Delft University of Technology

Design-for-Test for Intermittent Faults in STT-MRAMs

Yuan, Sicong; Yaldagard, Mohammad Amin; Xun, Hanzhi; Fieback, Moritz; Marinissen, Erik Jan; Kim, Woojin; Rao, Siddharth; Couet, Sebastien; Taouil, Mottaqiallah; Hamdioui, Said

DOI
10.1109/ets61313.2024.10567702

Publication date
2024

Document Version
Final published version

Published in
2024 IEEE European Test Symposium (ETS)

Citation (APA)
Yuan, S., Yaldagard, M. A., Xun, H., Fieback, M., Marinissen, E. J., Kim, W., Rao, S., Couet, S., Taouil, M., & Hamdioui, S. (2024). Design-for-Test for Intermittent Faults in STT-MRAMs. In *2024 IEEE European Test Symposium (ETS)* https://doi.org/10.1109/ets61313.2024.10567702

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.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher
is the copyright holder of this work and the author uses the
Dutch legislation to make this work public.

Design-for-Test for Intermittent Faults in STT-MRAMs

Sicong Yuan*‡ Mohammad Amin Yaldagard* Hanzhi Xun* Moritz Fieback* Erik Jan Marinissen‡
Woojin Kim‡ Siddharth Rao‡ Sebastien Couet‡ Mottaqiallah Taouil*† Said Hamdioui*†
*TU Delft, Delft, The Netherlands † CognitiveIC, Delft, The Netherlands ‡IMEC, Leuven, Belgium
{S.Yuan-4, M.A.Yaldagard, H.Xun, M.C.R.Fieback, M.Taouil, S.Hamdioui}@tudelft.nl
{Erik.Jan.Marinissen, Woojin.Kim, Siddharth.Rao, Sebastien.Couet}@imec.be

2024 IEEE European Test Symposium (ETS) | 979-8-3503-4932-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ETS61313.2024.10567702

***Abstract*—Guaranteeing high-quality test solutions for Spin-Transfer Torque Magnetic RAM (STT-MRAM) is a must to speed up its high-volume production. A high test quality requires maximizing the fault coverage. Detecting permanent faults is relatively simple compared to intermittent faults; the latter are faults (caused by non-environmental conditions) that appear and disappear as a function of time, and are therefore hard to detect. Testing for such faults in STT-MRAMs is even worse considering the Magnetic Tunneling Junction inherent property 'intrinsic switching stochasticity', which results in inevitable random write errors. This paper presents a novel Design-for-Testability (DFT) scheme for detecting intermittent faults in STT-MRAMs; it is based on monitoring the write current. The strength of the write current is inversely correlated to the write error rate; when the write current is smaller than the specification, the device is considered faulty. A reduction in the write current can be caused by any defect in the write path of the memory (e.g., interconnects and contacts). Simulation results based on industrial design show that applying DFT yields a superior coverage of intermittent faults compared to functional test methods, such as march tests.**

***Index Terms*—STT-MRAM, MTJ, test development, design for test, defect, fault, stochasticity.**

I. Introduction

The Spin-Transfer Torque Magnetic RAM (STT-MRAM) is on the way to commercialization, thanks to its different competitive advantages [1]. Since the initial product in 2006, world-leading foundries have entered the market, and STT-MRAMs are applied in a wide range of fields, like embedded systems, computing-in-memory systems, and are considered potential replacements for SRAMs [2–4]. However, further development of STT-MRAM mass production still faces critical challenges, such as its vulnerability to defects [5,6]. Therefore, high-quality and low-cost tests are required.

Work done on testing STT-MRAMs can be classified into two types: 1) testing based on the conventional memory test approach, and 2) testing based on a dedicated approach considering the unique properties of Magnetic Tunneling Junctions (MTJs). The conventional approach just extends the traditional method used for DRAMs and SRAMs [7,8] to STT-MRAMs [9–14]. In this approach, defects are modeled as linear resistors for test development. The approach fails to deliver high-quality test solutions as it does not properly consider the specific working mechanism of MTJs, which e.g., can cause unique defects that cannot be modeled as linear resistance [15,16]. Hence, a dedicated STT-MRAM test approach was developed [17]. In this approach, referred to as Device-Aware Test (DAT) [18], specific compact models for defective MTJs are generated; the models incorporate the impact of physical (manufacturing) defects on the electrical behavior. DAT has been shown to be very powerful in detecting unique defects in STT-MRAMs [17,19]. However, both approaches target mainly defects causing permanent faults; i.e., faults for which the behavior does not change with time. If the faults are not permanent (e.g., intermittent), additional effort is needed to guarantee their detection; these faults may be caused by small/ weak defects such as small resistance opens [20]. The situation becomes even worse when considering the MTJ property 'intrinsic switching stochasticity' [21], which leads to random write errors even in defect-free MTJs [21]. Regularly, Error Corrections Codes (ECC, typically an integral part of the memory) ensure the recovery of write errors as long as the Write Error Rate (WER) is within the defined specification [22]. However, the presence of weak defects combined with the 'intrinsic switching stochasticity" may cause the WER to violate the specification; hence causing intermittent faults. This calls for dedicated test solutions.

This paper presents a Design-for-Test (DFT) methodology dedicated to the detection of intermittent faults due to conventional defects (e.g., small interconnect and contact defects); it is based on the monitoring of write currents. The DFT provides a higher fault coverage for intermittent faults compared to functional test methods (e.g., march tests). In summary, the major contributions of the paper are:

- Introduce the 'intrinsic switching stochasticity' of MTJs.
- Calibrate the MTJ model with WER measurement data.
- Invesitage the limitations of applying function tests (e.g., march tests) on detecting intermittent faults, and estimate the associated escape rate..
- Present, design, and evaluate the DFT to detect intermittent faults, and assess its merits and cost.

The rest of this paper is organized as follows. Section II introduces the basics of STT-MRAM. Section III introduces the MTJ property 'intrinsic switching stochasticity', and its impact. Section IV demonstrates based on a case study and simulation the limitations of function tests to detect intermittent faults, and estimate the probability of escapes. Section V presents, designs, and evaluates the proposed DFT to detect intermittent faults. Section VI concludes this paper.

979-8-3503-4932-0/24/$31.00 ©2024 IEEE

Authorized licensed use limited to: TU Delft Library. Downloaded on July 12,2024 at 10:40:38 UTC from IEEE Xplore. Restrictions apply.

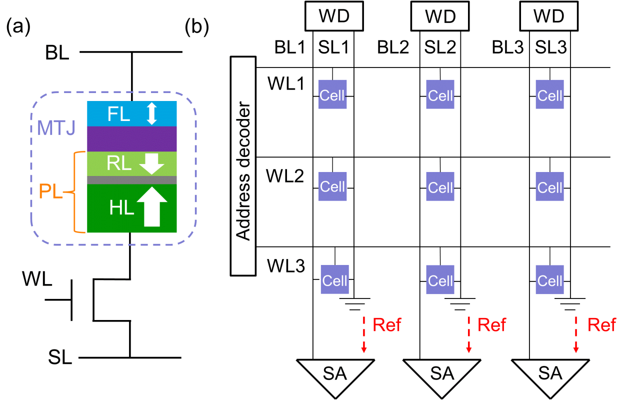

Fig. 1: (a) MTJ stack and 1T-1M cell, (b) STT-MRAM array.

II. BACKGROUND

Fig. 1(a) presents the schematic of the MTJ investigated in this work. The MTJ consists of a thin *Tunnel Barrier (TB)* sandwiched between a *Free Layer (FL)* and a *Pinned Layer (PL)*. The FL is a ferromagnetic layer, whose magnetization can be switched by write operations. The TB is made of MgO. The PL is composed of a *Reference Layer (RL)*, a metal spacer, and a *Hard Layer (HL)*. The MTJ has two stable resistance states depending on the magnetization direction of FL and RL. If the magnetization of the two layers is in parallel, the MTJ resistance is low (i.e., P or 0 state); if in anti-parallel, the MTJ resistance is high (i.e., AP or 1 state). Applying a current can switch the MTJ state between 'P' and 'AP'.

Fig. 1(a) also illustrates the structure of a 1 Transistor - 1 MTJ (1T-1M) cell with three terminals connecting to Bit Line (BL), Source Line (SL), and Word Line (WL). In write/read operations, the voltage of WL selects the cell, and the voltage between BL and SL performs the operation. For example, the $1w0$ operation is performed by connecting the BL to V_{DD} and the SL to the ground, generating a write current switching the MTJ state from AP to P. The tunneling electrons provide STT that switches the FL magnetization from parallel to anti-parallel to that of the RL. On the contrary, a $0w1$ operation refers to offering an opposite current I_{w1} by connecting the BL to the ground and the SL to V_{DD}. The MTJ state is switched from P to AP by the reversed STT. For write operations, a write current I_{w} larger than the critical current I_{c} is necessary to achieve a high write success rate, and the switching time t_{w} is inversely proportional to $I_{\mathrm{w}} - I_{\mathrm{c}}$ [21]. In read operations, a small read current I_{rd} is offered to detect the MTJ resistance.

Fig. 1(b) presents a 3×3 STT-MRAM array with associated peripheral circuits. Cells in the same row share the same WL, and cells in the same column share the same BL and SL. Peripheral circuits consist of e.g., the address decoder, Write Drivers (WD), and Sense Amplifiers (SA). The current flow of write and read operations are presented in Fig. 1(b).

III. INTRINSIC STOCHASTICITY IN MTJ SWITCHING

In this section, we first discuss the physical mechanism of the MTJ property - 'intrinsic switching stochasticity'. Then we present the impact of such this property on the MTJ electrical performance. Finally, the MTJ model will be calibrated (based on data measurements) to appropriately incorporate the impact of intrinsic switching stochasticity; the model will be used in the next section for circuit simulations.

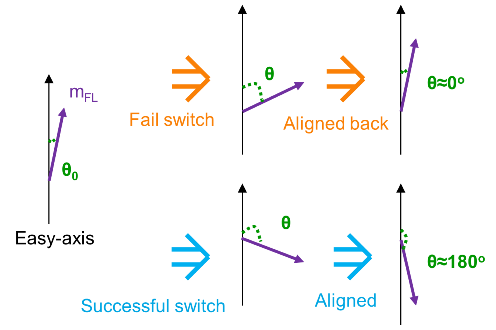

Fig. 2: Mechanism of MTJ switching process.

A. Intrinsic stochasticity mechanism

Fig. 2 presents a simplified physical model of the MTJ switching [21]. The switching process is viewed as the FL magnetization (m_{FL}) rotation. Initially, the angle θ between m_{FL} and easy-axis is around 0°. In write operations, m_{FL} rotates under the spin transfer torque effect, and θ increases. θ will be aligned to the easy-axis with $\theta{\approx}180°$ in successful switching; the MTJ switching time refers to the time that θ alters from $\approx 0°$ to $\approx 180°$. On the other hand, θ will be aligned back to easy-axis with $\theta{\approx}0°$ in case switching fails.

The stochasticity in the switching is attributed to the thermal fluctuation, which affects both the initial MTJ state and the switching process [21]. In the temperature of absolute zero, θ_0 keeps exactly 0°, and the MTJ can never be switched by STT [23]. In room temperature, the initial θ (i.e., θ_0) follows the Maxwell–Boltzmann distribution [23], and the stochasticity of the distribution introduces the varying of MTJ switching time. Besides, the MTJ switching depends both on θ at the write pulse ending and the thermal fluctuation. For example, if write pulses end with $\theta{=}90°$, the switching probability $P_{sw}{=}50\%$; if $\theta{\approx}180°$ at the write pulse ending, $P_{sw}{\approx}100\%$, yet with remaining a small probability of switching failure. This switching stochasticity is an inherent property that cannot be avoided; even in an ideal situation (e.g. temperature keeps constant at $22°C$), the stochasticity still exists. This stochasticity causes cycle-to-cycle random write errors [21].

Note that the MTJ intrinsic stochasticity is completely different from the well-known 'process variability'. The 'process variability' originates from the variations that occur during the fabrication process; it introduces the device-to-device variation [24]. The 'intrinsic switching stochasticity' is the intrinsic property of MTJs; it introduces the cycle-to-cycle variation. This paper only focuses on the 'intrinsic switching stochasticity', to clearly show how it introduces write errors, which are further involved in intermittent faults. In fact, both two issues require to be considered during the test development; This will be our future work.

Authorized licensed use limited to: TU Delft Library. Downloaded on July 12,2024 at 10:40:38 UTC from IEEE Xplore. Restrictions apply.

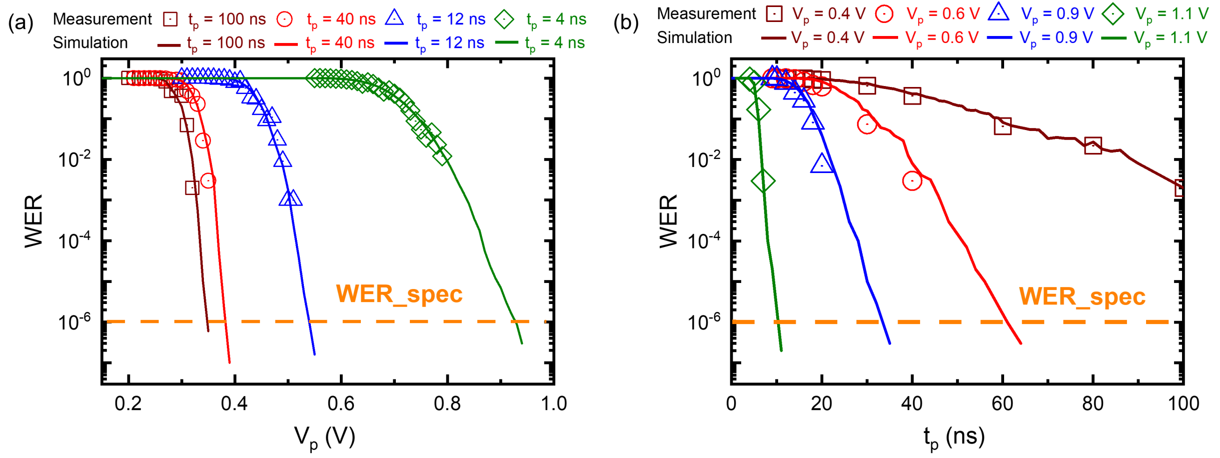

Fig. 3: WER measurement and model calibration for $1w0$, (a) varying V_p with $t_p = 4ns$; (b) varying t_p with $V_p = 1.1V$

B. Impact of 'intrinsic stochasticity' on device performance

The 'intrinsic switching stochasticity' can impact the MTJ electrical performance; it leads to *unavoidable random write errors*. Figures 3(a) and 3(b) present WER for $1w0$ (i.e., apply write 0 to a cell initialized to 1) by varying the pulse height V_p (i.e., the voltage between to FL and HL of MTJs) and pulse width t_p; these are measurements obtained from MTJs manufactured at IMEC, with the structure in Fig. 1(a) and a Critical Diameter CD=60 nm. As the figure shows, the WER can never be reduced to 0 regardless of the value V_p or t_p. Hence, additional efforts are needed to make the STT-MRAM design fault tolerance; e.g., ECC is necessary for high-quality MRAM systems [22]. However, the capability of the fault tolerance scheme is always limited. Therefore, it is a common practice to define a WER design spec (WER_spec); i.e., the tolerated WER according to the design specification. If the WER of the MTJ in one cell is lower than the WER_spec, then the used fault tolerance scheme can 'tolerate' the possible write errors; hence STT-MRAM cell is seen as fault-free. For industrial designs, it is very common to have WER_spec=10^{-6} (i.e., 1 write error per million writing operations) [25], which is also applied in this work.

C. Model calibration with WER measurement data

The MTJ switching mode be classified into three regimes according to the switching time (t_s): the precessional regime for t_s<10 ns, the dynamic regime for 10 ns<t_s<100 ns, and the thermal activation regime for t_s>100 ns [21]. For accurate simulation, the models of the three switching regimes of the MTJ should integrate the impact of the intrinsic stochasticity. We use the MTJ compact model presented in [26], and augment it with this impact. For example, the relationship of WER and the write pulses in the precessional regime is presented as follows [21]:

$$WER = 1 - \exp\left(-\frac{\Delta\pi^2 j}{4i\left(\exp\left(2\left(t_p/t_0\right)j\right)-1\right)}\right), \quad where \quad i = I_{MTJ}/I_c, \quad and \quad j = i-1 \tag{1}$$

Here, t_0 is the characteristic relaxation time, I_{MTJ} is the current through MTJs, I_c is the critical current, and Δ is the thermal stability [21]. We applied a distribution following Eq. 1 to the average switching time in the precessional regime in the used compact model. A similar process is conducted to the dynamic regime and thermal activation regime following the equations in [21]. The three models of the three switching regimes were calibrated using the same WER measurement data in Fig. 3. The figure shows both the measured data as well as the simulated results; it clearly shows that the calibrated models result in perfect fitting of the measured results.

IV. Limitation of Function Memory Tests

In this section, we will show the limitations of applying function test methods (e.g., march tests) to detect intermittent faults in STT-MRAMs, and analytically estimate the escape rate. First, we briefly present the simulation setup used in this work. Then, through an example (defect injection and circuit simulation for a selected defect - case study), we illustrate how march tests fail to guarantee the detection of intermittent faults. Finally, we estimate the escape rate.

A. Simulation set-up

The circuit simulations are performed with the 3×3 STT-MRAM array, as shown in Fig. 1(b). The augmented MTJ model of [26] is used and calibrated as explained in the previous section. Cadence Spectre is adopted for circuit-level simulations, and Predictive Technology Model (PTM) 40-nm transistor library is applied [27]. In our design spec, the used supply voltage is V_{DD}=1.8 V, and the write pulse width t_p is 4 ns. The circuit is simulated and verified for the correctness of the regular write/read operations being xwy and xrx where 'w' denotes the write operation, 'r' denotes the read operation, and $x, y \in \{0, 1\}$.

B. Illustrated example

We use a contact defect as an example to illustrate that it can cause intermittent faults and that a functional memory test will fail to guarantee the detection of this defect. We will follow the

Authorized licensed use limited to: TU Delft Library. Downloaded on July 12,2024 at 10:40:38 UTC from IEEE Xplore. Restrictions apply.

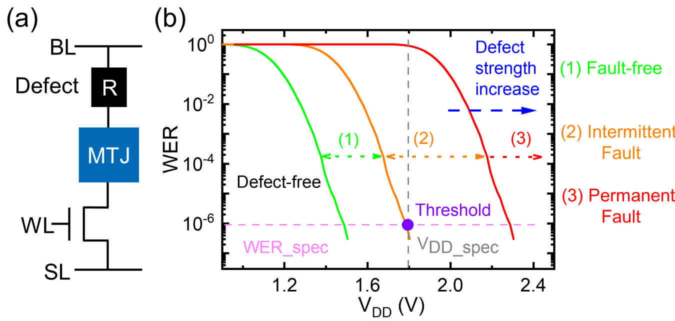

Fig. 4: (a) Defect injection; (b) Example of fault modeling.

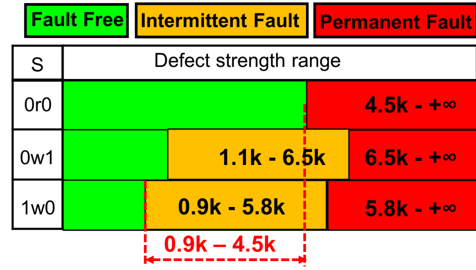

Fig. 5: Fault map.

traditional memory test approach which consists of three steps: defect modeling, fault modeling, and test generation [28].

1) Defect modeling: We model a contact defect as a linear resistor, and we inject it between BL and the MTJ, as presented in Fig. 4(q). The contact defect may physically originate from an unexpectedly high contact resistivity or the connection breaking [29]. The defect strength is modeled as the resistance value R_d, where $0<R_d \leq +\infty$.

2) Fault modeling: In this step, we vary R_d, apply sequences $S_i \in \{0w0, 0w1, 1w0, 1w1, 0r0, 1r1\}$ on the defective STT-MRAM cell while using the simulation setup (Fig. 1), and observe the way the defect manifests itself at the functional behavior. Regular fault modeling only performs the write operation once, and checks if the MTJ is switched successfully. This method misses intermittent faults due to the unavoidable random write errors. In order to capture the intermittent faults (if any) and the impact of WER, we perform each of the four *write* sequences 10 million times, while varying R_d and V_{DD}. We inspect (using Spice simulation) the memory state after each operation to identify if the performed (write) operations fail or pass. As an example of obtained results, Fig. 4(b) shows WER when performing $1w0$ operation with a constant t_p=4 ns and varying V_{DD} for three defect sizes:

- Defect-free case where $R_d = 0$. The obtained results are shown in the green line. In this case, WER < WER_spec (i.e., 10^{-6}) considering our design spec with the supply voltage V_{DD} =1.8 V.
- Borderline for fault-free case, where R_d=0.9 kΩ and at which the obtained WER = WER_spec considering our design spec (e.g., V_{DD}=1.8 V). These results are shown in the orange line. As Fig. 4(b) shows, when $0\,\Omega \leq R_d \leq 0.9\,\text{k}\Omega$ (region 1 in the figure), $WER<WER_spec$; and therefore the device is assumed to be fault free.
- Borderline for intermittent case, where R_d=5.8 kΩ, and at which the WER reaches 1 (i.e., permanent fault) considering our design spec. The obtained results are shown in the red line. As Fig. 4(b) shows, when $0.9\,\text{k}\Omega \leq R_d \leq 5.8\,\text{k}\Omega$) (region 2 in the figure), $WER_spec<WER<100\%$; hence, the defect causes intermittent faults.

It is worth noting that for all $R_d>5.8\,\text{k}\Omega$ (region 3 in the figure), WER=100%; hence, permanent faults are sensitized (considering our design spec such as V_{DD}=1.8 V). This is region 3 in the figure.

Fig. 5 shows the complete fault map for the considered defect in Fig. 4(a); only applied sequences ('S') which sensitized faults, together with defect strength range are shown. In addition, the map also gives which type of faults are sensitized. The 'red' boxes indicate ranges for permanent faults, the 'orange' boxes indicate ranges for intermittent faults, and the 'green' boxes indicate ranges of fault-free case (i.e., $WER<WER_spec$).

Fig. 5 reveals that when $0.9\,\text{k}\Omega<R_d<4.5\,\text{k}\Omega$, only intermittent faults are sensitized. When $4.5\,\text{k}\Omega<R_d<6.5\,\text{k}\Omega$, the defect sensitizes either intermittent or permanent faults, depending on the sensitizing sequence; i.e., $0w1$ and $1w0$ sensitize intermittent faults while $0r0$ sensitizes permanent faults. During $0r0$, the sensing current by the SA will be much smaller than expected due to the resistive defect in the read path (BL-MTJ-SL; see Fig. 4); the SA will return 1 instead of 0. Hence, the application of only read zero operation will guarantee the detection of the defect in this range. When $R_d>6.5\,\text{k}\Omega$, the defect sensitizes permanent faults; performing any of the three sequences ($0r0$, $0w1$, $1w0$) can guarantee the sensitization of the faults.

3) Test generation: The previous example shows that the defect sensitizes both intermittent as well as permanent faults. We need to guarantee the detection of both types of faults. Detecting permanent faults needs a simple march test, such as $\{w0; r0\}$; here we use the march notation as defined in [30]. However, the detection of intermittent faults requires additional effort. One way to deal with it is to apply a march test multiple times with the hope that the march algorithm will fail for at least one write operation. For example, the following march test can be applied:

$$\{\Updownarrow(w1); \Updownarrow(w0, r0, w1, r1)^n\} \tag{2}$$

This first element initializes the memory to state '1', while the second element repeats the four operations '$w0$, $r0$, $w1$, $r1$' for 'n' times. Given the random write errors, it is very hard, if not impossible, to guarantee the detection of intermittent faults with a reasonable n. The next section will show that even with very high values of n, such march test ends in a high escape rate of devices suffering from intermittent faults.

Authorized licensed use limited to: TU Delft Library. Downloaded on July 12,2024 at 10:40:38 UTC from IEEE Xplore. Restrictions apply.

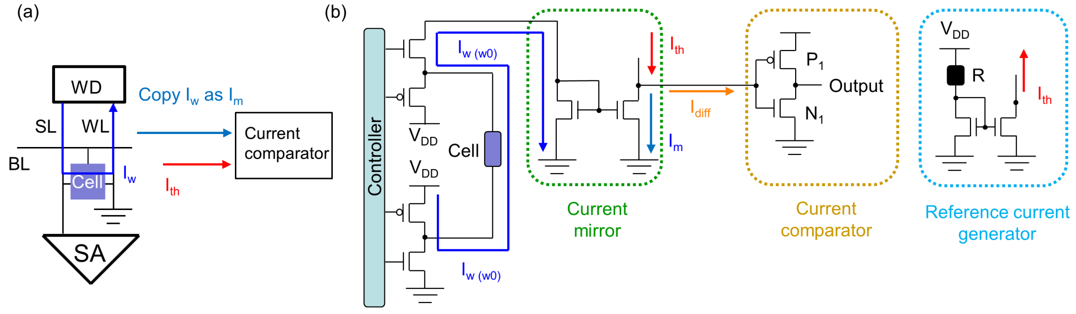

Fig. 6: (a) DFT methodology; (b) Circuit design of the DFT.

C. Escape rate calculation and verification

Let us assume that we apply $1w0$ for n times to a defective cell, and the defect of strength R_d=R_1 causes an intermittent fault with WER=$WER(R_d)$. Then the Fault Sensitisation Probability (FSP) can be estimated as:

$$FSP(R_1) = 1 - (1 - WER(R_1))^n \quad (3)$$

We assume that the defect causes intermittent faults for the range $R_1 \leq R_d < R_2$ (e.g., borderline cases in Fig. 4(b)), then the total FSP for all intermittent faults due to the defect can be obtained by applying an integration as:

$$FSP = \frac{\int_{R_1}^{R_2} 1 - (1 - WER(R_d))^n dR_d}{\int_{R_1}^{R_2} 1 dR_d} \quad (4)$$

$WER(R_d)$ is the function of how WER changes with R_d. For example, for our case study, R_1=0.9 kΩ and R_2=5.8 kΩ for $1w0$ can be extrated from Fig. 4(b). The calculation will give FSP=2.3% for n=1.

Next, we will calculate Fault Detection Probability (FDP) assuming the application of Eq. 2; the focus here is only on intermittent faults. We assume the same defect of strength R_d=R_1, which causes faults with WER=$WER_{0w1}(R_1)$ for $0w1$, and $WER_{1w0}(R_1)$ for $1w0$, then:

$$FDP(R_1) = 1 - [(1 - WER_{1w0}(R_1))(1 - WER_{0w1}(R_1))]^n \quad (5)$$

Similar to Eq. 4, FPD for the whole range of the defect strength being only able to sensitize intermittent faults for the targeted defect can be derived:

$$FDP = \frac{\int_{0.9K}^{4.5K} 1 - [(1 - WER_{1w0}(R))(1 - WER_{0w1}(R))]^n dR}{\int_{0.9K}^{4.5K} 1 dR} \quad (6)$$

The Escape Probability (EP) of the above defect can then be calculated as: EP=1-FDP. If we apply the march test of Eq. 2, EP=99.79% for n=1, and EP=87% for n=1000. Clearly, even though with very high n, EP is still very high.

Finally, we performed circuit simulations to verify the above estimation. E.g., for the same defect, we selected 15 different values of R_d within the range $[0.9\,\text{k}\Omega, 4.5\,\text{k}\Omega]$. The test of Eq. 2 is simulated in the presence of each of the 15 R_d values; and for each R_d values, we perform the test for n=1000 times. We assumed to have 1000 similar defective chips of each of the 15 R_d values (resulting in the representation of 15000 defective chips in total); hence we repeated the process of applying the test with n=1000 for 15000 times. The fault is assumed to be detected if any read operation within the test returns a wrong value. The simulation results show that only in 2133 simulations (among the total of $15 * 1000$) the fault was detected (about 14.2%), which is quite in line with our estimation.

V. DFT Methodology and Implementation

A. DFT concept

Fig. 6(a) presents the concept of the proposed DFT, which is based on: a) copying/mirroring the write current I_w to I_m, and b) comparing the copied current with a reference current I_{th} which is the write current derived at the 'borderline for fault-free case' considering our design spec (i.e., the write current corresponding to the purple dot in Fig. 4(b)). I_{th} is a constant value, since the relationship between WER and I_w is constant for defect-free MTJs (see Sec. III). Any defect leading to I_w<I_{th} suggests the cell is faulty.

It is important to note that there are four options to select I_{th}; i.e., based on I_w for $0w1$, $1w0$, $0w0$, or $1w1$. Selecting I_{th} for $0w1$ or $1w0$ will not do the targeted job; because of the intermittent behavior of the cell, the MTJ may or may not switch, hence I_w will change accordingly. As a consequence, comparing the changing I_w with a reference Ith will not work. In addition, we need to select the comparison of I_w for $1w1$ or for $0w0$; the idea is to select the case resulting in the maximal current deviation ΔI_w (in the presence of the defect) as compared with I_{th}. The differences in the current in the two cases can be given as:

$$\Delta I_w(0w0) \propto \frac{V_{DD}}{R_{M_low}} - \frac{V_{DD}}{R_{M_low} + R_d} \quad (7)$$

$$\Delta I_w(1w1) \propto \frac{V_{DD}}{R_{M_high}} - \frac{V_{DD}}{R_{M_high} + R_d} \quad (8)$$

Here, R_{M_low} and R_{M_high} refer to the low resistance, respectively, high resistance states of the MTJ cell. The analysis of the above equations reveals that the max current deviation is always obtained for the case of $0w0$ for lower R_d; hence this will be used for the DFT. The cell should be initialized to 0 before applying DFT.

B. DFT implantation

Fig. 6(b) presents the DFT circuit design consisting of three parts: the current mirror, the current comparator, and the reference current generator. At time 0 (i.e., the start of

Authorized licensed use limited to: TU Delft Library. Downloaded on July 12,2024 at 10:40:38 UTC from IEEE Xplore. Restrictions apply.

the testing), the current mirror copies I_w to I_m, and the reference current generator provides I_{th}. Then $I_{diff}{=}I_{th}{-}I_m$ is generated; it can be either positive of negative. It will drive the input of the current comparator (i.e., inverter consists of MOSFETs P_1 and N_1). If $I_m{>}I_{th}$, then $I_{diff}{<}\ 0$, and the inverter reports '0' at 'Output' (fault-free). Conversely, if $I_m{<}I_{th}$, then $I_{diff}{>}\ 0$, and the 'Output' of the inverter reports '1' (fault detected).

C. Verification and evaluation

The DFT is simulated while injecting different R_d values for our case study. The results confirm the superiority of the DFT in guaranteeing the detection of the targeted intermittent faults. Moreover, the proposed DFT also detects all permanent faults (R_d $>4.5\,\mathrm{k\Omega}$). Hence, it covers both permanent and intermittent faults!

Clearly, the proposed DFT is superior as compared with any functional march test (even with repeating march elements). In fact, the fault coverage of the march test is just a subset of that of DFT. Hence, the proposed DFT offers an outstanding alternative to increase fault coverage by considering intermittent faults. Although we cannot claim that all intermittent faults in the STT-MRAM memory can be detected, the proposed DFT clearly has an added value as it covers also intermittent faults and permanent faults that are caused by defects in *the write path*. Similar philosophy can be used to develop another DFT to cover intermittent faults caused by defects in the read path.

Nevertheless, the DFT comes at some additional cost; it requires only 7 MOSFETs per column, and some minor design effort. The larger the size of the memory array (in terms of rows), the smaller the overall area overhead.

VI. Conclusion

This paper sets up a step toward the improvement of STT-MRAM outgoing product quality by improving the fault coverage of intermittent faults. Although permanent faults could be dominant as compared with intermittent faults, missing the fraction of the fault coverage related to intermittent faults may end in high escape rates. This paper proves the occurrence of such faults in STT-MRAM, shows their potential contribution to a high escape rate, and proposes a powerful scheme to target them. The paper also demonstrates that intermittent faults in STT-MRAM need special attention due to their inherent 'intrinsic switching stochasticity", making the situation even worse.

Our future work will focus on how to include both 'process variability' and 'intrinsic switching stochasticity' during the test development. Two methods can be considered: 1) Use the worst-case scenario for the reference current (e.g., 3 or 6 sigma design); 2) Apply a multi-reference design to the DFT circuit (post calibration).

Acknowledgements

This work is supported by IMEC's Industrial Affiliation Program on STT-MRAM devices.

References

[1] Y. Huai *et al.*, "Spin-transfer torque MRAM (STT-MRAM): Challenges and prospects," *AAPPS bulletin*, vol. 18, pp. 33–40, 2008.

[2] D. Edelstein *et al.*, "A 14 nm embedded stt-mram cmos technology," in *IEDM*, 2020, pp. 11–5.

[3] A. Singh *et al.*, "Cim-based robust logic accelerator using 28 nm stt-mram characterization chip tape-out," in *AICAS*, 2022, pp. 451–454.

[4] S. Rao *et al.*, "STT-MRAM array performance improvement through optimization of Ion Beam Etch and MTJ for Last-Level Cache application," in *IMW*, 2021, pp. 1–4.

[5] E.I. Vatajelu *et al.*, "Challenges and solutions in emerging memory testing," *IEEE Trans. Emerg. Topics Comput.*, vol. 7, pp. 493–506, 2017.

[6] L. Wu *et al.*, "Survey on STT-MRAM testing: Failure mechanisms, fault models, and tests," *arXiv preprint arXiv:2001.05463*, 2020.

[7] S. Hamdioui *et al.*, "An experimental analysis of spot defects in SRAMs: realistic fault models and tests," in *ATS*, 2000.

[8] A.J. van de Goor *et al.*, "Disturb neighborhood pattern sensitive fault," in *VTS*, 1997, pp. 37–45.

[9] J. Azevedo *et al.*, "A complete resistive-open defect analysis for thermally assisted switching MRAMs," in *IEEE T. VLSI Syst.*, vol. 22, no. 11. IEEE, 2014, pp. 2326–2335.

[10] S.M. Nair *et al.*, "Defect injection, fault modeling and test algorithm generation methodology for STT-MRAM," in *ITC*, 2018, pp. 1–10.

[11] A. Chintaluri *et al.*, "A model study of defects and faults in embedded spin transfer torque (STT) MRAM arrays," in *ATS*, 2015, pp. 187–192.

[12] C. Münch *et al.*, "MBIST-based Trim-Search Test Time Reduction for STT-MRAM," in *VTS*, 2022, pp. 1–7.

[13] I. Yoon *et al.*, "Modeling and analysis of magnetic field induced coupling on embedded STT-MRAM arrays," *TCAD*, vol. 37, pp. 337–349, 2017.

[14] S.M. Nair *et al.*, "VAET-STT: Variation aware STT-MRAM analysis and design space exploration tool," *TCAD*, vol. 37, pp. 1396–1407, 2017.

[15] W. Kim *et al.*, "Experimental Observation of Back-Hopping With Reference Layer Flipping by High-Voltage Pulse in Perpendicular Magnetic Tunnel Junctions," *IEEE Trans. Magn.*, vol. 52, pp. 1–4, 2016.

[16] J.H. Lim *et al.*, "Origins and signatures of tail bit failures in ultrathin mgo based stt-mram," in *IRPS*, 2020, pp. 1–5.

[17] L. Wu *et al.*, "Electrical Modeling of STT-MRAM Defects," in *ITC*, 2018, pp. 1–10.

[18] M. Taouil *et al.*, "Device Aware Test for Memory Units," 2021, in European patent EP4026128A1.

[19] L. Wu *et al.*, "Characterization and fault modeling of intermediate state defects in stt-mram," in *2021 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2021, pp. 1717–1722.

[20] D. Zhou *et al.*, "Review on diagnosis techniques for intermittent faults in dynamic systems," *IEEE Trans. Ind. Electron.*, vol. 67, pp. 2337–2347, 2019.

[21] A.V. Khvalkovskiy *et al.*, "Basic principles of STT-MRAM cell operation in memory arrays," *J. Phys. D: Appl. Phys*, vol. 46, p. 139601, 2013.

[22] X. Guo *et al.*, "Sanitizer: Mitigating the impact of expensive ECC checks on STT-MRAM based main memories," *IEEE Trans. Comput.*, vol. 67, pp. 847–860, 2017.

[23] T. Kawahara *et al.*, "Spin-transfer torque RAM technology: Review and prospect," *Microelectron. Reliab.*, vol. 52, pp. 613–627, 2012.

[24] S. Taghipour *et al.*, "CD-DFT: A current-difference design-for-testability to detect short defects of STT-MRAM under process variations," *IEEE Trans. Device Mat. Rel.*, vol. 21, pp. 436–443, 2021.

[25] J. Kan *et al.*, "Systematic validation of 2x nm diameter perpendicular MTJ arrays and MgO barrier for sub-10 nm embedded STT-MRAM with practically unlimited endurance," in *IEDM*, 2016, pp. 27–4.

[26] Y. Wang *et al.*, "Compact model of magnetic tunnel junction with stochastic spin transfer torque switching for reliability analyses," *Microelectron. Reliab.*, vol. 54, pp. 1774–1778, 2014.

[27] ASU, "Predictive Technology Model (PTM)," 2012.

[28] S. Hamdioui *et al.*, "Memory fault modeling trends: A case study," *J. Electron. Testing*, vol. 20, pp. 245–255, 2004.

[29] A.P. Jacob *et al.*, "Scaling challenges for advanced CMOS devices," *Int. J. High Speed Electron. Syst.*, vol. 26, p. 1740001, 2017.

[30] A.J. Van de Goor, *Testing semiconductor memories: theory and practice*. John Wiley & Sons, Inc., 1991.

Authorized licensed use limited to: TU Delft Library. Downloaded on July 12,2024 at 10:40:38 UTC from IEEE Xplore. Restrictions apply.