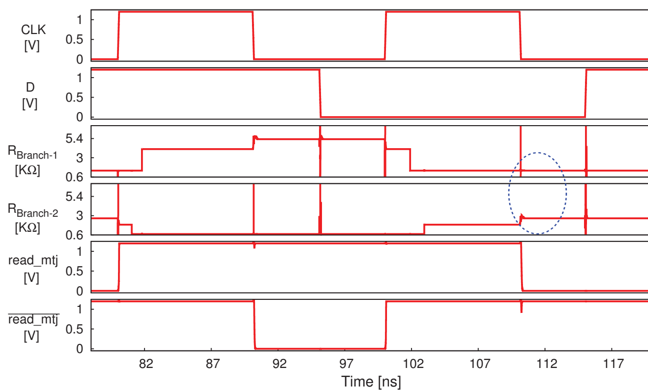

Fig. 6. Functionality of FTNV-L in the presence of short fault (for typical process corner and see Table I for setup information). Blue dotted circle indicates the worst resistance differences during read due to short fault.

TABLE II. Comparison of standard latch and 3MR design with proposed FTNV-L design (†:Transistor width are increased by 3.4 X compared to standard NV-latch)

Parameters	Standard NV Latch	3MR	Proposed FTNV-L
TMR_{eq} (%)	200	200	200
Restore Latency (ps)	83	203	89
Restore Energy (fJ)	12	42	15
Backup delay (ps)	4056	4056	4065
Backup energy (fJ)	390	1170	366
Transistor count	16	72	16 †
MTJ count	2	6	8

However, in the presence of defective MTJs, the standard latch design is not functional at all, whereas our proposed FTNV-L design as well as 3MR are able to generate fault-free output. Both of these designs can address a single MTJ fault per latch, but the 3MR has huge overheads because it uses three sets of standard NV latch designs. For instance, compared to FTNV-L, the 3MR design has around 2 X and 3 X overheads for the restore latency and energy, respectively. The voter circuit in 3MR adds 120 ps to the delay and consumes 6.4 fJ energy during restore. On the other hand, our proposed FTNV-L has almost similar results in comparison to the standard NV latch design. Nevertheless, the drive strength of the write components are increased by 3.4 X in FTNV-L design compared to the standard NV latch design. This is because, in FTNV-L several MTJs are stacked in series-parallel connection, hence more current needs to pass through write components to assure MTJ switching.

V. Conclusions

Non-volatile shadow latches are beneficial for leakage power reduction based on normally-off computing. However, the non-volatile MTJ storing device is susceptible to new failure mechanisms which can severely affect the manufacturing yield and correct in-field functionality of these systems. We proposed a *Fault Tolerant Non-Volatile Latch* (FTNV-L) to preserve the fault-free functionality of the latch in the presence of various faults. In our proposed FTNV-L, any single fault per latch can be tolerated at much reduced costs compared to traditional solutions based on triple module redundancy.

VI. Acknowledgement

This work was partly supported by the European Commission under the Seventh Framework Program as part of the spOt project (http://www.spot-research.eu/).

References

[1] International Technology Roadmap for Semiconductors. http://www.itrs.net, 2013.
[2] F. Oboril, et al. Evaluation of hybrid memory technologies using sot-mram for on-chip cache hierarchy. *TCAD*, 2015.
[3] S. Yamamoto, et al. Nonvolatile flip-flop using pseudo-spin-transistor architecture and its power-gating applications. In *ISCDG*, pages 17–20, 2012.
[4] Y. Lakys, et al. Low power, high reliability magnetic flip-flop. *EL*, 2010.
[5] R. Robertazzi, et al. Analytical mram test. In *ITC*, 2014.
[6] K. Sugiura, et al. Ion beam etching technology for high-density spin transfer torque magnetic random access memory. *JJAP*, 2009.
[7] G. Panagopoulos, et al. Modeling of dielectric breakdown-induced time-dependent stt-mram performance degradation. In *DRC*, 2011.
[8] M. Seyedhamidreza, et al. Impact Of Process-Variations in STTRAM and Adaptive Boosting for Robustness. In *DATE*, 2015.
[9] H. Naemi, C. Augustine, A. Raychowdhury, S. Lu, J. Tschanz. STTRAM Scaling And Retention Failure. *Intel Technology Journal*, 17, 2013.
[10] M. Kuepferling, et al. Vortex dynamics in co-fe-b magnetic tunnel junctions in presence of defects. *JAP*, 2015.
[11] A. Khvalkovskiy, et al. Basic principles of stt-mram cell operation in memory arrays. *JAP*, 2013.
[12] R. Bishnoi, et al. Read disturb fault detection in stt-mram. In *ITC*, 2014.
[13] T. Min, et al. Back-hopping after spin torque transfer induced magnetization switching in magnetic tunneling junction cells. *JAP*, 2009.
[14] A. Mejdoubi, et al. A compact model of precessional spin-transfer switching for mtj with a perpendicular polarizer. In *MIEL*, 2012.
[15] G. Prenat, et al. Hybrid cmos/magnetic process design kit and application to the design of high-performances non-volatile logic circuits. In *ICCAD*, 2011.