Nonpersistent Errors Optimization in Spin-MOS Logic and Storage Circuitry

Peiyuan Wang[1], Xiaobin Wang[2], Yaojun Zhang[1], Hai Li[3], Steven P. Levitan[1], and Yiran Chen[1]

[1]Electrical and Computer Engineering Department, University of Pittsburgh, Pittsburgh, PA 15261 USA
[2]Seagate Technology, Bloomington, MN 55435 USA
[3]Department of Electrical and Computer Engineering, Polytechnic Institute of New York University, Brooklyn, NY 11201 USA

By combining the flexibility of MOS logic and the nonvolatility of spintronic devices, Spin-MOS logic and storage circuitries offer a promising approach to implement a highly integrated, power-efficient, and nonvolatile computing and storage systems. Besides the persistent errors due to process variations, however, the functional correctness of Spin-MOS circuitries suffers from additional nonpersistent error that incurred by the randomness of spintronic device operations, i.e., thermal fluctuations. In this work, we quantitatively investigate the impacts of the thermal fluctuations on the operations of two typical Spin-MOS circuitries: one transistor and one magnetic tunnel junction (1T1J) spin-transfer torque random access memory (STT-RAM) cell and a nonvolatile flip–flop design. The possible design techniques to reduce thermal incurred nonpersistent error rate are also discussed. Our experimental results show that the optimization of nonpersistent and persistent errors are closely entangled with each other and should be conducted from both circuit design and magnetic device engineering perspectives simultaneously.

Index Terms—**Nonpersistent, process variation, Spin-MOS.**

I. Introduction

SPIN torque induced magnetization switching in magnetic tunneling junctions (MTJs) is the fundamental of modern spintronic memory, which features nanosecond access time, high programming endurance, nonvolatility, and zero standby power. By combining the flexibility of MOS logic and the nonvolatility of spintronic devices, Spin-MOS logic and storage circuitries make it possible to implement high-density, low-power, nonvolatile, and robust computing and storage systems. Besides the well-known spin-transfer torque random access memory (STT-RAM), spintronic devices have been also used in timing sequential circuitries and simple logics, such as latches [1] and lookup tables [2]. Moreover, prior arts show that compared to the conventional MOS logics whose functionalities are based on the operation of electrical charge, Spin-MOS circuitries are more resilient to soft errors, which are primarily generated by the Alpha particle emissions from chip packaging materials [3].

The function errors of a circuit can be categorized as nonpersistent and persistent ones [4]. An error is persistent if it happens deterministically and can be repeated after the chip is fabricated, such as the errors introduced by process variations. The nonpersistent errors include those introduced by soft-errors in CMOS circuitries or by thermal fluctuations in Spin-MOS circuitries. In Spin-MOS circuitries, thermal-induced nonpersistent errors demand specific optimization design techniques even soft-errors are eliminated. In this work, we quantitatively investigate the impacts of the thermal fluctuations on the operations of two typical Spin-MOS circuitries: 1T1J spin-transfer torque random access memory (STT-RAM) cell and a nonvolatile flip–flop design. On top of it, we exploit the possibility to minimize the thermal incurred nonpersistent error rate, while taking into account the adverse impacts on the persistent errors.

The rest of this paper is organized as follows. Sections II-A and II-B give a preliminary on the persistent and nonpersistent errors in Spin-MOS circuitries by using a STT-RAM cell as the example. Section II-C depicts the quantitative analysis on the impacts of nonpersistent error on the operation of Spin-MOS circuitries and its optimizations. Section II-D discusses the tradeoff between persistent and nonpersistent errors. In Section III, a more complicated case studya nonvolatile flip-flop is presented. Finally the work is concluded in Section IV.

II. Case Study on STT-RAM Cells

A. Persistent Errors in Spin-MOS Circuitries

Fig. 1 shows the popular 1T1J (one-transistor-one-MTJ) STT-RAM cell design, where a MTJ is connected to a NMOS transistor. The MTJ resistance can be changed between the high and the low state under a polarized switching current. It is well-known that the switching time of a MTJ is determined by the switching current: the increases on the switching time leads to the reduction on the switching current. Moreover, when the MTJ switching time is under 10 ns, the further scaling of switching time will cause the exponential increase in switching current, as shown in Fig. 2. Here, the switching current and time are achieved based on MTJ with a 45×90 nm ellipse shape.

Due to the process variation, e.g., the variations of NMOS transistor channel width (W), channel length (L), and threshold voltage $(V_{\rm t})$, the current provided by the NMOS transistor to the MTJ varies from memory cell to cell or even from chip to chip associated with the variations of MTJ switching time. Because the parameters of NMOS transistors are fixed after the chip is fabricated, the corresponding errors incurred by the transistor variations, i.e., MTJ fails to switch within the applied write pulse

Manuscript received February 21, 2011; revised April 23, 2011; accepted May 01, 2011. Date of current version September 23, 2011. Corresponding author: P. Wang (e-mail: wap15@pitt.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMAG.2011.2153838

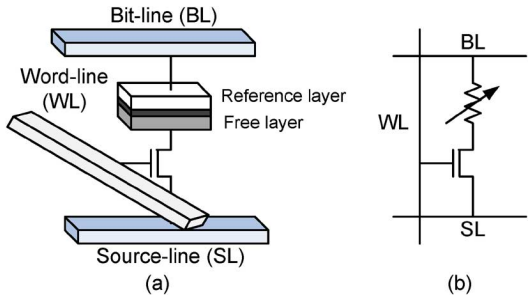

Fig. 1. 1T1J STT-RAM cell. (a) Cell view. (b) Equivalent schematic.

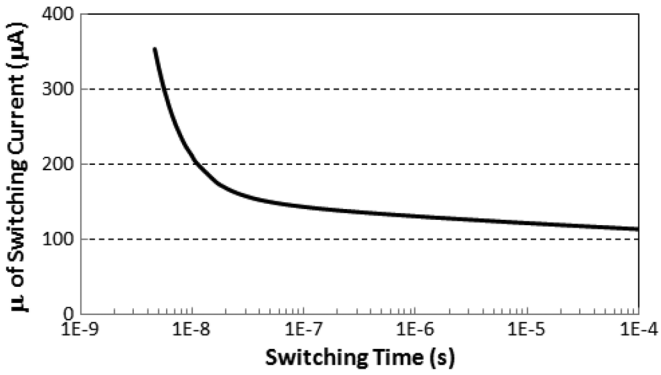

Fig. 2. Relationship between MTJ switching time and switching current.

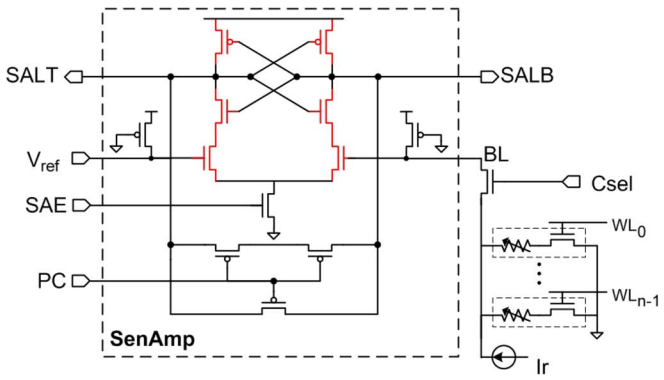

Fig. 3. Conceptual sense amplifier design.

width, are persistent. Similarly, the MTJ geometry and resistance variations, which may cause the MTJ driving current shifts by changing the bias conditions of NMOS transistor, are also fixed after the chip is fabricated. Therefore, the corresponding errors are also persistent.

Another important persistent error in STT-RAM design is the fault sensing due to the device mismatch in sense amplifier and/or the small sense margin. During the read operation, a read current $I_{\rm r}$ is injected into STT-RAM cell and generate the corresponding bitline voltage $V_{\rm BL}$. Then, the MTJ resistance state can be obtained by comparing $V_{\rm BL}$ to a reference voltage $V_{\rm ref}$ in the sense amplifier (SenAmp), as shown in Fig. 3. However, if the sizes and the threshold voltages of the six MOS transistors (highlighted in RED) deviate from their designed values too much, or the difference between $V_{\rm BL}$ and $V_{\rm ref}$ is too small, SenAmp may give a false result.

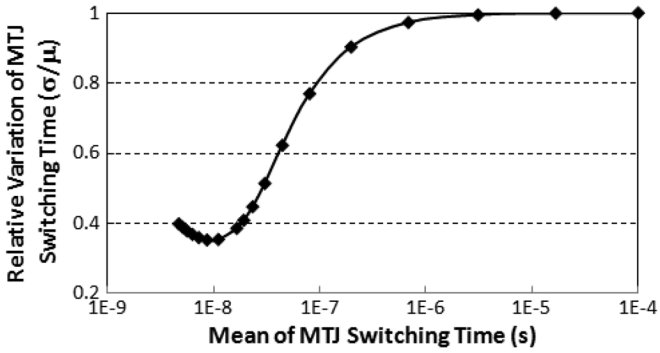

Fig. 4. Variations of MTJ switching time due to thermal fluctuations.

B. Nonpersistent Errors in Spin-MOS Circuitries

In a Spin-MOS circuitry, there are two major nonpersistent errors, which occur in write or read operations, respectively.

The first one is due to the thermal fluctuation in the write operation of STT-RAM cells. When a MTJ works in a long time region (>10 ns), the thermal fluctuation is dominated by the thermal component of internal energy; when MTJ works in sub-10 ns region, the thermal fluctuation is dominated by the thermally activated initial angle of procession [5]. The existence of thermal fluctuation causes the deviation of the MTJ switching time from its nominal value. Also, following the increases in MTJ switching current, the ratio between the standard deviation and the mean of MTJ switching time decreases first, mainly due to the increased impact of spin-torque on MTJ switching. Then, it increases again after the mean of MTJ switching time enters sub-10 ns region and the thermally activated initial angle of procession dominates, as shown in Fig. 4. If a MTJ cannot switch by the end of write pulse width, an error will be generated.

The second nonpersistent error is read disturbance, which denotes the undesired MTJ switching during the read operation. In [6], it is pointed out that the disturbance probability $(\Pr_{\rm dis})$ of a MTJ at a read current of $I_{\rm R}$ can be expressed as

$$\Pr_{\rm dis} = 1 - \exp\left\{-\frac{t}{\tau}\exp\left[-\Delta\left(1-\frac{I_{\rm R}}{I_{\rm C}}\right)\right]\right\} \qquad (1)$$

which has been proven in [7]. Here, t is the read current pulse width. Δ is the magnetic memorizing energy without applying any current or magnetic fields. τ is the inverse of the attempt frequency. $I_{\rm C}$ is the critical switching current, which is the minimum current amplitude to switch the MTJ resistance with a write pulse width of τ. Usually, the read current pulse width is fixed by the timing control circuit. Therefore, $\Pr_{\rm dis}$ is mainly determined by the read current amplitude. Fig. 5 shows the read disturbance probability of the simulated MTJ with 10 ns read pulse width under various read current. Compared to the nonpersistent error resulted by the thermal fluctuation in the write operation, the impact of read disturbance is much smaller.

C. Minimizing Nonpersistent Errors

The straightforward way to minimize the nonpersistent errors in the write operation of STT-RAM cell is increasing the switching current, or sizing up the NMOS transistor. As shown in Figs. 2 and 4, increasing switching current (e.g., by increasing

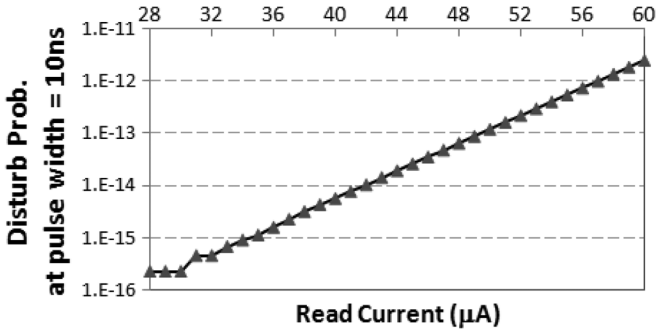

Fig. 5. The variations of MTJ read disturbance probability when read current amplitude changes.

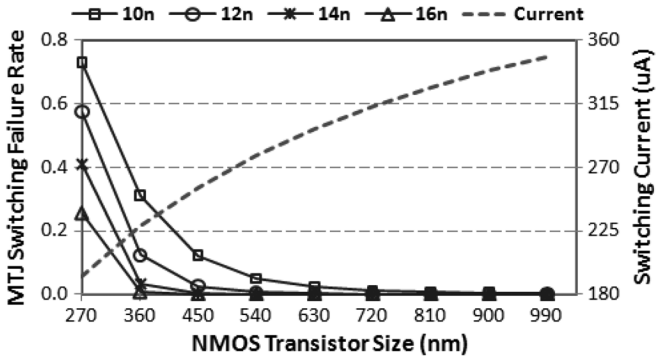

Fig. 6. The nonpersistent failure rate of MTJ as the transistor size varies.

NMOS transistor size) can produce a tighter distribution of the required switching time by reducing both the mean and the standard deviation of switching time.

Fig. 6 shows the simulated rate that MTJ fails to switch within the given switching pulse width by varying NMOS transistor size in the range from 270 to 990 nm. At each node, 1000 Monte Carlo simulations were conducted with the thermal fluctuation in consideration. Here, the switching pulse width increases from 10 to 16 ns with a step of 2 ns.

The simulation results show that increasing the size of the NMOS transistor, and hence, increasing switching current in a STT-RAM cell can effectively reduces the MTJ switching failure rate when NMOS is small. However, further increasing NMOS transistor size (e.g., when NMOS is bigger than 360 nm at 16 ns switching pulse width) does not improve MTJ switching failure rate much. Although the mean of MTJ switching time becomes below 10 ns when the NMOS channel width is larger than 360 nm, significant timing error rate can still be observed due to the variations of MTJ switching performance. In the practice of STT-RAM cell design, the target write error rate is usually predetermined by the memory specification.

The minimization of read disturbance probability is usually achieved by controlling the read current amplitude through clamping magnetic field which may be applied to enhance the MTJ stability during the read operation [8].

D. Tradeoff Between Persistent and Nonpersistent Errors

In STT-RAM designs, the amplitude of read current is usually controlled by a global read driver. We note that the sense margin of a STT-RAM cell ΔV is proportional to $I_{\mathrm{R}} \cdot \Delta V$, where ΔV is the difference between the high- and the low-resistance states of the MTJ. Thus, reducing read current I_{R} minimizes read disturbance probability, while simultaneously increasing the sensing error rate due to the degraded sense margin.

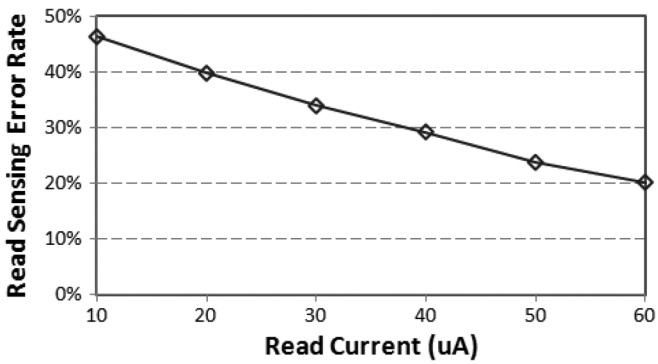

Fig. 7. Tradeoff between read disturbance probability and sensing errors.

The Monte Carlo simulation results are shown in Fig. 7. Here, we assume the standard deviations of NMOS transistor channel width and length are 5% of their nominal values and the standard deviation of $\mathrm{V_t}$ for minimum size is 33 mV. The variations of MTJ read disturbance probability with various read current amplitudes can reference Fig. 5. Our simulation shows that the sense margin degradation dominates within the given read current range $(< 60\,\mu\mathrm{A})$. For the given MTJ device, the read disturbance increases sharply after the read current exceed this value. Therefore, the nonpersistent errors due to the read disturbance start dominating, which is hard to control in design.

III. Case Study on Nonvolatile Flip–Flop

The nonvolatility of MTJ devices are also utilized in the other circuit component designs, i.e., Flip–Flop. Fig. 8(a) shows a recently proposed nonvolatile flip–flop design where two MTJs are embedded into the traditional flip–flop design with opposite stack structures [1]. In the normal operation, the whole flip–flop works as the conventional flip–flop. When the circuit is entering standby or power down mode, "EN" signal is raised and the stored value is written into the two MTJ by a current whose direction is controlled by the stored value.

One disadvantage of this design is that the write path always includes two NMOS and two MTJs. The large voltage drops across the MTJs degrade the driving ability of MOS transistors by reducing the voltage difference between their gate and source (V_{gs}).

In this work, we proposed a new flip–flop design with separated write paths of the two MTJs to overcome the above disadvantage, as shown in Fig. 8(b). Each MTJ has its own PMOS-NMOS transistor pair to supply the switching current during the write operation. Obviously, the size of PMOS-NMOS transistor pair must be sufficiently large to minimize the nonpersistent timing errors during the write operations.

When the nonvolatile flip–flop wakes up from the standby mode, the difference between the resistances of two MTJ is sensed. Similar as the SenAmp in STT-RAM designs, the device mismatch among the cross coupled inverters (M1–M4) may cause false sensing when the generated voltage difference

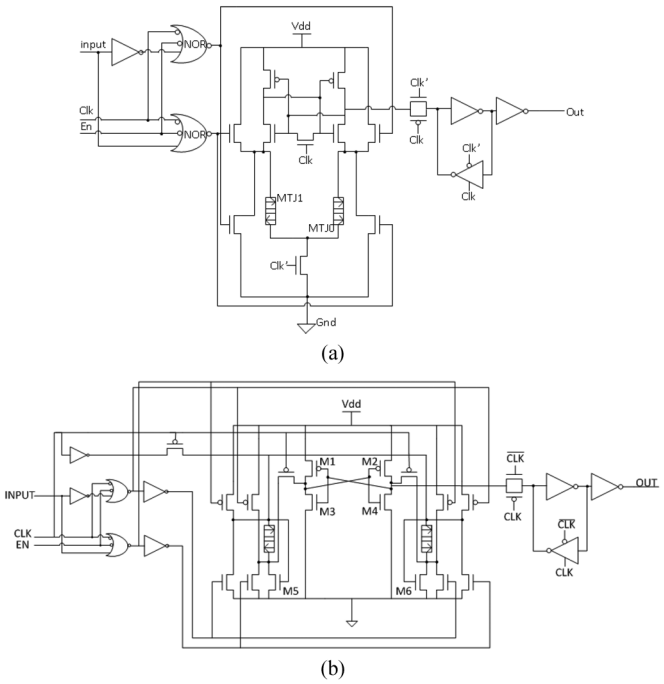

Fig. 8. Nonvolatile flip–flop designs. (a) Original design in [1]. (b) Our modified design.

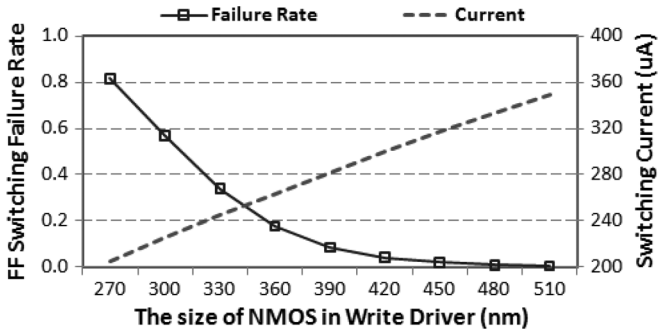

Fig. 9. The nonpersistent write failure rate of nonvolatile flip–flop when the transistor size varies.

on the two MTJs is too small. The new design connects MTJ and NMOS transistor (M5 or M6) in series to provide credible inputs for M1–M4. Depending on the data stored in two MTJs, one of M5 and M6 works in saturation region and another works in linear region. By properly sizing M5 and M6, the currents through MTJs can be adjusted, and consequently, the design can be more process-variation tolerant. However, similar to STT-RAM cell designs, increasing the read current may result in the increase in read disturbance probability.

Fig. 9 shows the nonpersistent write failure rate of our nonvolatile flip–flop when increasing the size of PMOS-NMOS write-driver pair. Here, we assume the PMOS transistor size is always twice of the NMOS transistor size. The failure rate follows the similar trend as the one of STT-RAM.

Fig. 10 shows the tradeoffs between the persistent and nonpersistent read errors when sizing up M5–M6. We use the same simulation setup as Section II-D for process variations. As expected, increasing the size of these transistors can produce the higher reading current through the MTJ pair. Hence, the persistent errors due to process variation can be reduced significantly.

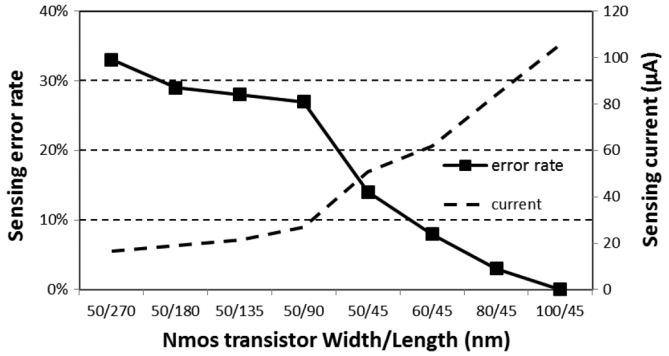

Fig. 10. Tradeoff between read disturbance probability and sensing errors of nonvolatile flip–flop when the transistor size varies.

IV. Conclusion

In this work, we thoroughly analyze the persistent and nonpersistent errors in Spin-MOS circuitries: the former mainly comes from process variations, and the later one is resulted by thermal fluctuations and read disturbance. On top of it, we quantitatively investigate the impacts of these variations and fluctuations on the operations of two typical Spin-MOS circuitries: 1T1J spin-transfer torque random access memory (STT-RAM) cell and a nonvolatile flip–flop design. The possible design techniques to reduce thermal incurred nonpersistent error rate are also discussed. Our experimental results show that the optimization of nonpersistent and persistent errors are closely entangled with each other and should be conducted from both circuit design and magnetic device engineering perspectives simultaneously.

References

[1] N. Sakimura *et al.*, "Nonvolatile magnetic flip-flop for standby-power-free SoCs," in *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2008, pp. 355–358.

[2] X. Guo, E. Ipek, and T. Soyata, "Resistive computation: Avoiding the power wall with low-leakage, STT-AM based computing," in *Proc. 37th Annu. Int. Symp. Comput. Architecture*, Jun. 2010, pp. 371–382.

[3] T. C. Chen, "Overcoming research challenges for CMOS scaling: Industry directions," in *Proc. 8th Int. Conf. Solid-State Integr. Circuit Technol.*, 2006, pp. 4–7.

[4] S. Lu, "Microprocessor memory circuits," in *Presentation in Workshop on Technology-Architecture Interaction: Emerging Technologies and Their Impact on Computer Architecture*, 2010, (Held in conjunction with 43rd Annu. IEEE/ACM Int. Symp. Microarchitecture).

[5] 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 *Proc, IEEE Int. Electron Devices Meeting*, Dec. 2009, pp. 1–4.

[6] Y. Higo *et al.*, "Thermal activation effect on spin transfer switching in magnetic tunnel junctions," *Appl. Phys. Lett.*, vol. 87, p. 082502, 2005.

[7] M. Hosomi *et al.*, "A novel nonvolatile memory with spin torque transfer magnetization switching: Spin-RAM," in *Proc. IEEE Int. Electron Device Meeting*, Dec. 2005, pp. 459–462.

[8] Y. Chen, H. Li, X. Wang, W. Zhu, W. Xu, and T. Zhang, "Combined magnetic- and circuit-level enhancements for the nondestructive self-reference scheme of STT-RAM," in *Proc. ACM/IEEE Int. Symp. Low Power Electron. Design*, Aug. 2010, pp. 1–6.