

A Survey of Test and Reliability Solutions for Magnetic Random Access Memories

Girard, Patrick; Cheng, Yuanqing; Virazel, Arnaud; Zhao, Weisheng; Bishnoi, Rajendra; Tahoori, Mehdi B.

DOI

[10.1109/JPROC.2020.3029600](https://doi.org/10.1109/JPROC.2020.3029600)

Publication date

2020

Document Version

Accepted author manuscript

Published in

Proceedings of the IEEE

Citation (APA)

Girard, P., Cheng, Y., Virazel, A., Zhao, W., Bishnoi, R., & Tahoori, M. B. (2020). A Survey of Test and Reliability Solutions for Magnetic Random Access Memories. *Proceedings of the IEEE*, 109 (2021)(2), 149-169. Article 9240959. <https://doi.org/10.1109/JPROC.2020.3029600>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



A Survey of Test and Reliability Solutions for Magnetic Random Access Memories

By PATRICK GIRARD^{ID}, *Fellow IEEE*, YUANQING CHENG^{ID}, *Senior Member IEEE*,
ARNAUD VIRAZEL^{ID}, *Member IEEE*, WEISHENG ZHAO^{ID}, *Fellow IEEE*, RAJENDRA BISHNOI,
AND MEHDI B. TAHOORI^{ID}, *Senior Member IEEE*

ABSTRACT | Memories occupy most of the silicon area in nowadays' system-on-chips and contribute to a significant part of system power consumption. Though widely used, nonvolatile Flash memories still suffer from several drawbacks. Magnetic random access memories (MRAMs) have the potential to mitigate most of the Flash shortcomings. Moreover, it is predicted that they could be used for DRAM and SRAM replacement. However, they are prone to manufacturing defects and runtime failures as any other type of memory. This article provides an up-to-date and practical coverage of MRAM test and reliability solutions existing in the literature. After some background on existing MRAM technologies, defectiveness and reliability issues are discussed, as well as functional fault models used for MRAM. This article is dedicated to a summarized description of existing test and reliability improvement methods developed

so far for various MRAM technologies. The last part of this article gives some perspectives on this hot topic.

KEYWORDS | Magnetic random access memory (MRAM); nonvolatile memories; reliability; spintronics; test.

I. INTRODUCTION

Spin electronics (Spintronics) is one of the most interesting and challenging topics in today's nanotechnology. It has pushed scientific research and microelectronic industry to build innovative electronic devices that rely on magnetic properties. Similar to other emerging resistive memory technologies, such as resistive random access memory (RRAM) [1] and phase-change memory (PCM) [2], magnetic or magneto-RRAM (MRAM) is a form of resistive memory technology where data are stored in terms of resistive states. Moreover, MRAM uses the spin of electrons for storage instead of their charge. Comparisons of different kinds of memory technologies are shown in Table 1. As illustrated in this table, MRAM technology has demonstrated the promise of universal memory. MRAM has several characteristics that make it is useful for many applications. Nonvolatility (ability to maintain memory contents without requiring power), performance (SRAM and DRAM-like speed with low latency), endurance (durability to support memory workloads without complex management), and reliability (robustness designed for extreme conditions) are these main characteristics. Moreover, an important feature of MRAM technology is that its fabrication process is CMOS-compatible [3].

Manuscript received January 21, 2020; revised July 9, 2020 and September 24, 2020; accepted October 1, 2020. The work of Yuanqing Cheng was supported in part by Science, Technology and Innovation Commission of Shenzhen Municipality under Grant JCYJ20180307123657364. (Corresponding author: Patrick Girard.)

Patrick Girard and **Arnaud Virazel** are with the Laboratory of Computer Science, Robotics and Microelectronics of Montpellier (LIRMM), University of Montpellier/CNRS, 34095 Montpellier, France (e-mail: girard@lirmm.fr; virazel@lirmm.fr).

Yuanqing Cheng is with the School of Microelectronics, Beihang University, Beijing 100191, China (e-mail: yuanqing@ieee.org).

Weisheng Zhao is with the Fert Beijing Institute, Beihang University, Beijing 100191, China, and also with the School of Microelectronics, Beihang University, Beijing 100191, China (e-mail: weisheng.zhao@buaa.edu.cn).

Rajendra Bishnoi is with the Computer Engineering Lab, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: r.k.bishnoi@tudelft.nl).

Mehdi B. Tahoori is with the Department of Computer Science, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany (e-mail: mehdi.tahoori@kit.edu).

Digital Object Identifier 10.1109/JPROC.2020.3029600

Table 1 Comparisons of Different Memory Technologies [5]–[7]

	SRAM	DRAM	NAND	PCRAM	RRAM	STT-MRAM
Cell size (F^2)	~ 150	~ 8	≤ 1	~ 5	~ 4	$6 \sim 8$
Non-Volatility	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Read Time	$\sim ns$	$\sim 30ns$	$\sim 50\mu s$	$\sim 30ns$	$20ns$	$10 \sim 20ns$
Write Time	$\sim ns$	$\sim 30ns$	$\sim ms$	$\sim 500ns$	$0.3 \sim 30ns$	$10 \sim 20ns$
Endurance	$> 10^{15}$	$> 10^{15}$	10^5	$\sim 10^7$	$10^6 \sim 10^{12}$	$> 10^{15}$
Byte Operation	<i>Yes</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>

High-density MRAM can be seen as a replacement for SRAM applications, such as cache memories. In addition to providing fast (low latency) caches, MRAM is nonvolatile and, thus, allows instant power on as it does not require capacitances to save data when power is removed. Because it can reduce latency for various devices, MRAM can help in data centers, at the edge of networks and network endpoints. This is especially important as high-speed wireless networks (e.g., 5G) become more common. The vast majority of these applications use limited energy sources (e.g., batteries), and hence, MRAM nonvolatility can be helpful [4]. MRAM can also be used as embedded memory, where macros are embedded or integrated with microcontroller units (MCUs). Finally, MRAM can be used for DRAM replacement in mission-critical enterprise applications, where power loss and lost memory can severely impact a client.

Automotive, the Internet of Things (IoT), and many other applications drive the market growth for MRAM. The global MRAM market is expected to reach \$4.8 billion by 2025 [8]. This growth is primarily attributed to the rising demand for power-efficient, cost-effective, and nonvolatile memory in many end-user industries, surging demand for flexible and wearable electronics, and increasing research and development activities. In addition, advancement in technology is further projected to augment market growth during the forecast period, as well as the huge demand for embedded applications.

MRAM performance, low power, and persistence are the major reasons for its use for many applications. For example, MRAM can be used in extreme-low-power designs, such as wearables, RFID-based applications, such as trackers, and performance-constrained domains, such as cloud applications. As mentioned earlier, an example is data centers, where power is the highest part of their operational costs.

Due to its nonvolatility, MRAM technology is also definitively recognized as the best alternative to the Flash memory technology that reaches its limits due to intrinsic variability issues or challenging cointegration with the CMOS process. It offers a number of advantages compared with the Flash technology, such as higher read/write speed, lower power consumption, longer endurance, better reliability (mainly due to its immunity to radiations), higher integration density and scalability, and new functionalities, such as computing-in-memory [9]–[11], neuromorphic

computing [12]–[14], true random number generation (TRNG), and physically unclonable function (PUF) for security [15], [16]. Magnetic devices can be used not only for standalone or embedded memories but also in the logic itself (e.g., nonvolatile decoders) [8].

Though MRAMs have the potential to mitigate almost all shortcomings of Flash and compete with other types of memory technology, however, they are as prone to defectivity and reliability issues as any other kind of memory. For this reason, a number of test and reliability improvement solutions targeting various MRAM technology generations have been developed in the last decade. These solutions were presented very often with interesting case studies and convincing experimental results.

This article surveys the existing MRAM test and reliability improvement solutions published so far. These solutions deal with the development of test algorithms (March-type, retention, thermal stability, and so on) and their implementation [e.g., memory built-in self-test (MBIST)]. Defect analysis and fault modeling are a prerequisite to the development of these solutions and, hence, are discussed in this article. Similarly, reliability concerns have been addressed in the literature through the development of Design-for-Reliability (DfR) solutions that target process variability, endurance degradation, and data disturbance. These solutions are summarized and discussed in this article. Note that some test and reliability solutions may also apply to other resistive emerging memory devices, such as RRAM and PCM. We will discuss the relevance and generalities of these solutions in appropriate places to provide the reader with an overall perspective of test and reliable design for emerging nonvolatile memory technologies.

The rest of this article is organized as follows. Section II presents an overview of the various MRAM technologies existing today. Section III discusses the defectiveness and reliability issues that may occur in MRAMs. Section IV surveys all existing test solutions for this type of memory. Similarly, Section V gives an overview of all existing reliability improvement solutions. Section VI concludes this article.

II. MRAM TECHNOLOGIES

The magnetic tunnel junction (MTJ) is the basic block of an MRAM and uses the magnetoresistance property of a material to change the value of its electrical resistance. Based on this principle, several MRAM technologies have

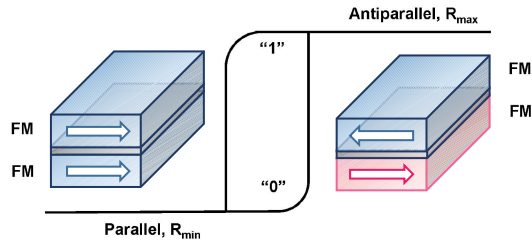


Fig. 1. MTJ in parallel and antiparallel states.

been proposed in the last years. They are presented in the following.

A. Magnetic Tunnel Junction

MTJs are spintronic devices that can be used in numerous applications, such as sensors and oscillators, and the basic building block of nonvolatile memories [17]. An MTJ usually consists of two ferromagnetic (FM) layers separated by an ultrathin insulating layer in which electrons can tunnel. One of the FM layers is pinned and acts as a reference layer. The other one is free and can be switched between at least two stable states. Hence, this device presents a magnetoresistive effect, which depends on the relative magnetization state of the FM layers. Fig. 1 shows a basic MTJ device. These states are parallel or antiparallel with respect to the reference layer. MTJ offers maximum resistance to electric current (R_{\max}) when the magnetizations of the FM layers are in an antiparallel configuration. Conversely, this device offers minimum resistance (R_{\min}) in a parallel configuration. The tunnel magnetoresistance (TMR), which is the resistive effect that occurs in these magnetic devices, quantifies the difference between R_{\min} and R_{\max} and can be defined as follows [18]:

$$\text{TMR} = \frac{R_{\max} - R_{\min}}{R_{\min}} \times 100\%. \quad (1)$$

A read operation consists in determining the magnetization state of the MTJ and can be done by voltage or current sensing across the MTJ stack. A CMOS-based sense amplifier is employed to retrieve the stored bit information. High TMR allows simple and stable sense amplifiers, improving the reading accuracy. A write operation is usually performed by using magnetic fields or spin-polarized current depending on the MRAM technology. This is discussed in Section II-B.

B. Existing MRAM Technologies

MRAM technologies can be classified by considering the switching method used to write data [19]. The first-generation MRAM refers to methods using magnetic fields to program (write-in) the array. As shown in Fig. 2(a), field-induced magnetic switching (FIMS) and FIMS-Toggle can be written by applying magnetic

fields generated by two current lines [20], [21]. The energy required to reverse the magnetization state in these MRAMs is minimized by concurrently applying these two perpendicular magnetic fields. A major advantage of field switching is the unlimited write endurance since reversing the magnetization of the free layer with a magnetic field does not induce any wear-out effects. On the other side, the FIMS drawback is the selectivity issue, which is the ability to write the selected MTJs without disturbing the other MTJs (this is called write margin). Another drawback is the scalability issue, which is mainly due to the magnitude of the required switching currents and the complexity of the memory cell geometry.

Another technology based on the same writing principle is the thermally assisted switching (TAS) MRAM technology [22]. Here, the MTJ is modified by inserting an anti-FM (AFM) layer that pins the storage layer while below its blocking temperature T_B . As shown in Fig. 2(b), when the temperature of the MTJ rises above T_B , the storage layer is freed. Hence, it can be reversed under the application of a small magnetic field provided by a single field line. This field is maintained beyond the heating voltage pulse to ensure the correct pinning of the storage layer. Though it has several advantages compared with Toggle-MRAM, such as reduced selectivity issues and improved integration density due to: 1) thermal stability and 2) the need for only one field line, TAS-MRAM still suffers from structure complexity and area cost issues.

An evolution of the TAS-MTJ is the magnetic logic unit (MLU) MTJ in which the AFM layer is replaced by a nonmagnetic (NM) layer. The formerly fixed reference layer, now called the sense layer, may change its magnetization state by the application of an external magnetic field. The writing process does not change with respect to TAS-MTJ, but the read procedure now relies on a sample and hold approach. The main advantages are an increasing density with less neighboring cell disturbances and an enhanced read margin [23].

The second-generation MRAM uses a spin-polarized current through the MTJ to write data. Spin transfer torque (STT) switching was the first proposed principle for this type of MRAM. As shown in Fig. 3(a), the magnetic moment in the fixed layer is fixed in one direction, while the direction of the magnetic moment in the free layer can be changed according to the difference and polarity of potential across the MTJ. Indeed, when a potential difference is applied across the MTJ, a spin-polarized current passes through the MTJ and polarizes the current in its preferred direction of the magnetic moment. The angular momentum of the electrons in the free layer creates a torque, which causes a change in the direction of the magnetization inside the free layer. Depending on the magnetization direction in the free layer, the resistance of the MTJ is modified. STT switching can be achieved with acceptable efficiency by using MTJ devices having either planar or perpendicular magnetization. The main advantages of STT-MRAM are fast read/write speed, high density, high

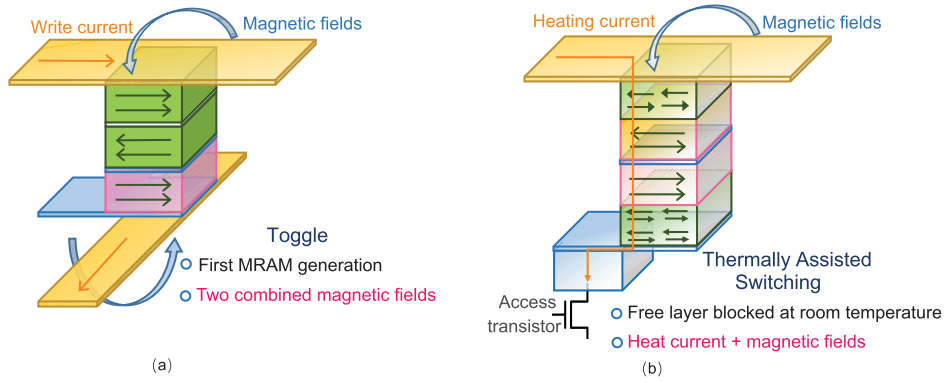


Fig. 2. MTJ structures of (a) Toggle-MRAM and (b) TAS-MRAM.

endurance, and high reliability. Especially, the TMR ratio in the common stacks involved with MgO/CoFeB interfaces has recently been demonstrated to be higher than 200% [24], which can reach the technological requirements [25]. Despite several remaining challenges, such as patterning process, read and write error rates, or long-term data retention, STT-MRAM technology is now entering into high-volume mass production serving markets from 2018. A number of IoT products embedded with STT-MRAM have been commercialized to provide long standby duration, such as the smartwatch. However, applications are limited to some niche markets as its lifetime is not yet fully satisfactory [26]. This reliability issue is due to the intrinsic mechanism of STT where the switching current should pass through the tunneling barrier. Many studies are focusing on the circuit- and system-level controls to reduce the number of switching operations for STT-MRAM-based computing systems, but it is difficult to achieve both fast speed and high reliability.

To mitigate some of the remaining issues of STT-MRAM, spin-orbit torque (SOT) MRAM has been proposed. As shown in Fig. 3(b), SOT-MRAM uses a three-terminal MTJ-based concept to isolate read and write paths, significantly improving device endurance and read stability. Unlike STT, magnetization reversal by spin-orbit torque is performed using in-plane currents. Due to the nature

Table 2 Comparisons of Different MRAM Technologies [30], [31]

	FIMS	TAS-MRAM	STT-MRAM	SOT-MRAM
Read Latency	35ns	~30ns	<10ns	<10ns
Write Latency	35ns	~30ns	<10ns	<1ns
Retention	>10y	>10y	>10y	>10y
Endurance	10^{15}	10^{15}	> 10^{15}	> 10^{15}
Write Energy	High	Medium	Low	Low
Density	Medium	Medium	High	Medium
TMR	>100%	>100%	>100%	~90%
Scalability	Medium	Medium	High	High

of the spin torque injection geometry (perpendicular to the easy axis), the incubation time is negligible and allows reliable switching operation in the sub-nanosecond timescale. However, in an MTJ with perpendicular magnetic anisotropy, an external magnetic field is necessary for SOT to realize deterministic switching. Similar to FIMS, this external magnetic field in MRAM will severely limit its scalability and reliability. To solve this problem, several switching solutions were proposed [27]–[29]. Among them, the magnetic switching realized by the interplay of STT and SOT could not only solve the external field problem of SOT but also reduce the amplitude and duration of STT current [27]. Therefore, it can have a longer endurance since the current-induced barrier breakdown is mitigated. The comparison of FIMS, TAS-MRAM, STT-MRAM, and SOT-MRAM is summarized in Table 2.

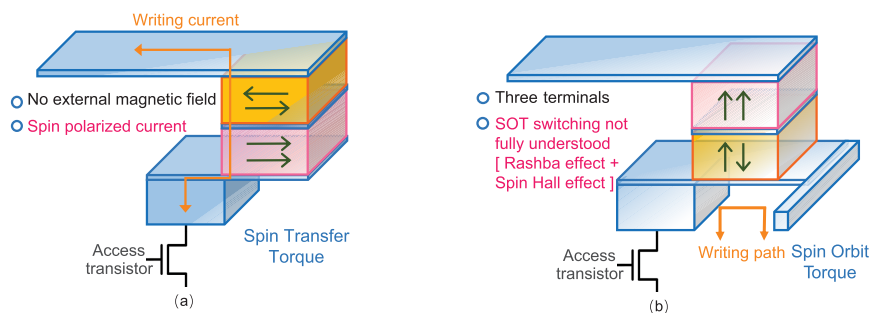


Fig. 3. MTJ structures of (a) STT-MRAM and (b) SOT-MRAM.

The third-generation MRAM refers to the potential use of other physical phenomena, including voltage-controlled magnetic anisotropy (VCMA), voltage-controlled magnetism (VCM), and spin Hall effect (SHE) [32]. The idea behind the use of these phenomena is that switching could be performed with little or no electrical current passing through the MTJ device, hence improving MRAM scaling and performance. However, each physical phenomenon has challenges to overcome for use in practical MRAM circuits [19]. For example, VCMA alone would not lead to deterministic switching between two stable states. Instead, it is more likely to be applied together with another innovation for use in an MRAM array. A better understanding of reliability issues, such as wear-out and parameter drift, is needed for practical VCM devices. Similarly, SHE is not compatible with efficient switching in devices having perpendicular magnetization. In many cases, a three-terminal cell configuration would be needed for these third-generation devices, which is not compatible with high-density memory arrays. However, the possibility of switching with little or no charge current passing through the tunnel barrier is highly motivating due to the further possible use in low-power circuits and high-endurance applications.

Besides these existing MRAM technologies, and in a wider perspective that intends to provide a complete logic/memory family with better energy and delay performances, a new concept has emerged recently. It consists of a scalable, CMOS-compatible, nonvolatile spintronic logic device that operates via spin-orbit transaction combined with magnetoelectric switching. No TMR or STT is used. The proposed magnetoelectric spin-orbit (MESO) logic enables a new paradigm to continue the scaling of logic device performance. More details can be found in [33]. It is important to note that racetrack/skyrmion and multilevel cells are interesting research directions beyond MRAM development. However, these technologies are still under the first step of R&D to demonstrate a working device [34], and there are no racetrack/skyrmion and multilevel devices showing electrical read-out switching without magnetic field up until now [35]. There are no test and reliability methods for the chip level on racetrack/skyrmion and multilevel cells.

III. DEFECTIVENESS AND RELIABILITY ISSUES OF MRAMs

Defects and failures in MRAM may occur during the manufacturing process or during the lifetime of the memory. In the first case, we refer to defectiveness issues. In the second case, we refer to reliability issues. Note that these reliability issues may be the consequence of wear-out or aging mechanisms or inherent MTJ device parameter instability but can also be provoked by parameter deviations coming from an improper but acceptable manufacturing process and that will evolve over time. All these aspects are discussed in the following. Of course, there are differences in manufacturing steps and processes of different flavors of

MRAM, which can lead to unique manufacturing defects and runtime failures for each of these specific technologies. Nevertheless, although most of the defectiveness and reliability issues discussed in this section have been reported in studies dealing with STT-MRAM, they can be considered as general issues that may occur in all types of MRAM, regardless of the specific technology. Besides, some kinds of defects, such as resistive-open and resistive-short, can also be observed in RRAM and PCM due to immature fabrication process or process variations [36].

A. Defectiveness Issues

Defectiveness issues of MRAM can be attributed to imperfect manufacturing process during which strong or weak defects may occur or process variability is too severe and drastically impacts device parameters. These issues can be fundamentally different from those observed on other memory technologies since both working principles and development processes for this technology are completely dissimilar. Zhao *et al.* [37], [38] classified MRAM errors into “hard errors” (similar to strong or weak defects) and “soft errors” (similar to parametric failures).

1) *Strong or Weak Defects*: These defects are the consequence of issues that occur during the fabrication steps. They can be caused by the deposition of dust particles, voids occurring during the electrode polishing, oxide barrier breakdown, or improper etching followed by sideways material redeposition [39]. These defects can be strong or weak defects depending on their influence on the system. The strong defects prevent the device to operate correctly and lead to permanent failures (hard faults). Examples are open, bridge, or parasitic coupling defects. Weak defects prevent the device to operate correctly depending on the size of the defect and also lead to permanent failures. Examples are resistive open, resistive bridge, capacitive defects, and so on. These defects can be caused by the same phenomena (voids, improper etching, and so on) except that, in this case, the MTJ device can still work but only for a given range of defect sizes.

Other strong or weak defects may occur because of manufacturing issues in the magnetic layers or due to loss of margin in the CMOS circuitry. In this case, the magnetic orientation of the MTJ cells is fixed to a specific configuration so that their magnetic orientation (i.e., resistances) cannot be changed [40]. Note that these defects are sometimes classified in the literature as intracell (within a cell) or intercell (cell-to-cell) defects [41], [42]. Intracell defects refer to resistive opens and shorts on lines inside a cell. Intercell defects refer to resistive shorts between the nodes of the victim cell and those of an aggressor cell.

2) *Parametric Failures*: Manufacturing process variation is another source of defectiveness issues. As the manufacturing of MRAM requires two different fabrication process technologies, namely a magnetic process for the MTJs and a CMOS process for the access transistors and peripheries, the characteristics of this memory technology can be

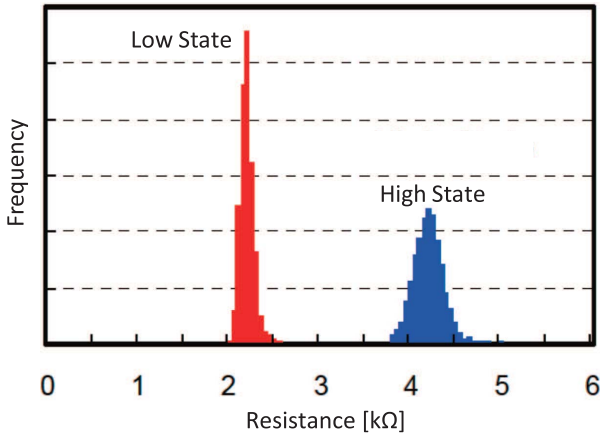


Fig. 4. Low- and high-resistance distributions of MTJs [48].

affected by variations due to the combined effect of these two processes. Permanent faults in MRAM can be caused by extreme parametric variations, as described in [43] and [44]. These variations come from changes in both material and lithographic properties, transistor electrical properties, and noise generated by thermal effects [41], [42]. MTJ material parameters that may suffer from variations are magnetic anisotropy, saturation magnetization, TMR ratio, and oxide thickness of the ultrathin insulating layer. Lithographic parameters that may suffer from variations are the planar dimensions of the MTJ and the length and width of the access transistor. The main transistor electrical parameter that may suffer from variations is the threshold voltage of the access transistor. These variations not only significantly affect the memory operations, such as read/write delays [45]–[47] and retention capabilities, but also lead to hard faults making the cell permanently damaged.

B. Reliability Issues

Reliability is a key issue for the commercial success of MRAM technology. Various reliability issues associated with MRAM are discussed subsequently.

1) *Impact of Process Variations*: The access behavior of MRAM highly depends on the properties of the MTJ device, which can be affected by variations in its geometrical dimensions, such as its cross-sectional area, the tunneling oxide thickness, and the volume of the FM layer. As described previously, variations in these parameters result in significant deviations in MTJ conductance and switching threshold current. For instance, as per the resistance distributions shown in Fig. 4, the resistance values can be deviated up to 4% [48] from its nominal values. Due to the influence of process variation, the MTJ properties, such as resistance, TMR ratio, and switching delay, can be affected [45]–[47]. At the manufacturing level, the reliability of an MRAM cell can be degraded

by: 1) device parameter deviations due to errors of lithography or etching process; 2) thermal disturbance [38]; and 3) dielectric breakdown. These phenomena may lead to TMR fluctuation, endurance degradation, data disturbance, retention failure, and so on [49], [50], which, in turn, may lead to access errors during the MRAM operations.

2) *Stochastic Switching*: The switching of an MTJ is stochastic in nature [51]–[53], and the switching time is randomly distributed. Nevertheless, the write time, which is primarily dependent on the clock period, has a fixed value for a synchronous design. Hence, due to the stochastic switching nature, some memory cells may not complete state transitions in a given write period, which can result in write errors [54], [55]. The write error rate can be expressed as follows [52]:

$$\text{WER}_{\text{bit}}(t_w) = 1 - \exp \left[\frac{-\pi^2 \cdot (I - 1) \cdot \Delta}{4(I \cdot e^{C(I-1)t_w} - 1)} \right] \quad (2)$$

$$I = \frac{I_w}{I_c} \quad (3)$$

where t_w is the write latency, C is a technology-dependent parameter, I is the ratio of the write current (I_w) to the critical current (I_c), and Δ is the thermal stability factor. The switching success probability is a function of switching current, write period, thermal stability factor, and other material parameters. The write latency distribution for a single STT-MRAM cell used as a case study is shown in Fig. 5. As shown in the figure, these distributions have very long tails [51], and hence, a large write period is required to achieve the specified write reliability. Note that increasing the write current and/or write pulse duration are the most effective methods to reduce write failure.

3) *Retention Failure*: Retention failure is most likely due to the inherent thermal instability of MRAM, which can lead to MTJ resistance state switching without memory access. The retention failure mechanism can be formally

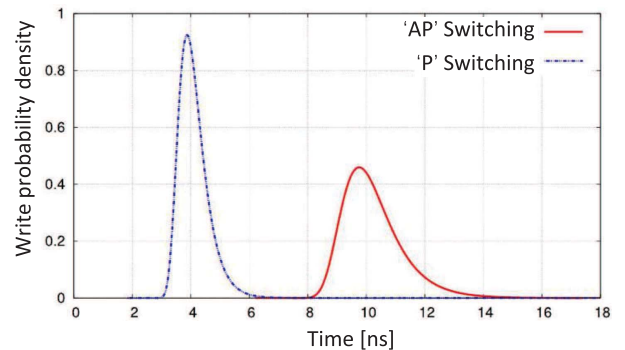


Fig. 5. Write latency distribution due to stochastic behavior for both antiparallel and parallel switching of a single-bit cell [51].

Table 3 Typical FFM in an STT-MRAM [41], [42]

Fault Model	Affects	Key Cause
Transition Fault (TF)	Write	Relative weak write current due to stray resistive paths
Coupling Fault (CF)	Write	Neighboring cells switching
Stuck-at Fault (SF)	Write	Intermediate node, word-line stuck-at VDD or GND
Incorrect Read Fault (IRF)	Read	Current miscorrelation due to defects affecting word-line and bit-line
Read Disturb Fault (RDF)	Read	Electrical disturbance at intermediate node due to larger than normal read current

expressed as follows:

$$P_{rf} = 1 - \exp\left(-\frac{t_{rf}}{\tau \cdot e^{\Delta}}\right) \quad (4)$$

where P_{rf} is the failure probability for a specific retention time, that is, t_{rf} , τ is the attempt period (about 1 ns), Δ is thermal stability factor, and $\tau \cdot e^{\Delta}$ is the average retention time. As shown in the formula, a higher thermal stability Δ can result in a longer retention time.

4) *Read Disturbance*: A read disturbance occurs when an MTJ cell is accidentally switched during a read operation [56], [57]. Note that an incorrect read value may also occur due to a read decision failure. The read disturbance can be described as

$$P_{rd} = 1 - \exp\left(-\frac{t_{read}}{\tau} \times \exp\left[\frac{\Delta(I_{read} - I_c)}{I_c}\right]\right) \quad (5)$$

where I_{read} is the read current, I_c is the critical switching current, t_{read} is the read pulse duration, and Δ is the thermal stability factor.

5) *Other Transient Failures*: Some other issues may occur during the manufacturing process that may lead to the above-mentioned transient failures. Unlike permanent failures, these failures can be recovered after their occurrences. They are called transient faults, and the correct state of the cell can be eventually recovered in the following memory accesses [39]. These faults can be caused by defects introduced during the polishing process or the magnetic stack deposition and annealing process [58].

To summarize, uncertainties in reliability can lead to performance degradation, higher production cost, and time-to-market penalties. Therefore, it is necessary to pinpoint and address these reliability issues for MRAM chip design to guarantee the final product yield.

C. MRAM Fault Models

In general, faults in memories are modeled as functional faults in which functional tests can be used to detect those faults. These functional tests for a given set of functional fault models (FFMs) can be performed using a systematic approach that is essential to increase the yield and reliability of memories. Functional faults refer to the deviation of the observed memory behavior from the functionally

specified one under a set of operations. There are a list of operations that are also known as operation sequence and a list of corresponding deviations that are termed faulty behavior for each FFM [51].

FFMs in MRAMs can be classified as static or dynamic and single or double cells impacting, depending on their behavior. Static single-cell FFM describe faults sensitized using a single operation on a faulty cell. Conversely, dynamic FFM are faults sensitized by performing more than one operation in sequence. On the other hand, double-cell FFM are faults consisting of two-cell fault primitives in which the victim cell is the one that shows the faulty behavior, whereas the aggressor cell is the one that produces this behavior. A comprehensive description of all these FFM for SRAM that also applies to MRAM can be found in [59].

The most common types of MRAM FFM are the following: Stuck At "0" (SAF0), Stuck At "1" (SAF1), and transition fault (TF), which leads to permanent faults. Undefined write fault (UWF), write disturb fault (WDF), read disturb fault (RDF), incorrect read fault (IRF), and retention fault (RF) are those leading to disturb faults [39], [41], [60]. A detailed description of the nonfunctional behavior associated with each of these FFM can be found in [39]. As an example, Chintaluri *et al.* [41], [42] have identified FFM that may occur during read and write operations in an STT-MRAM and how resistive and capacitive defects induce faults. Table 3 summarizes the FFM and how they contribute to read/write failures. Moreover, in [61], based on the defect injection scheme, defect characterization and fault modeling considering layouts for STT-MRAM are presented. In this analysis, a dynamic read fault behavior is demonstrated in which multiple vectors for excitation and detection are required. This read fault behavior happens in the case of intercell coupling faults (CFs).

Note that several other FFM can be found in the literature for different types of MRAM technologies. In [62], two new FFM related to the magnetic junction behavior of Toggle-MRAM were identified. They were proposed to model multivictim fault (MVF) and kink fault (KF). In MVF, a cluster of cells can change their magnetization state due to the impact of process variations, whereas, in KF, the MTJ resistivity changes due to the shrinkage of the hysteresis loop because of its relation with the cell shape. In [63], transition CFs and incorrect read CFs were found and modeled for TAS-MRAM. Inversion CF (ICF) in two

cells occurs when the logic value of the victim cell is inverted after a transition due to a write operation that is performed on the aggressor cell, which was identified in STT-MRAM [41].

IV. EXISTING TEST SOLUTIONS FOR MRAMs

Due to their regular structures, there exists a typical test development methodology for SRAMs [59]. The three main steps of this methodology are defect analysis, fault modeling, and development of test (e.g., March) algorithms. Defect analysis is usually done by using information collected after physical failure analysis (PFA) of defective memories. Note that this analysis can also be done by using a physical model of the memory and by subsequently performing defect injection campaigns. Fault modeling consists in finding an appropriate FFM for each type of defect encountered during defect analysis. Finally, the March test algorithms are developed to cover all possible FFMs that can be found in a given memory technology. The main goal here is to have the lowest complexity for the March algorithms with the highest FFM coverage. Note that this approach has also been used for other types of memories, such as DRAM, Flash, RRAM, or PCM [64]–[66].

In the context of MRAM, the same methodology can apply. This has been done for Toggle-MRAMs that are already in mass production, and the embedded STT-MRAM will soon enter into mass production [67]–[69]. Details will be given in Sections IV-A–IV-C. Concerning TAS-MRAM, the same methodology has been used to develop test solutions, despite the fact that TAS-MRAM has never (and will never) been into mass production. To this purpose, fault injection campaigns instead of PFA information have been used during defect analysis. Details will be given in Section IV-B. However, an interesting feature in the development of test solutions for these memories is that, very often, the same FFMs have been identified, even if defects or misbehaviors at the origin of the observed errors may differ depending on the MRAM technology. Consequently, the March test algorithms developed for a given technology, including TAS-MRAM, are likely to be utilized to test other memory technologies.

A. Test Solutions for Toggle-MRAM

The only study about Toggle-MRAM fault modeling and testing was presented in [62], [70], and [71]. In these articles, the authors first performed a classification and analysis of defects and their behavior and proposed corresponding fault models [62]. Defect injections were done by considering

- 1) Resistive shorts, such as wordlines and bitlines, are shorted to either V_{dd} or V_{ss} .
- 2) Stuck-on and stuck-off on the read access transistor of the MTJ cell.
- 3) Line break (open) on the write wordlines, read wordlines, or bitlines of the memory array.

- 4) Coupling on the same layer, that is, between the write wordlines, between the read wordlines, and between the bitlines.

- 5) MTJ device defects, such as MTJ opens, MTJ tunneling defects, and rough junction defects.

Defect injections were done on a real layout considering the physical model of a toggle MTJ cell, with parameters gathered from an experimental process. Simulation results in terms of the correlation between defects and fault models showed that most of the defects can be modeled by stuck-at faults, except MVF and KF for which two new FFMs were proposed, as mentioned in Section III-C. The proposed test solution for these FFMs consisted of a March C- algorithm applied with a 100% write current to detect stuck-at and MVF faults and the same algorithm applied with a 90% write current for KF detection.

In their subsequent study, the authors presented chip measurement results to prove the existence of write disturbance faults [70], [71]. They proposed the WDF model for toggle MRAM to represent the behavior of faults that affect data stored in MRAM cells when an excessive magnetic field is applied during the write operation on the neighboring cells. Then, they proposed a SPICE macro-model for the MTJ cell of the toggle MRAM to carry out circuit simulations. Finally, they developed an MRAM fault simulator, called RAMSES-M, able to derive a March $17N$ test algorithm (N being the number of cells in the memory array), which is described in the following:

$$\begin{aligned} &\uparrow(w0); \uparrow(r0, w1, r1); \downarrow(r1); \uparrow(r1, w0, r0); \downarrow(r0); \\ &\downarrow(r0, w1, r1); \downarrow(r1); \downarrow(r1, w0, r0); \downarrow(r0) \end{aligned}$$

where \uparrow (resp. \downarrow) denotes an increasing (resp. decreasing) addressing order of the various cells in the memory array for the corresponding march element composed of one or more read/write operations, $\uparrow\downarrow$ denotes an irrelevant address order, $w0$ ($w1$) denotes a write 0 (write 1) in a cell, and $r0$ ($r1$) denotes reading a cell with expected value 0 (value 1). A March element (nine in the above algorithm) is a sequence of operations applied to each cell in the memory before proceeding to the next cell.

The algorithm was proposed not only to test WDF in addition to SAF, TE, coupling, and address decoder faults but also to distinguish WDF from other faults, especially SAF0. The test was shown to be more efficient compared with the conventional March C algorithm. Since other emerging nonvolatile memory technologies, such as RRAM and PCM, may suffer from SAF or WDF as well, similar testing solutions can also be applied effectively [36], [72].

B. Test Solutions for TAS-MRAM

The only study about TAS-MRAM fault modeling and testing was presented in [23], [63], [73]–[77]. A preliminary step in this study was to develop a TAS-MTJ model in order to, subsequently, be able to perform electrical simulations [78]. Magnetization dynamics, as well

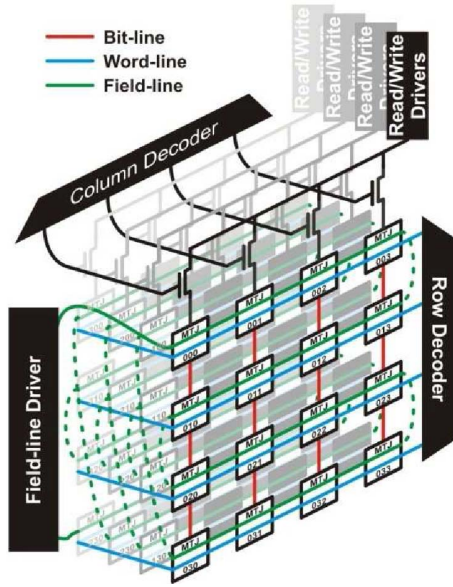


Fig. 6. TAS-MRAM architecture [73]–[77].

as dependencies of tunneling conductance, were considered during the development of this model, which was validated and calibrated with silicon data provided by Crocus Technologies. During this step, heat diffusion in such devices was also studied as TAS-MTJ relies on the blocking temperature concept.

The next step was to analyze the failure mechanisms of TAS-MRAMs through defect injections (resistive-open, resistive-bridge, and coupling-based resistive-open defects) in a representative TAS-MRAM architecture. Fig. 6 depicts a typical TAS-MRAM architecture, organized in a square matrix with 2^{MR} rows and 2^{NC} columns, for a total storage capacity of 2^{MR+NC} bits per page, where MR and NC are the number of bits used to specify the row and column address, respectively. Each cell in the array is connected to one of the row-lines (namely, wordlines) and connected to one of the column-lines (namely, bitlines). A particular set of MTJs can be accessed for a read or write operation by selecting its wordline and bitline. There is only one field line that connects all MTJs serially: 1) row by row and 2) passing through all pages in the architecture. Defect injections were made taking into account both magnetic and CMOS fabrication processes, as well as architecture properties, such as wires and cell neighboring. Not surprisingly, the analysis of failure mechanisms showed that both read and write operations may be affected by the injected defects.

Fig. 7 illustrates how and where resistive-open defects on various interconnects of the memory array were injected. The defects are those that directly impact MTJ's heat current (from df0 to df2), those that indirectly impact the MTJ's heat current (from df3 to df6), and those that impact field-line current (df7 and df7'). The TAS-MRAM operation may be affected by these

resistive-open defects in several ways, and all details can be found in [75]. Similar injection campaigns were carried out for resistive-bridge and coupling defects [63], [76].

From the results of the fault injection campaigns, the following step consisted in performing fault modeling. Standard static fault models, such as SAF and TF models, were identified. Moreover, a WDF model related to the write procedure that requires heating the device above its blocking temperature was also identified. In addition, a less restrictive static CF was also observed in which, irrespective of the stored data in the aggressor cell, the victim cell fails to perform a “write 0” operation. Moreover, both dynamic transition coupling and dynamic incorrect read CFs were observed. In those cases, the aggressor cell should be accessed immediately before the victim cell. Finally, a dynamic write disturb CF was also observed. However, in this case, the aggressor cell should be accessed immediately after the victim cell.

Based on the observed failure mechanisms, a full set of FFMs specific to TAS-MRAM was identified, and a March 14N test algorithm targeting these specific FFMs was proposed. It is reported in the following:

$$\begin{array}{ccccccc} \uparrow (w0); \uparrow (r0, w1); \uparrow (w1, r1, w0, r0, w1); \uparrow (w0, r0); & & & & & & \\ M1 & M2 & M3 & M4 & & & \\ & & \uparrow (w1, r1); \downarrow (r1, w1) & & & & \\ & & M5 & M6 & & & \end{array}$$

where M1 is the initialization march element used to reset all cells to 0, and M2–M6 are the various march elements of the proposed algorithm used to test all the target FFMs. In this algorithm, each static FFM is covered as follows.

- 1) *SAF1*: If $r0$ in M2 returns “data 1.”
- 2) *WDF1*: $w1$, $w1$, and $r1$ sequence (M3 with the help of $w1$ in M2) when the second $w1$ toggles the cell to “data 0,” which is observed by the read operation.

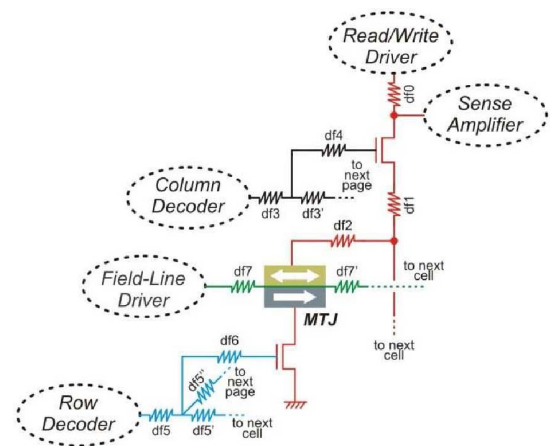


Fig. 7. Resistive-open defects injection [75].

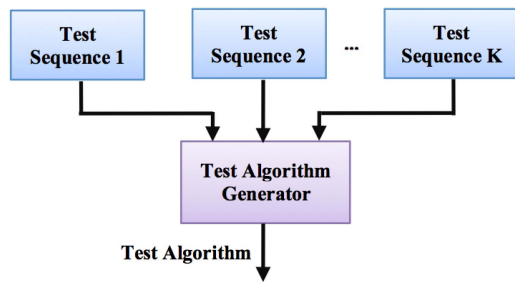


Fig. 8. Test algorithm generation flow [61], [79].

- 3) *SAFO*: If r_1 in M3 returns “data 0.”
- 4) *TFO* and *CFtr0*: w_0 and r_0 sequence in M3 when the read operation returns “data 1.”

Regarding dynamic FFM, they can be covered in a similar manner (details can be found in [77]). Finally, faults in the address decoder are covered by M1, M2, and M6.

C. Test Solutions for STT-MRAM

Compared with Toggle-MRAM and TAS-MRAM, a higher number of test solutions for STT-MRAM can be found in the literature. These solutions can be classified as: 1) test generation methods; 2) BIST techniques; and 3) Design-for-Testability (DfT) techniques. They are briefly described in the following.

1) *Test Generation Methods*: A test generation method has been proposed in [61] and [79] to cover all faults that are specific to STT-MRAM. As a preliminary step, the impact of process variations and test conditions, namely the temperature and voltage, on the defect manifestation were fully quantified. The spot defects in the layout and their manifestation as resistive opens and shorts in the netlist, as well as the impact of MTJ defects on the functionality of memory arrays, have been analyzed. This analysis showed the existence of dynamic read fault behavior requiring multiple test vectors for fault excitation and detection. This type of fault happens in the case of intercell CFs. In addition, write faults have been shown to be very sensitive to the test voltage and temperature, and a low-voltage and low-temperature condition is representing the worst case scenario in their framework.

Based on the results of this analysis, and after appropriate fault modeling, the authors constructed an efficient test algorithm that provides full coverage of the observed faults. They built this algorithm based on the set of test sequences identified during the fault modeling process. To this end, several March test sequences were constructed to detect different classes of STT-MRAM specific faults. Finally, a combined test algorithm for testing all the modeled faults was developed (see Fig. 8).

2) *BIST Techniques*: BIST techniques presented in [80]–[82] were developed to perform *in situ*, statistical, retention failure testing of large STT-MRAM arrays.

A retention failure is the consequence of a bit-flip in a cell, a stochastic phenomenon, caused by thermal noise. Since the retention time is exponentially proportional to the thermal stability, testing of retention failures consists in measuring the thermal stability of a cell by applying a weak write current to that cell. Naeimi *et al.* [40] have proposed a retention test method with weak write current based on the thermal activation model proposed in [83]. Since the thermal activation is a stochastic phenomenon, a lot of successive tests are required to acquire statistically significant data. Thus, retention test time for cells with high thermal stability increases exponentially. Although performing a parallel test at a subarray level can reduce retention test time, the test time is still a major bottleneck for this method.

In order to alleviate the retention test time problem, Yoon *et al.* [80], Yoon and Raychowdhury [81], and Hamdioui *et al.* [82] proposed a new MBIST architecture that performs retention testing of large STT-MRAM arrays in a time-efficient manner. The proposed MBIST scheme reduces retention time considerably by: 1) applying weak write current to multiple rows in an array and 2) conducting a read operation only when a fault is detected within the rows under test.

The retention test is divided into two phases: error detection (ED) and error search (ES). In the MBIST architecture, the corresponding logic for ED and ES is included in the control logic. Based on the outputs of the MBIST circuit, an ED (err_det) signal is asserted by ED logic, and while this signal is asserted, the ES phase is performed. ES logic controls which rows to assert to localize the error and it outputs error location to the output of control logic once it identifies error locations. A $search_done$ signal is asserted if ES is finished and it resets err_det signal. An IWW (weak write current) bus controls the voltage of bitline and wordline, leading to different magnitudes of IWW current. Columns with different resistors serve as references to find errors in blocks of rows, and temperature sensors are placed inside a subarray to monitor the temperature. Each characterization test, which determines the thermal stability, is qualified by temperature. The proposed scheme enables parallelism in the test process and allows a fine tradeoff between the localization of weak cells and test time. More explanations about the ED and ES processes can be found in [80] and [81]. The proposed MBIST shows a 93.75% improvement in retention test time compared with the brute-force approach in [40] with less than 5% estimation error.

3) *DfT Techniques*: The DfT technique presented in [84] targets RDF detection. Read disturb is a major reliability issue in which a read operation on a given cell can lead to a bit-flip because read and write currents share the same path. As read disturb has a dependence on various important design parameters, such as write current, read current, retention, and readability, a reduction in the read disturb rate always leads to design compromises. Moreover, with

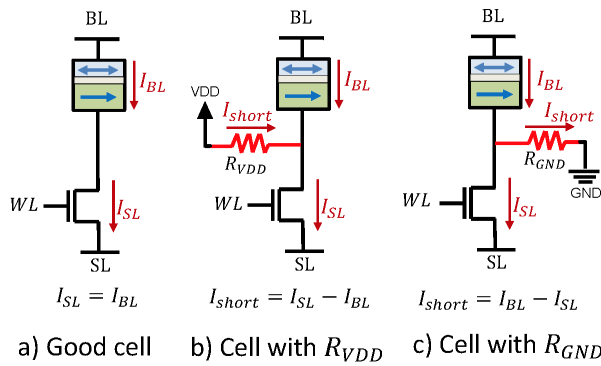


Fig. 9. STT-MRAM cell with different short defects [85], [86].
(a) Good cell. (b) Cell with R_{VDD} . (c) Cell with R_{GND} .

a reduction technique, it is not possible to eliminate read disturb entirely. Hence, it must be detected to attain a reliable memory. Therefore, the authors proposed a dynamic circuit-level approach that tracks the read current and, thus, is able to detect read disturb. This is possible as an RDF changes the resistance of the affected bit-cell, which, in turn, affects the read current. As a consequence, the ratio of the actual read current to the reference current of the sense amplifier will flip. This observation is exploited by the proposed detection circuitry to create an error signal, which indicates the occurrence of a read disturb. Since the read current is unidirectional, read disturb can only affect one logic value. Therefore, the read disturb detection (RDD) circuit is only activated for that particular logic value, which results in a very low power penalty. Moreover, there is no timing penalty as the read detection circuit is isolated from the actual read process by using a current mirror.

The whole RDD circuit consists of five parts: a basic equalizer circuit, a sense amplifier used to read the bit-cell content, the detection circuit itself, a control logic used to enable the detection circuit only for read operations that can be affected by a read disturb, and a self-test mechanism to test the functionality of the RDD circuit. More details about the implementation and operational modes of this RDD circuit can be found in [84].

Experimental results show that the proposed RDD technique can detect up to 95% of the total RDFs and imposes negligible area and power overhead.

Another DfT technique has been proposed in [85] and [86] to test bridging defects in STT-MRAMs. The authors have modeled resistive-short defects between the internal node of an STT-MRAM cell and an external node. These types of defects are modeled by a resistive connection between the internal node of the cell and either V_{dd} (R_{Vdd}) or Gnd (R_{Gnd}). This test method is based on the fact that a resistive-short defect between an external node and the internal node of the memory cell changes the amount of the current flowing into the cell (through the bitline terminal) and outside of the cell (through the source-line terminal). Fig. 9 shows the behavior of a cell

with different short defects. A fault-free cell behaves as a single current path, in which the current from the read circuitry at the bitline terminal flows through the MTJ and the access transistor until the source-line terminal. In the case of R_{Vdd} short, the current is injected to the cell so that I_{SL} becomes greater than I_{BL} . Similarly, in the case of R_{Gnd} short, its current is removed from the cell, and I_{BL} becomes greater than I_{SL} . A faulty read operation happens when the voltage generated for reading an antiparallel state ("1") becomes lower than the reference voltage or when the voltage generated for reading a parallel state ("0") becomes larger than the reference voltage.

Based on this observation, the authors proposed to change the basic columnwise readout circuitry so that it can measure the difference between the current flowing into and out of any cell in the respective column. A large current difference signals the presence of a defect.

The modified read circuitry exploits differential current amplifiers that are placed between the reference part and the memory cell part of the read circuitry. The modified circuit has three modes of operation: normal mode, Test mode 1 to detect an R_{Vdd} short, and Test mode 2 to detect an R_{Gnd} short. This method is robust to process variations and can detect resistive open and short defects. More details can be found in [85] and [86].

4) *Test Solutions for SOT-MRAM:* To the best of our knowledge, there is no solution published so far for testing SOT-MRAM.

V. EXISTING RELIABILITY IMPROVEMENT SOLUTIONS FOR MRAMs

To enhance the reliability of MRAM, many innovative and effective solutions have been proposed from the device level, circuit level, and architectural level. In the following, we first summarize the related work that has been done in terms of reliability modeling and evaluation. The second part of this section describes existing reliability solutions for STT-MRAM, which is the mainstream MRAM technology, and the final part will be dedicated to reliability solutions for more advanced technology, for example, SOT-MRAM, racetrack memory, and skyrmion. Note that although the following reliability enhancement solutions were proposed for MRAMs, similar solutions can be applied or extended to other emerging nonvolatile memories (e.g., RRAM and PCM) as well.

A. Reliability Modeling and Evaluation Methodologies of MRAM

1) *PVT Modeling and Evaluation Techniques:* An accurate and efficient reliability modeling and evaluation method is essential to pinpoint the bottleneck and the most fatigue part in the system. Especially, the characteristics of FM materials, such as TMR, are very sensitive to environmental temperature, which has already been observed in many experiments, such as in [87]. Wu et al. [88]

proposed a thermal model of MTJ validated by published experimental measurements and found a significant read disturbance in the deep submicrometer regime. A body-biased feedback sensing amplifier was also proposed to improve read reliability at high temperatures. Similarly, Zhang *et al.* [89] quantitatively investigated read/write errors of a single STT-MRAM cell caused by process variations and temperature with commercial EDA tools on a 45-nm technology node. Based on this work, a thermal-aware sensing circuit was proposed to reduce the read errors due to PVT variations [90]. Considering the heating mechanism of PCM, the thermal disturbance should also be mitigated [2]. Indeed, with technology scaling and shrinking of the distance between cells, the thermal disturbance will become more severe for PCM. Some reliability solutions mentioned in this article may be extended to PCM to alleviate this problem.

Xie *et al.* [91] claimed that the stochastic switching of the MTJ under the STT effect is severely affected by the thermal noise under the room temperature. They proposed the numerical Fokker–Planck-based simulation framework to study the thermal effect in MTJ switching of STT-MRAM. A comparison with other simulation methods was also given in this article. In addition, thermal variation and Joule heating on the reliability of STT-MRAM were explored [92]. Extensive simulations revealed the close relationship of STT-MRAM read/write failures and thermal fluctuations. Kong *et al.* [93] investigated the impact of process variations on the MTJ with the Object-Oriented MicroMagnetic Framework (OOMMF) and SpinFlow3D and showed that geometrical modifications can greatly improve the MTJ performance. An STT-MRAM cache reliability evaluation framework was built in [94] not only considering the correlation of retention failure, read disturbance, and write failure but also taking into account the diversity of running benchmarks and process variations among different memory cells. Their experimental results indicated that the error rate may vary up to $32\times$ among different benchmarks and process variations of memory cells contribute another $6.5\times$ difference in cache vulnerability.

Moreover, reliability is especially challenging in some applications, such as automotive, military, and aerospace, which may suffer from extreme temperature conditions. The self-heating effect of MTJ stack has been observed in [95] and investigated through 1-D numerical thermal simulations. Unlike the TAS approach in which the MTJ is heated by an external element, here, the MTJ can be heated by itself due to the Joule heating. Despite the efforts dedicated to technology optimization in the past years, a high current density flowing through the MTJ is always demanded by most of the switching mechanisms. This leads to a significant self-heating effect that may cause functional errors in hybrid MTJ/CMOS circuits [96].

2) *Dielectric Breakdown of STT-MRAM:* Dielectric breakdown is another critical reliability issue that determines the lifetime of devices (transistors or MTJs). As MTJ

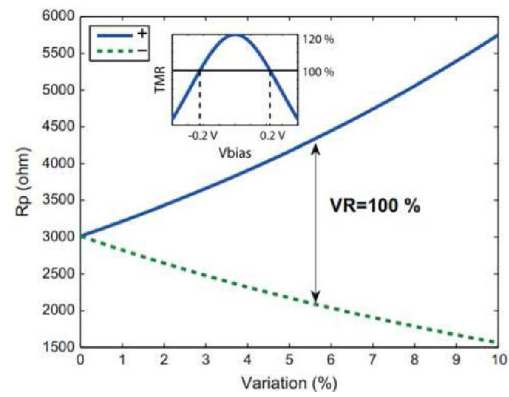


Fig. 10. Relationship of: 1) resistance of MTJ and oxide barrier thickness variation and 2) TMR and bias voltage for reading (the embedded subplot) [37].

is a memristive device and its resistance mainly comes from the oxide barrier, the voltage applied on MTJ is almost imposed on the insulator (MgO). A dielectric breakdown has attracted significant attention for other resistive memories, such as RRAM and PCM [97]–[99]. With the shrinking of technology nodes and oxide dielectric thickness (~ 1 nm), the breakdown voltage also scales down, and it is necessary to mitigate time-dependent dielectric breakdown (TDDB) of the MTJ caused by write operations [100].

Several experiments have analyzed the TDDB effect in STT-MRAM [101], [102], and other experiments have been performed to explore the physical mechanism behind the TDDB phenomenon in STT-MRAM [103]. It has been found that TDDB depends on a variety of factors, such as annealing temperature, oxide material purity, tunnel barrier thickness, stress voltage, temperature, and stress duration. Variability of oxide barrier thickness also leads to reliability degradation, as the resistance has an exponential dependence on the thickness. Furthermore, a bias voltage for reading can greatly reduce the TMR [37], as shown in Fig. 10. Ho *et al.* [104] investigated the reliability issue caused by dielectric breakdown or TDDB effect and built an accurate TDDB model to analyze the time to breakdown and the postbreakdown currents. Simulation results revealed that new design constraints need to be imposed for better reliability of STT-MRAM. Munira *et al.* [52] analyzed the factors affecting the reliability of the writing process in STT-MRAM array, including process variations, thermally activated initial angle, thermal fluctuations, and the voltage across the MTJ. A quasi-analytical model was built to calculate the current and energy during the write operation for TDDB evaluations.

3) *Radiation Effect on STT-MRAM:* With the emergence of MRAM products, some new reliability issues have been observed, such as radiation effect. Hirose *et al.* [105] estimated the risk of irradiative particle bombardments, that is, alpha particles and neutrons, which are the well-known

soft error sources on the ground, with respect to both frequencies and the hazardous effects of bombardments. The effect of proton and Cr ion radiation on MTJs was investigated in [106], which concluded that an in-plane MTJ is robust to proton irradiation but the properties of an MTJ can be degraded by Cr ion irradiation.

4) *Reliability Measurements on Prototypes*: Some reliability modeling studies with prototype measurements were also reported. An error behavior model used to characterize read/write errors of STT-MRAM was proposed in [107]. The proposed model, which was validated by measurements on Everspin MRAM chips, revealed that, in a normal environment, the error rate is very low and dominated by read error, but the write error rate dramatically increases with the magnetic disturbance. Their research highlighted the necessity of protecting STT-MRAM from magnetic disturbance or attack. An 8-Mb embedded STT-MRAM prototype was tested under package-level reliability stress, magnetic stress, and radiation stress [108]. The measurements showed that their prototype had a negligible fail bit count (FBC) even without ECC protection and was suitable for mass production. These studies indicate that the STT-MRAM fabrication process is approaching commercial maturity.

B. Reliability Solutions From the Device and Fabrication Process Perspectives

There are many works trying to improve the MRAM reliability from device fabrication optimization and material engineering perspectives. Mahawar *et al.* [109] observed that the conventional fabrication process of MRAM is based on bulk or SOI technology, which may not be suitable for MRAM that requires a large switching current. They improved the fabrication process by introducing the fully depleted (FD) silicon carbide (4H-SiC) substrate NMOS technology to increase the driving current effectively. The experimental results showed that the new process has a very low probability of thermal fatigue and device failure and can reduce the write error rate by 45%. Plasma oxidation was proved to be an effective method to produce Al-metal-based MTJ with ultrauniform resistance [110]. Interface engineering is another efficient way to improve the write and read reliability [111]. The performance of MTJs can be significantly enhanced through proper modulation of heavy metal/FM metal interface, such as perpendicular magnetic anisotropy, TMR, and magnetic damping.

Gonçalves *et al.* [110] established a simulation framework based on the Landau–Lifshitz–Gilbert (LLG) equations that can solve the magnetic dynamics self-consistently. With this framework, the authors explored different magnetic materials to construct the MTJ stack and illustrated the requirement for the coupled free-layer MTJ stacks in scaled technology nodes. In addition, Augustine *et al.* [112] presented a design space exploration framework for STT-MRAM, in which they performed

Table 4 Comparison Results for BCH and Hamming Coding Schemes With an Encoded Number of Bits of 128 [113]

	Hamming (Correction-bit = 1)	BCH (Correction-bit = 2)
Latency (ns)	1.3	3.6
Area (μm^2)	4907	106700
Dynamic power (mW)	0.11	1.10
Leakage (mW)	0.12	2.24

a numerical study on four types of MTJ stacks and evaluated their advantages, as well as the limitations from the perspective of memory applications.

C. Reliability Solutions From the Coding Theory Perspective

Error correction code (ECC) is widely used in SRAM and DRAM to enhance the access reliability, and among many, Hamming and Bose–Chaudhuri–Hocquenghem (BCH) are the most popular coding theories. Hamming codes are to detect two bits (with the help of an additional parity bit) and correct only a single bit in the code-word, whereas BCH is typically employed when multiple bit correction capabilities are required. These coding schemes require an additional encoding/decoding mechanism, and their synthesized results for the same are shown in Table 4. Moreover, the impact on read and write latencies of L1 and L2 MRAM-based caches due to various coding schemes is illustrated in Table 5. The unidirectional read disturbance switching and asymmetrical $0 \rightarrow 1$ and $1 \rightarrow 0$ switching in the write operation provide opportunities for ECC code optimization in STT-MRAM. Mei *et al.* [115] proposed a polar code to reduce the error rate caused by process variation and thermal fluctuation. Compared with Hamming, BCH, and LDPC codes, the proposed polar code can induce lower decoding complexity and approach flexible code rates to be adaptive to different raw error rates of STT-MRAM chips. Sayed *et al.* [116] also took advantage of unidirectional read disturbance switching and asymmetrical $0 \rightarrow 1$ and $1 \rightarrow 0$ switching and devised a unidirectional ED code instead of conventional ECC code to reduce the latency and storage overhead. With the proposed technique, both reliability and access performance of STT-MRAM can be improved dramatically.

Table 5 Read and Write Latencies of 64-Bit STT-MRAM L1 and L2 Caches With 16- and 512-KB Capacities [114]

			No ECC	ECC1 SECDED	ECC2 BCH	ECC3 BCH	ECC4 BCH
Storage overheads			0%	11%	18.9%	25.5%	31%
L1	Write [ns]	ECC Encoding	—	0.400	0.525	0.530	0.545
		Memory Write	11.610	6.456	4.909	4.100	3.628
		Overall	11.610	6.856	5.434	4.630	4.173
	Read [ns]	ECC Decoding	—	0.580	2.459	3.698	4.714
		Memory Read	0.898	0.899	0.899	0.899	0.900
		Overall	0.898	1.479	3.358	4.597	5.614
L2	Write [ns]	ECC Encoding	—	0.400	0.525	0.530	0.545
		Memory Write	21.00	12.50	9.50	7.80	6.70
		Overall	21.00	12.90	10.025	8.33	7.245
	Read [ns]	ECC Decoding	—	0.580	2.459	3.698	4.714
		Memory Read	1.120	1.121	1.121	1.121	1.122
		Overall	1.120	1.701	3.580	4.819	5.836

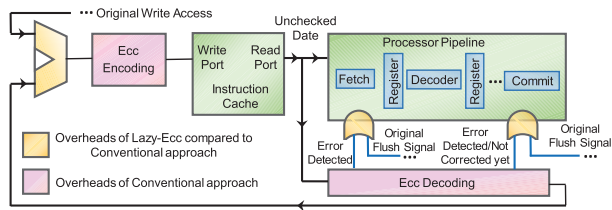


Fig. 11. Block diagram of Lazy-ECC for instruction cache [114].

A novel cascaded channel model for fast error rate simulation was proposed in [117]. Based on this model, the authors proposed a two-stage hybrid decoding scheme extended from the Hamming code to improve STT-MRAM access reliability. Kang *et al.* [118] proposed one-step majority-logic-decodable (OS-MLD) code to correct multi-bit error in STT-MRAM. The code requires low encoding/decoding latency and circuit complexity. The authors further optimized the decoder implementation to increase the parallelism. A hybrid MTJ/CMOS memory array design was used to validate the effectiveness of the proposed coding scheme. Aliagha *et al.* [119] took advantage of the unidirectional property of the read error and the asymmetrical switching between 0 and 1 states to devise an error-rate-aware coding scheme. With this coding scheme, the number of error-prone $0 \rightarrow 1$ transitions can be minimized. Simulation results showed that the read disturbance and write error can be reduced by 58%–71% when running different workloads with only less than 1% hardware overhead.

Sayed *et al.* [114], [120] combined the early termination of a write operation with ECC protection to enable STT-MRAM as an L1 cache. To reduce the ECC decoding latency, the authors separated the ECC detection from correction and performed speculative computing on unchecked data. The ED can ensure error confinement and data integrity. In addition, in the above studies, a fast and robust ECC scheme called Lazy-ECC was proposed to guarantee the reliability of STT-MRAM when used in the fast upper level cache. As shown in Fig. 11, it separates ED and correction and performs speculative computation on unchecked data. This enables a very effective unidirectional ED code for STT-MRAM, which exactly matches the asymmetric switching characteristic of this technology.

Furthermore, an online adaptive approach was proposed to eliminate process variations and temperature effects on the write latency and reliability. Besides, Das and Toubia [121] proposed an ECC scheme based on orthogonal Latin square (OLS) code that is used in SRAM to reduce both hard and soft errors in STT-MRAM. The soft errors can be corrected by the error-correcting OLS, while the hard errors can be masked during the decoding procedure. The one-step decoding has very low latency and a negligible impact on cache performance.

ECC is also effective and widely used in other emerging nonvolatile memory technologies. As for PCM, several ECC coding schemes were proposed, such as DIN coding [122],

flipping coding [123], and WOM coding [124], to enhance the endurance with little performance and hardware overheads. Wang *et al.* [125] investigated the configurable ECC in RRAM, and the proposed technique had the adaptive ECC correction capability depending on different error modes.

D. Reliability Solutions From the Circuit Design Perspective

At the circuit design level, many effective techniques have been proposed to improve STT-MRAM reliability. Chen *et al.* [126] classified the access errors as persistent errors and nonpersistent errors. The former category is mostly caused by process variations, while the latter one is due to the thermal fluctuation of the MTJ. Then, the authors proposed a stochastic circuit design methodology to consider the above errors and demonstrated that such methodology was essential for spintronic logic and memory design [126]. Meanwhile, Toshiba and The University of Tokyo proposed a write-verify-write strategy to improve write reliability in their 4-Mb STT-MRAM prototype [127]. Chen *et al.* [128] exploited STT-MRAM technology to build nonvolatile FPGA for radiation protection. By replacing CMOS LUT in FPGA with MRAM, both the static power consumption and radiation-induced errors can be reduced considerably. This is one of the first attempts to incorporate STT-MRAM in FPGA design. Zhang *et al.* [90] proposed a thermal compact model of the MTJ and investigated the thermal impact on the access transistor and the MTJ in a memory cell. The authors also proposed a thermal-aware sensing circuit design to form a feedback mechanism to compensate for the driving current loss due to the temperature increase. Experimental results showed that the new sensing circuit design can effectively reduce the error rate due to thermal fluctuations. On the other hand, the radiation effect on peripheral circuitry of STT-MRAM was considered in [129], and a comprehensive approach was proposed to analyze the soft errors due to radiation from the device-level modeling to circuit-level analysis. A lookup table was constructed to store the failure probability as a reference for STT-MRAM designers. A thorough error rate analyses on write and sensing circuits were also performed considering the process variations and radiation effect.

Considering that the read current is approaching the write current below 32-nm technology node, a read disturbance-free scheme was proposed in [130]. As mentioned in this work, a large and short-period read current can be applied without flipping the MTJ state due to the rapid increase in the critical current at a pulse duration of less than 10 ns. Meanwhile, Noguchi *et al.* [131] proposed a hierarchical bitline design and 2T2MTJ cell design to suppress the read current magnitude and duration. Measurements on the fabricated 1-Mb MRAM prototype showed that the proposed technique can reduce read disturbance by $10\times$. Kang *et al.* [132] proposed to utilize a multilevel

cell (MLC) structure in an adaptive STT-MRAM design that can work either in the high-reliability mode or high-density mode depending on the requirements of specific applications. Appropriate read/write sensing circuits and proper control circuits were devised to support such an adaptive design.

Some bit-cell flip detection circuits were proposed in [51] and [133], which can determine the completion of the write operation dynamically. These circuits actually trace the current during the write operations and generate a bitwise acknowledgment signal when they complete their respective write operations. With this approach, not only high-energy efficiency is achieved but also the overall write errors can be detected that can reduce the overhead of ED and correction. Moreover, the duration of the write current flow through the MTJ stack can be reduced using this approach, which can also reduce the TDDDB effects. A current boosting technique was proposed in [134] to mitigate the negative impact of the stochastic switching behavior by reducing the write current delay margin.

Bose and Der Jeij [135] first predicted the sensitivity of the application performance to write latency considering the write error rate. Then, based on the write frequency, the write current level was adjusted at run time to trade off the performance and write reliability. Sayed *et al.* [136] developed a cross-layer framework to analyze retention failures and proposed an adaptive scrubbing scheme that considers process variation and temperature to mitigate retention failures. This was achieved with a group of wise allocations of the scrubbing intervals in the memory array. Another PVT-aware operational transconductance amplifier (OTA) design technique was proposed in [130]. With the multivalued resistor obtained by serial-parallel connections of MTJs, the OTA's operating point can be tuned postprocess and be more immune to PVT variations. Zhang *et al.* [137] analyzed the impact of process variation on the reliability of the STT-MRAM cell and proposed a statistical design flow to reduce persistent and nonpersistent errors.

For other nonvolatile memory technologies, such as RRAM and PCM, some similar circuit design techniques can be used to improve their endurance and reliability. For example, García-Redondo *et al.* [138] proposed a reconfigurable writing technique for RRAM to mitigate the write errors due to thermal variations. The read and write circuits of RRAM were revised in [139] to detect and resolve the performance degradation caused by soft errors when RRAM is used in neuromorphic systems. Since this article focuses on the reliability solutions of MRAM, we do not delve into the reliable circuit design details of other emerging nonvolatile memories.

E. Reliability Solutions From the Computer Architecture Perspective

Cheshmikhani *et al.* [140] investigated the STT-MRAM cache reliability issue from the cache replacement policy perspective. The authors observed that the highest

temperature increase occurs during the energy-consuming write operation. By adjusting the replacement location, the proposed thermal-aware cache replacement can make sure consecutive write operations be separated far away in physical locations, so the temperature increase can be alleviated and the write error rate can be improved. Wu *et al.* [141] noticed the tradeoff between thermal reliability and MRAM write energy and proposed a thermal-aware NUCA architecture design for STT-MRAM-based last-level cache. Guo *et al.* [142] proposed a dynamic voltage adjustment technique called DOVA and applied the write voltage depending on the criticality of the cacheline to make a tradeoff between performance and reliability of STT-MRAM-based L1 cache. Wen *et al.* [143] focused on the reliability of MLC-based STT-MRAM cache. MLC can achieve high integration density. However, it has a soft bit and a hard bit that require different sensing schemes and have different sensing reliability. The authors first proposed a novel cell structure R2M, which is based on the 2T-2MTJ cell design to enhance the sensing reliability of the soft bit. Then, they proposed another scheme, R2M-C, to improve the reliability by exploiting data locality and redundancy. By combining the two-level optimizations, the reliability of the MLC STT-MRAM cache can be effectively improved.

Cheshmikhani *et al.* [144] observed that, although the parallel access of tag and data array in the cache can improve the cache performance, it may accumulate read disturbance errors in STT-MRAM cache due to the parallel reading of unused data. Based on this observation, the authors proposed the Read Error Accumulation Preventer cache (REAP-cache) to reduce the read disturbance. Architectural-level simulations showed that the read disturbance-induced MTTF can be improved by $171\times$ with less than 1% area overhead and 2.7% energy overhead. Ahari *et al.* [130] proposed a low-cost architectural solution to reduce the timing margin. They used a handshaking protocol between the memory and its controller to dynamically evaluate the write latency (see Fig. 12). Results showed that their approach can not only significantly reduce the write error rate but also improve the overall system performance. A process variation-aware STT-MRAM design framework was proposed in [46] and [47]. The authors utilized the hybrid analytical and Monte Carlo simulation-based approach to guide the STT-MRAM memory array optimization to make a tradeoff between performance, energy, and reliability. In another relevant work [145], a 3T-3MTJ cell structure was proposed to approach a tradeoff between storage density and access latency. To overcome the high write energy and latency of STT-MRAM, Sun *et al.* [146] proposed to add several small SRAM buffers between the core and L1 cache. These buffers can provide high data bandwidth and alleviate the performance degradation caused by the L1 STT-MRAM cache. The proposed architecture can reduce power consumption and improve radiation immunity effectively.

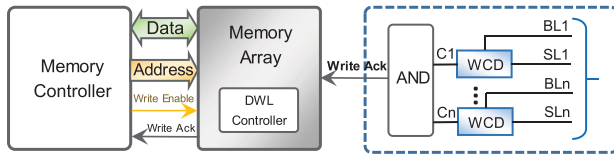


Fig. 12. Handshaking policy between the memory array and the memory controller [130].

Similar architectural reliability solutions may also be applied to other nonvolatile emerging memory, such as RRAM and PCM. For example, Swami and Mohanram [147] proposed a data compression and alignment technique to mitigate the write disturbance problem in PCM. Similarly, another work [148] proposed a synergistic approach that incorporated data compression, differential write, and wear-leveling to prolong the lifetime of PCM. As for RRAM devices, a dynamic writing driver architecture was proposed in [138], which can select different writing strategies depending on working temperature, data retention time, endurance requirement, and power consumption constraint. In addition, since RRAM is widely used in neuromorphic computing systems, many fault-tolerant techniques were proposed to make a tradeoff between neural network accuracy and lifetime of RRAM-based hardware platform [149]–[151].

F. Reliability Solutions for More Advanced Spintronic Memory Technologies

Although STT-MRAM has many advantages and attracts enormous attention, it induces high write energy and latency. The read disturbance problem also aggravates as the technology node shrinks down. Another way to deal with these challenges is SOT-MRAM technology. It utilizes the SOT (Rashba effect or SHE) to switch MTJ state. Due to the separation between read and write path, read disturbance can be eliminated. The write speed can also be faster than STT-MRAM.

Oboril *et al.* [152] exploited SRAM/SOT-MRAM to build a hybrid cache system, where SRAM was used as an L1 data cache and SOT-MRAM was used as L1 instruction cache and L2 cache. Extensive architectural simulations showed the benefits of using SOT-MRAM as an on-chip cache from the energy, performance, and reliability perspectives. Wang *et al.* [153], [154] noticed that the shortened write period may be comparable to the radiation-induced noise pulse, so the write circuit should also be protected from radiation in addition to the sensing amplifier. The simulation results on an SOT-MTJ compact model showed that the proposed write circuitry can effectively improve the radiation immunity.

Besides, to overcome the endurance bottleneck of conventional STT-MRAM, a NAND-SPIN architecture was proposed in [155], as shown in Fig. 13, where the fast erasing and programming of MTJ were implemented with two

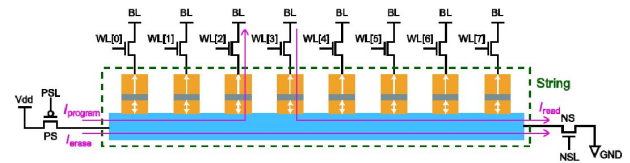


Fig. 13. Structures and operation of the nand-SPIN string [153].

unidirectional currents generating SOT and STT, respectively. By sharing the SOT-induced erase operation, this new memory can have a higher density and lower write energy. More importantly, the STT current for initializing MTJs from P state to AP state, which is larger than that of the inverse case, can be replaced by SOT current. Consequently, the tunneling current can be reduced to obtain a longer endurance.

Moreover, to deal with the write error problem of STT-MRAM, a promising method is the toggle spin torques (TSTs) MRAM whose write operation is performed by applying two currents to the MTJ and a heavy metal layer, respectively, in a toggle-like manner [156]. Neither of them can separately achieve magnetic switching, significantly improving the writing reliability. In order to improve endurance, several SOT-based designs that can be adjusted based on the switching activity rate were proposed [157], [158]. In these designs, the write current can flow through these devices only if the value to be stored is different from the value that is already stored. Bishnoi *et al.* [159] have proposed a multiport memory architecture using SOT-MRAM in which simultaneous read and write operations can be performed on the same cell without affecting each other's functionality. When employing a similar architecture in GPU (as a register file), a higher energy efficiency than SRAM and STT-RAM can be achieved [160] while maintaining the same performance as that of SRAM. The same architecture can also be used to verify write operations dynamically without any delay penalty; thus, write error rate can be improved.

VI. CONCLUSION AND PERSPECTIVES

In this article, we have presented a survey of existing test and reliability improvement solutions for MRAM memories. After a short discussion about the motivations for using such type of memory and their numerous fields of application, we have presented and discussed the various types of MRAM technologies existing today. Although STT-MRAMs have been widely investigated in the last few years and are now entering into high-volume mass production, some other promising technologies, such as SOT or TST, have emerged recently. In Section III, we discussed the main defectiveness and reliability issues of MRAMs, as well as the existing related FFMs used for test solutions development. Sections IV and V represent the core of this article and propose a summarized description of the main test and reliability improvement techniques proposed so

far in the literature. Although these solutions are mainly used for MRAM, some techniques mentioned in this article may also be extended or revised to deal with testing and reliability issues of other emerging resistive memory technologies, such as RRAM and PCM.

For an emerging technology making a debut in the customer side and get out of the foundry doors, it is extremely important to have very low defect per million (DPM) and high quality and reliability. For MRAM having new sets of fabrication materials, processes, masks, and so on, as well as new and fundamental operation modes compared with conventional CMOS-based memories, the yield in the ramp-up phase would be very low. This puts more pressure on the role of testing in order to have very high coverage in order to guarantee an acceptable level of DPM. Given the unique failure modes and mechanisms of MRAM and the resulting new faults and failures, it is important to develop MRAM-specific fault modeling, ATPG, DfT, and BIST schemes.

At the same time to boost the manufacturing yield of MRAM chips, it is imperative to devise new defect-tolerant schemes that are based on a combination of redundant rows and columns, error-correcting codes, and postfabrication trimming.

Similarly, MRAM is subject to new runtime failure mechanisms. The switching and write operations in MRAM is inherently stochastic. The read margin due to small differences of parallel and antiparallel resistance values, aggravated by process and runtime variations, is also very low. These runtime failure mechanisms mandate new fault-tolerant schemes for MRAM to ensure reliable operation in the field. Existing single error correction double ED (SECDED) schemes used in conventional CMOS memories may not be enough for MRAM, and more robust ECC schemes with multiple bit error correction may be required

for MRAM. Moreover, efficient implementations of such schemes to hide performance, power, and area overheads are of great importance.

Artificial intelligence (AI) and machine learning (ML) applications are one of the major target applications and markets of emerging MRAM technologies. In particular, the MRAM technology, due to its resistive and nonvolatility features, is a promising candidate for the emerging computing paradigms, such as computation-in-memory and neuromorphic computing, where computational tasks are actually performed within the memory itself, for better support of AI and ML applications. The regular structure of neural networks allows them to be mapped to a resistive crossbar architecture and, consequently, profit from its characteristic benefits. Resistive crossbars are widely investigated to hold the synaptic weight of the network [161]–[165]. MTJs are a particularly well-suited resistive memory technology for weight storage, as they offer a high integration density, low access delay, and a low power consumption [166]. The resistive states of the MTJ can be used to implement binary encoding of the weights of the artificial neural network. However, the use of MRAM technology in the implementation of these emerging computing paradigms comes with unique test and reliability challenges that mandate a new set of techniques for defect characterization, fault modeling, and test generation [167], [168] to fault-tolerant neuromorphic computing based on MRAM technology [169], [170]. While, on the one hand, the manifestation of defects is higher due to reduced margin, on the other hand, these architectures have inherent defect and fault tolerance due to approximate computing. Such tradeoffs for test and reliability solutions need to be explored in future research directions on these topics. ■

REFERENCES

- [1] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 453, p. 80, 2008.
- [2] H. S. P. Wong et al., "Phase change memory," *Proc. IEEE*, vol. 98, no. 12, pp. 2201–2227, 2010.
- [3] T. Coughlin. (Aug. 2019). *A Universal Memory?* [Online]. Available: <https://www.forbes.com/sites/tomcoughlin/2019/08/12/a-universal-memory/#228ac38673d8>
- [4] M. Lapedus. (May 2019). *Challenges in Making and Testing STT-MRAM*. [Online]. Available: <https://semiengineering.com/challenges-in-making-and-testing-mram>
- [5] G.-H. Koh, "Challenges and prospects of memory scaling," in *Short Course of Proc. Symposia VLSI Technol. Circuits, Samsung Electron., Virtual Conf.*, Jun. 2020.
- [6] T.-C. Chang, K.-C. Chang, T.-M. Tsai, T.-J. Chu, and S. M. Sze, "Resistance random access memory," *Mater. Today*, vol. 19, no. 5, pp. 254–264, 2016.
- [7] F. Zahoor, T. Z. Azni Zulkifli, and F. A. Khanday, "Resistive random access memory (RRAM): An overview of materials, switching mechanism, performance, multilevel cell (mlc) storage, modeling, and applications," *Nanos. Res. Lett.*, vol. 15, no. 1, p. 90, Dec. 2020.
- [8] E. Y. Deng, G. Prenat, L. Anghel, and W. S. Zhao, "Non-volatile magnetic decoder based on MTJs," *Electron. Lett.*, vol. 52, no. 21, pp. 1774–1776, Oct. 2016.
- [9] S. Jain, A. Ranjan, K. Roy, and A. Raghunathan, "Computing in memory with spin-transfer torque magnetic RAM," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 3, pp. 470–483, Mar. 2018.
- [10] D. Fan, S. Angizi, and Z. He, "In-memory computing with spintronic devices," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI*, Bochum, Germany, Jul. 2017, pp. 683–688.
- [11] Z. He, Y. Zhang, S. Angizi, B. Gong, and D. Fan, "Exploring a SOT-MRAM based in-memory computing for data processing," *IEEE Trans. Multi-Scale Comput. Syst.*, vol. 4, no. 4, pp. 676–685, Oct. 2018.
- [12] A. F. Vincent et al., "Spin-transfer torque magnetic memory as a stochastic memristive synapse for neuromorphic systems," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 2, pp. 166–174, Apr. 2015.
- [13] A. Sengupta, M. Parsa, B. Han, and K. Roy, "Probabilistic deep spiking neural systems enabled by magnetic tunnel junction," *IEEE Trans. Electron Devices*, vol. 63, no. 7, pp. 2963–2970, Jul. 2016.
- [14] G. Burr et al., "Neuromorphic computing using non-volatile memory," *Adv. Phys. X*, vol. 2, no. 1, pp. 89–124, May 2017.
- [15] K. Yang et al., "A 28 nm integrated true random number generator harvesting entropy from MRAM," in *Proc. IEEE Symp. VLSI Circuits*, Hong Kong, Jun. 2018, pp. 171–172.
- [16] E. I. Vatajelu, G. Di Natale, and P. Prinetti, "Security primitives (PUF and TRNG) with STT-MRAM," in *Proc. IEEE 34th VLSI Test Symp.*, Las Vegas, NV, USA, Apr. 2016, pp. 1–4.
- [17] I. L. Prejbeanu, S. Bandiera, J. Alvarez-Héault, R. C. Sousa, B. Dieny, and J.-P. Nozières, "Thermally assisted MRAMs: Ultimate scalability and logic functionalities," *J. Phys. D, Appl. Phys.*, vol. 46, no. 7, Feb. 2013, Art. no. 074002.
- [18] D. D. Tang and Y. J. Lee, *Magnetic Memory—Fundamentals and Technology*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
- [19] D. Apalkov, B. Dieny, and J. M. Slaughter, "Magnetoresistive random access memory," *Proc. IEEE*, vol. 104, no. 10, pp. 1796–1830, Oct. 2016.
- [20] D. Apalkov, B. Dieny, and J. M. Slaughter, "Magnetoresistive random access memory," *Proc. IEEE*, vol. 104, no. 10, pp. 1796–1830, Oct. 2016.
- [21] I. L. Prejbeanu, "Development of thermally assisted MRAMs: From basic concepts to industrialization," Ph.D. dissertation, Université Grenoble Alpes, France, 2015.
- [22] I. L. Prejbeanu, B. Dieny, and J. M. Slaughter, "Magnetoresistive random access memory," *IEEE Trans. Magn.*, vol. 40, no. 4, pp. 2625–2627, Apr. 2004.
- [23] J. Azevedo et al., "Dynamic compact model of

- self-referenced magnetic tunnel junction,” *IEEE Trans. Electron Devices*, vol. 61, no. 11, pp. 3877–3882, Nov. 2014.
- [24] M. Wang et al., “Current-induced magnetization switching in atom-thick tungsten engineered perpendicular magnetic tunnel junctions with large tunnel magnetoresistance,” *Nature Commun.*, vol. 9, no. 1, pp. 1–7, Feb. 2018.
- [25] N. Strelkov et al., “Impact of joule heating on the stability phase diagrams of perpendicular magnetic tunnel junctions,” *Phys. Rev. B, Condens. Matter*, vol. 98, no. 21, Dec. 2018, Art. no. 214410.
- [26] M. Lapedus. (2019). *Challenges in Making and Testing STT-MRAM*. [Online]. Available: <https://semiengineering.com/challenges-in-making-and-testing-mram/>
- [27] M. Wang et al., “Field-free switching of a perpendicular magnetic tunnel junction through the interplay of spin-orbit and spin-transfer torques,” *Nature Electron.*, vol. 1, no. 11, pp. 582–588, Nov. 2018.
- [28] Y. Oh et al., “Field-free switching of perpendicular magnetization through spin-orbit torque in antiferromagnet/ferromagnet oxide structures,” *Nature Nanotechnol.*, vol. 11, no. 10, pp. 878–884, Oct. 2016.
- [29] G. Yu et al., “Switching of perpendicular magnetization by spin-orbit torques in the absence of external magnetic fields,” *Nature Nanotechnologies*, vol. 9, no. 7, pp. 548–554, Jul. 2014.
- [30] S. Senni, “Exploration of non-volatile magnetic memory for processor architecture,” Ph.D. dissertation, Univ. Montpellier, Montpellier, France, 2015.
- [31] K. Garello et al., “SOT-MRAM 300 nm integration for low power and ultrafast embedded memories,” in *Proc. IEEE Symp. VLSI Circuits*, Jun. 2018, pp. 81–82.
- [32] S. Bhatti, R. Sbiaa, A. Hirohata, H. Ohno, S. Fukami, and S. N. Piramanayagam, “Spintronics based random access memory: A review,” *Mater. Today*, vol. 20, no. 9, pp. 530–548, Nov. 2017.
- [33] S. Manipatruni, D. E. Nikonov, R. Ramesh, H. Li, and I. A. Young, “Spin-orbit logic with magnetoelectric nodes: A scalable charge mediated nonvolatile spintronic logic,” 2015, *arXiv:1512.05428*. [Online]. Available: <http://arxiv.org/abs/1512.05428>
- [34] W. Kang, Y. Huang, X. Zhang, Y. Zhou, and W. Zhao, “Skyrmion-electronics: An overview and outlook,” *Proc. IEEE*, vol. 104, no. 10, pp. 2040–2061, Oct. 2016.
- [35] K. M. Song et al., “Skyrmion-based artificial synapses for neuromorphic computing,” *Nature Electron.*, vol. 3, no. 3, pp. 148–155, Mar. 2020, doi: 10.1038/s41928-020-0385-0.
- [36] M. Fieback, M. Taouil, and S. Hamdioui, “Testing resistive memories: Where are we and what is missing?” in *Proc. IEEE Int. Test Conf.*, Phoenix, AZ, USA, Oct. 2018, pp. 1–9.
- [37] W. S. Zhao et al., “Failure and reliability analysis of STT-MRAM,” *Microelectron. Rel.*, vol. 52, nos. 9–10, pp. 1848–1852, Sep. 2012.
- [38] W. S. Zhao, T. Devolder, Y. Lakys, J. O. Klein, C. Chappert, and P. Mazoyer, “Design considerations and strategies for high-reliable STT-MRAM,” *Microelectron. Rel.*, vol. 51, nos. 9–11, pp. 1454–1458, Sep. 2011.
- [39] E. I. Vatajelu, P. Pouyan, and S. Hamdioui, “State of the art and challenges for test and reliability of emerging nonvolatile resistive memories,” *Int. J. Circuit Theory Appl.*, vol. 46, no. 1, pp. 4–28, Jan. 2018.
- [40] H. Naeimi, C. Augustine, A. Raychowdhury, S.-L. Lu, and J. Tschanz, “STTRAM scaling and retention failure,” *Intel Technol. J.*, vol. 17, no. 1, pp. 54–75, Jan. 2013.
- [41] A. Chintaluri, H. Naeimi, S. Natarajan, and A. Raychowdhury, “Analysis of defects and variations in embedded spin transfer torque (STT) MRAM arrays,” *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 6, no. 3, pp. 319–329, Sep. 2016.
- [42] A. Chintaluri, A. Parihar, S. Natarajan, H. Naeimi, and A. Raychowdhury, “A model study of defects and faults in embedded spin transfer torque (STT) MRAM arrays,” in *Proc. IEEE 24th Asian Test Symp.*, Mumbai, India, Nov. 2015, pp. 187–192.
- [43] S. M. Nair, R. Bishnoi, and M. B. Tahoori, “A comprehensive framework for parametric failure modeling and yield analysis of STT-MRAM,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 27, no. 7, pp. 1697–1710, Jul. 2019.
- [44] S. M. Nair, R. Bishnoi, and M. B. Tahoori, “Parametric failure modeling and yield analysis for STT-MRAM,” in *Proc. IEEE/ACM Design, Automat. Test Eur.*, Dresden, Germany, Mar. 2018, pp. 265–268.
- [45] R. Bishnoi, F. Oboril, and M. B. Tahoori, “Design of defect and fault-tolerant nonvolatile spintronic flip-flops,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 4, pp. 1421–1432, Apr. 2017.
- [46] S. M. Nair, R. Bishnoi, M. S. Golanbari, F. Oboril, F. Hameed, and M. B. Tahoori, “VAET-STT: STT-MRAM analysis and design space exploration tool,” *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 7, pp. 1396–1407, Jul. 2017.
- [47] S. M. Nair, R. Bishnoi, M. S. Golanbari, F. Oboril, and M. B. Tahoori, “VAET-STT: A variation aware estimator tool for STT-MRAM based memories,” in *Proc. Design, Autom. Test Eur. Conf. Exhib.*, Lausanne, Switzerland, Mar. 2017, pp. 1456–1461.
- [48] M. Hosomi et al., “A novel nonvolatile memory with spin torque transfer magnetization switching: SPIN-RAM,” in *IEDM Tech. Dig.*, Washington, DC, USA, Dec. 2005, pp. 459–462.
- [49] A. Bosio et al., “Rebooting computing: The challenges for test and reliability,” in *Proc. IEEE Int. Symp. Defect Fault Tolerance VLSI Nanotechnol. Syst.*, Noordwijk, The Netherlands, Oct. 2019, pp. 8138–8143.
- [50] A. Raychowdhury, D. Somasekhar, T. Karnik, and V. De, “Design space and scalability exploration of 1T-1STT MTJ memory arrays in the presence of variability and disturbances,” in *IEDM Tech. Dig.*, Baltimore, MD, USA, Dec. 2009, pp. 1–4.
- [51] R. Bishnoi, M. Ebrahimi, F. Oboril, and M. B. Tahoori, “Improving write performance for STT-MRAM,” *IEEE Trans. Magn.*, vol. 52, no. 8, pp. 1–11, Aug. 2016.
- [52] K. Munira, W. H. Butler, and A. W. Ghosh, “A quasi-analytical model for energy-delay-reliability tradeoff studies during write operations in a perpendicular STT-RAM cell,” *IEEE Trans. Electron Devices*, vol. 59, no. 8, pp. 2221–2226, Aug. 2012.
- [53] A. Vatankehaghadi, S. Huda, and A. Sheikholeslami, “A survey on circuit modeling of spin-transfer-torque magnetic tunnel junctions,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 9, pp. 2634–2643, Sep. 2014.
- [54] T. Devolder, C. Chappert, and K. Ito, “Subnanosecond spin-transfer switching: Comparing the benefits of free-layer or pinned-layer biasing,” *Phys. Rev. B, Condens. Matter*, vol. 75, no. 22, Jun. 2007, Art. no. 224430.
- [55] H. Tomita et al., “Unified understanding of both thermally assisted and precessional spin-transfer switching in perpendicularly magnetized giant magnetoresistive nanopillars,” *Appl. Phys. Lett.*, vol. 102, no. 4, Jan. 2013, Art. no. 042409.
- [56] A. Raychowdhury, D. Somasekhar, T. Karnik, and V. K. De, “Modeling and analysis of read (RD) disturb in 1T-1STT MTJ memory bits,” in *Proc. 68th Device Res. Conf.*, South Bend, IN, USA, Jun. 2010, pp. 43–44.
- [57] A. Raychowdhury, “Pulsed READ in spin transfer torque (STT) memory bitcell for lower READ disturb,” in *Proc. IEEE/ACM International Symposium on Nanoscale Architectures*, Brooklyn, NY, USA, Jul. 15–17, 2013, pp. 34–35.
- [58] B. Dieny, R. Goldfarb, and K.-J. Lee, *Introduction to Magnetic Random Access Memory*. Hoboken, NJ, USA: Wiley, 2016.
- [59] A. Bosio, L. Dilillo, P. Girard, S. Pravossoudovitch, and A. Virazel, *Advanced Test Methods for SRAMs*. New York, NY, USA: Springer, 2009.
- [60] C.-L. Su, C.-W. Tsai, C.-W. Wu, C.-C. Hung, Y.-S. Chen, and M.-J. Kao, “Testing MRAM for write disturbance fault,” in *Proc. IEEE Int. Test Conf.*, Santa Clara, CA, USA, Oct. 2006, pp. 1–9.
- [61] S. M. Nair et al., “Defect injection, fault modeling and test algorithm generation methodology for STT-MRAM,” in *Proc. IEEE Int. Test Conf.*, Phoenix, AZ, USA, Oct. 2018, pp. 1–10.
- [62] C.-L. Su et al., “MRAM defect analysis and fault modeling,” in *Proc. IEEE Int. Test Conf.*, Charlotte, NC, USA, Oct. 2004, pp. 124–133.
- [63] J. Azevedo et al., “Coupling-based resistive-open defects in TAS-MRAM architectures,” in *Proc. IEEE Eur. Test Symp.*, Annecy, France, May/Jun. 2012, p. 1.
- [64] B. Cockburn, “Tutorial on DRAM fault modeling and test pattern design,” in *Proc. IEEE Int. Workshop Memory Technol., Design Testing*, San Jose, CA, USA, Aug. 1998, p. 66.
- [65] O. Gineez, J.-M. Daga, P. Girard, C. Landrault, S. Pravossoudovitch, and A. Virazel, “Embedded flash testing: Overview and perspectives,” in *Proc. IEEE Int. Conf. Design Test Integr. Syst.*, San Jose, CA, USA, Sep. 2006, pp. 86–92.
- [66] C.-Y. Chen et al., “RRAM defect modeling and failure analysis based on march test and a novel squeeze-search scheme,” *IEEE Trans. Comput.*, vol. 64, no. 1, pp. 180–190, Jan. 2015.
- [67] S. Tehrani, “Status and prospect for MRAM technology,” in *Proc. IEEE Hot Chips Symp. Nonvolatile Memory Seminar*, Stanford, CA, USA, Aug. 2010, pp. 1–23.
- [68] *Everspin Begins Production of 1Gb STT-MRAM*. Accessed: Jun. 24, 2019. [Online]. Available: <https://www.anandtech.com/show/14580/everspin-begins-production-of-1gb-sttmram>
- [69] *Intel STT-MRAM Technology is Ready for Mass Production*. Accessed: Feb. 21, 2019. [Online]. Available: <https://www.tomshardware.com/news/intel-stt-mram-mass-production,38665.html>
- [70] C.-L. Su, C.-W. Tsai, C.-W. Wu, C.-C. Hung, Y.-S. Chen, and M.-J. Kao, “Testing MRAM for write disturbance fault,” in *Proc. IEEE Int. Test Conf.*, Santa Clara, CA, USA, Oct. 2006, pp. 1–9.
- [71] C.-L. Su et al., “Write disturbance modeling and testing for MRAM,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 16, no. 3, pp. 277–288, Mar. 2008.
- [72] Z. Zhang, W. Xiao, N. Park, and D. J. Lilja, “Memory module-level testing and error behaviors for phase change memory,” in *Proc. IEEE 30th Int. Conf. Comput. Design*, Montreal, QC, Canada, Sep. 2012, pp. 358–363.
- [73] J. Azevedo et al., “Analysis of resistive-open defects in TAS-MRAM array [poster],” in *Proc. IEEE Int. Conf.*, Anaheim, CA, USA, Sep. 20–22, 2011.
- [74] J. Azevedo et al., “Impact of resistive-open defects on the heat current of TAS-MRAM architectures,” in *Proc. IEEE/ACM Design, Automat. Test Eur. Conf. Exhib.*, Dresden, Germany, Mar. 2012, pp. 532–537.
- [75] J. Azevedo et al., “A complete resistive-open defect analysis for thermally assisted switching MRAMs,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 22, no. 11, pp. 2326–2335, Nov. 2014.
- [76] J. Azevedo et al., “Impact of resistive-bridge defects in TAS-MRAM architectures,” in *Proc. IEEE 21st Asian Test Symp.*, Niigata, Japan, Nov. 2012, pp. 125–130.
- [77] J. Azevedo, “Test and reliability of MRAMs,” Ph.D. dissertation, University of Montpellier, Montpellier, France, Oct. 2013.
- [78] M. E. Baraji, V. Javerliac, W. Guo, G. Prenat, and B. Dieny, “Dynamic compact model of thermally-assisted switching magnetic tunnel

- junctions," *J. Appl. Phys.*, vol. 106, no. 12, Dec. 2009, Art. no. 123906.
- [79] S. M. Nair, R. Bishnoi, M. B. Tahoori, H. Grigoryan, and G. Tshagharyan, "Variation-aware fault modeling and test generation for STT-MRAM," in *Proc. IEEE 25th Int. Symp. On-Line Test. Robust Syst. Design (IOLTS)*, Rhodes, Greece, Jul. 2019, pp. 80–83.
- [80] I. Yoon, A. Chintaluri, and A. Raychowdhury, "EMACS: Efficient MBIST architecture for test and characterization of STT-MRAM arrays," in *Proc. IEEE Int. Test Conf.*, Fort Worth, TX, USA, Nov. 2016, pp. 4.3.1–4.3.10.
- [81] I. Yoon and A. Raychowdhury, "Test challenges in embedded STT-MRAM arrays," in *Proc. 18th Int. Symp. Qual. Electron. Design*, Santa Clara, CA, USA, Mar. 2017, pp. 35–38.
- [82] S. Hamdioui, P. Pouyan, H. Li, Y. Wang, A. Raychowdhury, and I. Yoon, "Test and reliability of emerging non-volatile memories," in *Proc. IEEE Asian Test Symp.*, Taipei, Taiwan, Nov. 2017, pp. 175–183.
- [83] R. Heindl, W. H. W. H. Rippard, S. E. Russek, M. R. Pufall, and A. B. Kos, "Validity of the thermal activation model for spin-transfer torque switching in magnetic tunnel junctions," *J. Appl. Phys.*, vol. 109, no. 7, p. 073910, Apr. 2011.
- [84] R. Bishnoi, M. Ebrahimi, F. Oboril, and M. B. Tahoori, "Read disturb fault detection in STT-MRAM," in *Proc. IEEE Int. Test Conf.*, Seattle, WA, USA, Oct. 2014, pp. 23.3.1–23.3.7.
- [85] A. F. Gomez, F. Forero, K. Roy, and V. Champac, "Robust detection of bridge defects in STT-MRAM under process variations," in *Proc. IFIP/IEEE Int. Conf. Very Large Scale Integr.*, Verona, Italy, Oct. 2018, pp. 65–70.
- [86] V. Champac, A. Gomez, F. Forero, and K. Roy, "Analysis of bridge defects in STT-MRAM cells under process variations and a robust DFT technique for their detection," in *Proc. IFIP/IEEE Int. Conf. Very Large Scale Integr.*, Verona, Italy, Oct. 2018, pp. 207–231.
- [87] S. Ikeda et al., "Tunnel magnetoresistance of 604% at 300K by suppression of Ta diffusion in CoFeB/MgO/CoFeB pseudo-spin-valves annealed at high temperature," *Appl. Phys. Lett.*, vol. 93, no. 8, p. 082508, Aug. 2008.
- [88] B. Wu, Y. Cheng, J. Yang, A. Todri-Sanial, and W. Zhao, "Temperature impact analysis and access reliability enhancement for 1T1MTJ STT-MRAM," *IEEE Trans. Rel.*, vol. 65, no. 4, pp. 1755–1768, Dec. 2016.
- [89] L. Zhang et al., "Quantitative evaluation of reliability and performance for STT-MRAM," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2016, pp. 1150–1153.
- [90] L. Zhang et al., "Addressing the thermal issues of STT-MRAM from compact modeling to design techniques," *IEEE Trans. Nanotechnol.*, vol. 17, no. 2, pp. 345–351, Mar. 2018.
- [91] Y. Xie, B. Behin-Aein, and A. Ghosh, "Numerical Fokker-Planck simulation of stochastic write error in spin torque switching with thermal noise," in *Proc. 74th Annu. Device Res. Conf.*, Newark, DE, USA, Jun. 2016, pp. 1–2.
- [92] L. Zhang et al., "Reliability and performance evaluation for STT-MRAM under temperature variation," in *Proc. 17th Int. Conf. Thermal, Mech. Multi-Phys. Simul. Exp. Microelectron. Microsyst.*, Montpellier, France, Apr. 2016, pp. 1–4.
- [93] J. F. Kong, K. Eason, K. P. Tan, and R. Shiba, "Parameter variation investigation of magnetic tunnel junctions," in *Proc. IEEE Asia-Pacific Magn. Recording Conf.*, Singapore, Oct./Nov. 2012, pp. 1–2.
- [94] E. Cheshmikhani, H. Farbeh, and H. Asadi, "A system-level framework for analytical and empirical reliability exploration of STT-MRAM caches," *IEEE Trans. Rel.*, vol. 69, no. 2, pp. 594–610, Jun. 2019.
- [95] R. C. Sousa et al., "Tunneling hot spots and heating in magnetic tunnel junctions," *J. Appl. Phys.*, vol. 95, no. 11, pp. 6783–6785, Jun. 2004.
- [96] L.-B. Faber, W. Zhao, J.-O. Klein, T. Devolder, and C. Chappert, "Dynamic compact model of spin-transfer torque based magnetic tunnel junction (MTJ)," in *Proc. 4th Int. Conf. Design Technol. Integr. Syst. Nanoscale Era*, Cairo, Egypt, Apr. 2009, pp. 130–135.
- [97] E. Wu, B. Li, J. H. Stathis, R. Achanta, R. Filippi, and P. McLaughlin, "A time-dependent clustering model for non-uniform dielectric breakdown," in *IEDM Tech. Dig.*, Washington, DC, USA, Dec. 2013, pp. 15.3.1–15.3.4.
- [98] S. Long et al., "A model for the set statistics of RRAM inspired in the percolation model of oxide breakdown," *IEEE Electron Device Lett.*, vol. 34, no. 8, pp. 999–1001, Aug. 2013.
- [99] J. Suñé, E. Miranda, D. Jimenez, S. Long, and M. Liu, "From dielectric failure to memory function: Learning from oxide breakdown for improved understanding of resistive switching memories," in *Proc. 11th Annu. Non-Volatile Memory Technol. Symp.*, Shanghai, China, Nov. 2011, pp. 1–6.
- [100] J. J. Kan, "Engineering of metallic multilayers and spin transfer torque devices," Ph.D. dissertation, Univ. California, San Diego, CA, USA, 2014.
- [101] A. A. Khan, J. Schmalhorst, A. Thomas, O. Schebaum, and G. Reiss, "Dielectric breakdown in Co-Fe-B/MgO/Co-Fe-B magnetic tunnel junction," *J. Appl. Phys.*, vol. 103, no. 12, Dec. 2008, Art. no. 123705.
- [102] S. Amara-Dababi et al., "Charge trapping-detraping mechanism of barrier breakdown in MgO magnetic tunnel junctions," *Appl. Phys. Lett.*, vol. 99, no. 8, Aug. 2011, Art. no. 083501.
- [103] K.-S. Kim, Y. M. Jang, C. H. Nam, K.-S. Lee, and B. K. Cho, "Stress polarity dependence of breakdown characteristics in magnetic tunnel junctions," *J. Appl. Phys.*, vol. 99, no. 8, p. 08K705, Aug. 2006.
- [104] C. Ho, G. D. Panagopoulos, S. Y. Kim, Y. Kim, D. Lee, and K. Roy, "A physical model to predict STT-MRAM performance degradation induced by TDD," in *Proc. IEEE Device Res. Conf.*, Notre Dame, IN, USA, Jun. 2013, pp. 59–60.
- [105] K. Hirose, D. Kobayashi, T. Ito, and T. Endoh, "Memory reliability of spintronic materials and devices for disaster-resilient computing against radiation-induced bit flips on the ground," *Jpn. Soc. Appl. Phys.*, vol. 56, no. 8, p. 0802A5, Aug. 2017.
- [106] J.-Y. Park, J.-M. Kim, J. Ryu, J. Jeong, and B.-G. Park, "Effects of proton and ion beam radiation on magnetic tunnel junctions," *Thin Solid Films*, vol. 686, Sep. 2019, Art. no. 137432.
- [107] X. Shi, F. Wu, X. Guan, and C. Xie, "Error behaviors testing with temperature and magnetism dependency for MRAM," in *Proc. IEEE 34th Int. Conf. Comput. Design*, Scottsdale, AZ, USA, Oct. 2016, pp. 356–359.
- [108] Y. Ji et al., "Reliability of 8 Mbit embedded-STT-MRAM in 28nm FDSOI technology," in *Proc. IEEE Int. Rel. Phys. Symp.*, Monterey, CA, USA, Mar./Apr. 2019, pp. 1–3.
- [109] S. Mahawar, S. Verma, P. K. Pal, and B. K. Kaushik, "Highly reliable STT MRAM using fully depleted body and buried 4H-SiC NMOS," in *Proc. IEEE Int. Conf. Electron Devices Solid-State Circuits*, Singapore, Jun. 2015, pp. 705–708.
- [110] O. Gonçalves, G. Prenat, and B. Dieny, "Radiation hardened MRAM-based FPGA," *IEEE Trans. Magn.*, vol. 49, no. 7, pp. 4355–4358, Jul. 2013.
- [111] S. Peng et al., "Modulation of heavy metal/ferromagnetic metal interface for high-performance spintronic devices," *Adv. Electron. Mater.*, vol. 5, no. 8, Aug. 2019, Art. no. 1900134.
- [112] C. Augustine, A. Raychowdhury, D. Somasekhar, J. Tschanz, V. De, and K. Roy, "Design space exploration of typical STT MTJ stacks in memory arrays in the presence of variability and disturbances," *IEEE Trans. Electron Devices*, vol. 58, no. 12, pp. 4333–4343, Dec. 2011.
- [113] Z. Pajouhi, X. Fong, and K. Roy, "Device/circuit/architecture co-design of reliable STT-MRAM," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, 2015, pp. 1437–1442.
- [114] N. Sayed, R. Bishnoi, and M. B. Tahoori, "Fast and reliable STT-MRAM using nonuniform and adaptive error detecting and correcting scheme," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 27, no. 6, pp. 1329–1342, Jun. 2019.
- [115] Z. Mei, K. Cai, and B. Dai, "Polar coding for spin-torque transfer magnetic random access memory (STT-MRAM)," in *Proc. IEEE Int. Magn. Conf.*, Singapore, Apr. 2018, p. 1.
- [116] N. Sayed, F. Oboril, R. Bishnoi, and M. B. Tahoori, "Leveraging systematic unidirectional error-detecting codes for fast STT-MRAM cache," in *Proc. IEEE 35th VLSI Test Symp. (VTS)*, Las Vegas, NV, USA, Apr. 2017, pp. 1–6.
- [117] K. Cai and K. A. S. Imminck, "Cascaded channel model, analysis, and hybrid decoding for spin-torque transfer magnetic random access memory," *IEEE Trans. Magn.*, vol. 53, no. 11, pp. 1–11, Nov. 2017.
- [118] W. Kang, W. Zhao, L. Yang, J.-O. Klein, Y. Zhang, and D. Ravclosa, "One-step majority-logic-decodable codes enable STT-MRAM for high speed working memories," in *Proc. IEEE Non-Volatile Memory Syst. Appl. Symp.*, Chongqing, China, Aug. 2014, pp. 1–6.
- [119] E. Aliagha, A. M. H. Monazzah, and H. Farbeh, "REACT: Read/write error rate aware coding technique for emerging STT-MRAM caches," *IEEE Trans. Magn.*, vol. 55, no. 5, pp. 1–8, May 2019.
- [120] N. Sayed, M. Ebrahimi, R. Bishnoi, and M. B. Tahoori, "Opportunistic write for fast and reliable STT-MRAM," in *Proc. IEEE/ACM Design, Autom. Test Eur. Conf. Exhib.*, Lausanne, Switzerland, Mar. 2017, pp. 554–559.
- [121] A. Das and N. A. Toubia, "Online correction of hard errors and soft errors via one-step decodable OLS codes for emerging last level caches," in *Proc. IEEE Latin Amer. Test Symp.*, Santiago, Chile, Mar. 2019, pp. 1–6.
- [122] L. Jiang, Y. Zhang, and J. Yang, "Mitigating write disturbance in super-dense phase change memories," in *Proc. 44th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw.*, Atlanta, GA, USA, Jun. 2014, pp. 216–227.
- [123] M. Imran, T. Kwon, J. M. You, and J.-S. Yang, "Flipcy: Efficient pattern redistribution for enhancing MLC PCM reliability and storage density," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Denver, CO, USA, Nov. 2019, pp. 1–7.
- [124] P. M. Palangappa, J. Li, and K. Mohanram, "WOM-code solutions for low latency and high endurance in phase change memory," *IEEE Trans. Comput.*, vol. 65, no. 4, pp. 1025–1040, Apr. 2016.
- [125] M. Wang, N. Deng, H. Wu, and Q. He, "Theory study and implementation of configurable ECC on RRAM memory," in *Proc. 15th Non-Volatile Memory Technol. Symp. (NVMTS)*, Beijing, China, Oct. 2015, pp. 1–3.
- [126] Y. Chen, Y. Zhang, and P. Wang, "Probabilistic design in spintronic memory and logic circuit," in *Proc. 17th Asia South Pacific Design Autom. Conf.*, Sydney, NSW, Australia, Jan. 2012, pp. 323–328.
- [127] H. Noguchi et al., "4Mb STT-MRAM-based cache with memory-access-aware power optimization and write-verify-write/read-modify-write scheme," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, San Francisco, CA, USA, Jan./Feb. 2016, pp. 132–133.
- [128] E. Y. Chen et al., "Comparison of oxidation methods for magnetic tunnel junction material," *J. Appl. Phys.*, vol. 87, no. 9, pp. 6061–6063, May 2000.
- [129] J. Yang et al., "Radiation-induced soft error analysis of STT-MRAM: A device to circuit approach," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 35, no. 3, pp. 380–393, Mar. 2016.
- [130] A. Ahari, M. Ebrahimi, F. Oboril, and M. Tahoori, "Improving reliability, performance, and energy efficiency of STT-MRAM with dynamic write latency," in *Proc. 33rd IEEE Int. Conf. Comput.*

- Design, New York, NY, USA, Oct. 2015, pp. 109–116.
- [131] H. Noguchi et al., “A 3.3 NS-access-time 71.2 μ W/MHz 1Mb embedded STT-MRAM using physically eliminated read-disturb scheme and normally-off memory architecture,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, San Francisco, CA, USA, Feb. 22–26, 2015, pp. 136–138.
- [132] W. Kang et al., “DFSTT-MRAM: Dual functional STT-MRAM cell structure for reliability enhancement and 3-D MLC functionality,” *IEEE Trans. Magn.*, vol. 50, no. 6, pp. 1–7, Jun. 2014.
- [133] T. Zheng, J. Park, M. Orshansky, and M. Erez, “Variable-energy write STT-RAM architecture with bit-wise write-completion monitoring,” in *Proc. Int. Symp. Low Power Electron. Design*, Beijing, China, Sep. 2013, pp. 229–234.
- [134] N. Sayed, R. Bishnoi, F. Oboril, and M. B. Tahoori, “A cross-layer adaptive approach for performance and power optimization in STT-MRAM,” in *Proc. IEEE/ACM Design, Automat. Test Eur. Conf. Exhib.*, Dresden, Germany, Mar. 2018, pp. 791–796.
- [135] B. Bose and L. Der Je, “Systematic unidirectional error-detecting codes,” *IEEE Trans. Comput.*, vol. C-34, no. 11, pp. 1026–1032, Nov. 1985.
- [136] N. Sayed, S. M. Nair, R. Bishnoi, and M. B. Tahoori, “Process variation and temperature aware adaptive scrubbing for retention failures in STT-MRAM,” in *Proc. 23rd Asia South Pacific Design Autom. Conf.*, Jeju, South Korea, Jan. 2018, pp. 203–208.
- [137] Y. Zhang, X. Wang, and Y. Chen, “STT-RAM cell design optimization for persistent and non-persistent error rate reduction: A statistical design view,” in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design*, San Jose, CA, USA, Nov. 2011, pp. 471–477.
- [138] F. García-Redondo, P. Royer, M. López-Vallejo, H. Aparicio, P. Iturero, and C. A. López-Barrio, “Reconfigurable writing architecture for reliable RRAM operation in wide temperature ranges,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 4, pp. 1224–1235, Apr. 2017.
- [139] A. M. S. Tosson, S. Yu, M. Anis, and L. Wei, “Mitigating the effect of reliability soft-errors of RRAM devices on the performance of RRAM-based neuromorphic systems,” in *Proc. Great Lakes Symp. VLSI*, Banff, AB, Canada, May 2017, pp. 53–58.
- [140] E. Cheshmikhani, H. Farbeh, S. G. Miremadi, and H. Asadi, “TA-LRW: A replacement policy for error rate reduction in STT-MRAM caches,” *IEEE Trans. Comput.*, vol. 68, no. 3, pp. 455–470, Mar. 2019.
- [141] B. Wu et al., “A novel high performance and energy efficient NUCA architecture for STT-MRAM LLCs with thermal consideration,” *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 39, no. 4, pp. 803–815, Apr. 2019.
- [142] X. Guo, P. Girard, J. Chen, K. Liu and Y. Cheng, “DOVA: A dynamic overwriting voltage adjustment for STT-RAM I1 cache,” in *Proc. IEEE Int. Symp. Quality Electron. Design*, Santa Clara, CA, USA, Mar. 2020, pp. 1–6.
- [143] W. Wen, Y. Zhang, and J. Yang, “Read error resilient MLC STT-MRAM based last level cache,” in *Proc. IEEE Int. Conf. Comput. Design*, Boston, MA, USA, Nov. 2017, pp. 455–462.
- [144] E. Cheshmikhani, H. Farbeh, and H. Asadi, “Enhancing reliability of STT-MRAM caches by eliminating read disturbance accumulation,” in *Proc. Design, Autom. Test Eur. Conf. Exhib.*, Florence, Italy, Mar. 2019, pp. 854–859.
- [145] L. Xue et al., “An adaptive 3T-3MTJ memory cell design for STT-MRAM-based LLCs,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 3, pp. 484–495, Mar. 2018.
- [146] H. Sun, C. Liu, W. Xu, J. Zhao, N. Zheng, and T. Zhang, “Using magnetic RAM to build low-power and soft error-resilient I1 cache,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 1, pp. 19–28, Jan. 2012.
- [147] S. Swami and K. Mohanram, “Adam: Architecture for write disturbance mitigation in scaled phase change memory,” in *Proc. IEEE/ACM Design, Autom. Test Eur. Conf. Exhib.*, Dresden, Germany, Mar. 2018, pp. 1235–1240.
- [148] A. Jadidi, M. Arjomand, M. K. Tavana, D. R. Kaeli, M. T. Kandemir, and C. R. Das, “Exploring the potential for collaborative data compression and hard-error tolerance in PCM memories,” in *Proc. 47th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw.*, Denver, CO, USA, Jun. 2017, pp. 85–96.
- [149] M. Liu, L. Xia, Y. Wang, and K. Chakrabarty, “Design of fault-tolerant neuromorphic computing systems,” in *Proc. IEEE 23rd Eur. Test Symp.*, Bremen, Germany, May 2018, pp. 1–9.
- [150] B. Li, L. Xia, P. Gu, Y. Wang, and H. Yang, “Merging the interface: Power, area and accuracy co-optimization for RRAM crossbar-based mixed-signal computing system,” in *Proc. IEEE/ACM Design Automat. Conf.*, San Francisco CA, USA, Jun. 2015, pp. 1–6.
- [151] J.-Y. Hu, K.-W. Hou, C.-Y. Lo, Y.-F. Chou, and C.-W. Wu, “RRAM-based neuromorphic hardware reliability improvement by self-healing and error correction,” in *Proc. IEEE Int. Test Conf. Asia*, Harbin, China, Aug. 2018, pp. 19–24.
- [152] F. Oboril, R. Bishnoi, M. Ebrahimi, and M. B. Tahoori, “Evaluation of hybrid memory technologies using SOT-MRAM for on-chip cache hierarchy,” *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 34, no. 3, pp. 367–380, Mar. 2015.
- [153] B. Wang, Z. Wang, C. Hu, Y. Zhao, Y. Zhang, and W. Zhao, “Radiation-hardening techniques for spin orbit torque-MRAM peripheral circuitry,” *IEEE Trans. Magn.*, vol. 54, no. 11, pp. 1–5, Nov. 2018.
- [154] B. Wang, Z. Wang, C. Hu, Y. Zhao, Y. Zhang, and W. Zhao, “Novel radiation hardening read/write circuits using feedback connections for spin-orbit torque magnetic random access memory,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 66, no. 5, pp. 1853–1862, May 2019.
- [155] Z. Wang et al., “High-density NAND-like spin transfer torque memory with spin orbit torque erase operation,” *IEEE Electron Device Lett.*, vol. 39, no. 3, pp. 343–346, Mar. 2018.
- [156] Z. Wang et al., “Proposal of toggle spin torques magnetic RAM for ultrafast computing,” *IEEE Electron Device Lett.*, vol. 40, no. 5, pp. 726–729, May 2019.
- [157] A. Gebregiorgis, R. Bishnoi, and M. B. Tahoori, “Spintronic normally-off heterogeneous system-on-chip design,” in *Proc. Design, Autom. Test Eur. Conf. Exhib.*, Dresden, Germany, Mar. 2018, pp. 113–118.
- [158] R. Bishnoi, F. Oboril, and M. B. Tahoori, “Non-volatile non-shadow flip-flop using spin orbit torque for efficient normally-off computing,” in *Proc. 21st Asia South Pacific Design Autom. Conf. (ASP-DAC)*, Jan. 2016, pp. 769–774.
- [159] R. Bishnoi, F. Oboril, and M. B. Tahoori, “Low-power multi-port memory architecture based on spin orbit torque magnetic devices,” in *Proc. ACM Int. Great Lakes Symp. VLSI*, Boston, MA, USA, 2016, pp. 409–414.
- [160] S. Mittal et al., “Architecting SOT-RAM based GPU register file,” in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI*, Bochum, Germany, Jul. 2017, pp. 38–44.
- [161] E. Vianello et al., “Resistive memories for spike-based neuromorphic circuits,” in *Proc. IEEE Int. Memory Workshop (IMW)*, May 2017, pp. 1–6.
- [162] A. F. Vincent et al., “Spin-transfer torque magnetic memory as a stochastic memristive synapse for neuromorphic systems,” *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 2, pp. 166–174, Apr. 2015.
- [163] G. Srinivasan, A. Sengupta, and K. Roy, “Magnetic tunnel junction based long-term short-term stochastic synapse for a spiking neural network with on-chip STDP learning,” *Sci. Rep.*, vol. 6, no. 1, Sep. 2016, 29545.
- [164] D. Soudry, D. Di Castro, A. Gal, A. Kolodny, and S. Kvatinsky, “Memristor-based multilayer neural networks with online gradient descent training,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2408–2421, Oct. 2015.
- [165] N. Zheng and P. Mazumder, “Learning in memristor crossbar-based spiking neural networks through modulation of weight-dependent spike-timing-dependent plasticity,” *IEEE Trans. Nanotechnol.*, vol. 17, no. 3, pp. 520–532, May 2018.
- [166] G. Prenat et al., “Ultra-fast and high-reliability SOT-MRAM: From cache replacement to normally-off computing,” *IEEE Trans. Multi-Scale Comput. Syst.*, vol. 2, no. 1, pp. 49–60, Jan. 2016.
- [167] S. M. Nair, C. Münch, and M. B. Tahoori, “Defect characterization and test generation for spintronic-based compute-in-memory,” in *Proc. IEEE Eur. Test Symp.*, Tallin, Estonia, May 2020, pp. 1–10.
- [168] R. Bishnoi et al., “Special session—Emerging memristor based memory and CIM architecture: Test, repair and yield analysis,” in *Proc. IEEE VLSI Test Symp.*, San Diego, CA, USA, Apr. 2020, pp. 1–10.
- [169] C. Münch, R. Bishnoi, and M. B. Tahoori, “Tolerating retention failures in neuromorphic fabric based on emerging resistive memories,” in *Proc. ACM/IEEE Asia South Pacific Design Automat. Conf.*, Beijing, China, Jan. 2020, pp. 393–400.
- [170] C. Münch, R. Bishnoi, and M. B. Tahoori, “Reliable in-memory neuromorphic computing using spintronics,” in *Proc. ACM/IEEE Asia South Pacific Design Automat. Conf.*, Tokyo, Japan, Jan. 2019, pp. 230–236.

ABOUT THE AUTHORS

Patrick Girard (Fellow, IEEE) received the Ph.D. degree in microelectronics from the University of Montpellier, Montpellier, France, in 1992.

He is currently the Research Director of the French National Center for Scientific Research (CNRS), Paris, France. He is also with the Microelectronics Department, Laboratory of Computer Science, Robotics and Microelectronics of Montpellier (LIRMM), Montpellier. He is also the Director of the International Associated Laboratory “LAFISI”



(French-Italian Research Laboratory on Hardware-Software Integrated Systems), Montpellier. He is also the Deputy Director of the French Scientific Network dedicated to research in the fields of system-on-chip, embedded systems, and connected objects (SOC2). His research interests include all aspects of digital and memory testing, with an emphasis on critical constraints, such as timing and power. Robust design of neuromorphic circuits and machine learning for test and diagnosis are also part of his new research activities. He has supervised 40 Ph.D. dissertations and has published eight books or book chapters, 80 journal articles, and more than 250 conference and symposium papers in these fields.

Yuanqing Cheng (Senior Member, IEEE) received the Ph.D. degree from the Key Laboratory of Computer System and Architecture, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2012.

After spending one year as a Postdoctoral Researcher at the Laboratory of Computer Science, Robotics and Microelectronics of Montpellier (LIRMM), Montpellier, France, he joined Beihang University, Beijing, as an Assistant Professor. His research interests include a reliable and physical design for 3-D integrated circuits, reliable, and low-power design for emerging technologies, such as spintronics and carbon nanotube technologies.



Arnaud Virazel (Member, IEEE) received the Ph.D. degree in microelectronics from the University of Montpellier, Montpellier, France, in 2001.

He is currently an Associate Professor with the University of Montpellier, where he is with the Microelectronics Department, Laboratory of Informatics, Robotics and Microelectronics of Montpellier (LIRMM), and he is responsible for the Test and dEpendability of microelectronic integrated SysTems (TEST) Team. He is the Deputy Head of the Electrical Engineering Master Program (about 200 students) in charge of the first year and the Integrated Electronic Systems specialization at the University of Montpellier. He has published three books or book chapters, 40 journal articles, and more than 140 conference and symposium papers spanning diverse disciplines, including Design-for-Testability (DFT), built-in self-test (BIST), diagnosis, reliability, delay testing, power-aware testing, and memory testing. His teaching topics are mainly focusing on digital circuit design, test, and reliability.



Weisheng Zhao (Fellow, IEEE) received the Ph.D. degree from the University of Paris Sud, Orsay, France, in 2007.

He was nominated as a tenured Research Scientist at the French National Center for Scientific Research (CNRS), Paris, France, from 2009 to 2013. He is currently the Deputy Vice-Dean of the School of Microelectronics, Beihang University, Beijing, China, where he is also the Director of the Fert Beijing Research Institute and the Beihang-Goertek Joint Microelectronics Institute.

Dr. Zhao was a recipient of the Chinese 1000 Young Plan in 2013 and the prestigious IEEE Guillemin-Cauer Award in 2017.



Rajendra Bishnoi received the Ph.D. degree in computer science from the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2017.

He was a Research Leader of the MRAM Group, Chair of Dependable Nano Computing, KIT, for more than two years. From 2006 to 2012, he was a Design Engineer with Freescale, Noida, India, where he was a part of the Technical Solution Group in memory and SoC flow. He is currently an Assistant Professor with the Computer Engineering Laboratory, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology (TU-Delft), Delft, The Netherlands.

Dr. Bishnoi was a recipient of the EDAA Outstanding Dissertation Award for the year 2017.



Mehdi B. Tahoori (Senior Member, IEEE) received the B.S. degree in computer engineering from the Sharif University of Technology, Tehran, Iran, in 2000, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 2002 and 2003, respectively.

From 2002 to 2003, he was a Research Scientist with the Fujitsu Laboratories of America, Sunnyvale, CA, USA, where he was involved in the area of advanced computer-aided research, engaged in reliability issues in deep-submicrometer mixed-signal very-large-scale integration (VLSI) designs. In 2003, he joined the Electrical and Computer Engineering Department, Northeastern University, Boston, MA, USA, as an Assistant Professor, where he was promoted to the rank of Associate Professor with tenure in 2009. From August 2015 to December 2015, he was a Visiting Professor with the VLSI Design and Education Center (VDEC), The University of Tokyo, Tokyo, Japan. He is currently a Full Professor and the Chair of Dependable Nano-Computing (CDNC) with the Department of Computer Science, Institute of Computer Science & Engineering (ITEC), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.

