Impact of Self-Heating on Reliability of a Spin-Torque-Transfer RAM Cell

Subho Chatterjee, Sayeef Salahuddin, Satish Kumar, and Saibal Mukhopadhyay, *Member, IEEE*

***Abstract*—This paper estimates the temperature distribution within a spin-torque-transfer RAM (STTRAM) cell due to self-heating using a thermal simulation based on the finite volume method. The analysis shows that, due to high switching current and small volume of the magnetic tunnel junction (MTJ), there can be significant rise in temperature in the MTJ as well as the silicon transistor. The impacts of the increased temperature on operational reliability metrics of the STTRAM cell, i.e., read disturb, write failure, and sensing accuracy, are evaluated. It is shown that, due to the self-heating effect, the operational reliability of an STTRAM cell depends on the read–write history of that cell.**

***Index Terms*—Magnetic tunnel junction (MTJ), read disturb, self-heating, sensing accuracy, spin-torque-transfer random access memory (STTRAM).**

I. Introduction

SPIN-torque-transfer RAM (STTRAM) has emerged as a strong candidate for future embedded memory devices [1]–[6]. An STTRAM cell consists of a magnetic tunnel junction (MTJ) connected in series with an n-channel MOS (NMOS) access device. This cell is connected between the bit lines (BLs) and the source lines (SLs) and is accessed using the word line [see Fig. 1(a)]. These memory cells are replicated in a highly dense manner to create a memory array [see Fig. 1(b)]. The MTJ consists of two ferromagnetic layers separated by a dielectric layer referred to as the spacer (usually MgO) [see Fig. 1(c)] [1], [2]. The magnetization of one layer is fixed, whereas that of the other can be controlled by the injection of spin polarized electrons. The MTJ offers different resistances in the two modes of magnetization: A parallel configuration offers lower resistance R_L, and an antiparallel configuration offers higher resistance R_H. This resistance difference is used to store and detect logic states in the cell. When a write current greater than a critical switching current flows through the MTJ in the proper direction, the state of the MTJ switches from antiparallel to parallel or vice versa. The critical switching current is a property of the MTJ ($\sim$100 μA for a write pulsewidth of $\sim$5–10 ns). The proper direction of the write current is from BL to SL for antiparallel-to-parallel switching and SL to BL for parallel-to-antiparallel switching.

The high write current and small device volume can result in very high power density within the MTJ device and the STTRAM cell. To understand the problem, consider an MTJ device with a diameter of 100 nm and a thickness of 10 nm (surface area $\propto$ 50 nm $\times$ 50 nm, and volume $\propto$ 50 nm $\times$ 50 nm $\times$ 10 nm) connected to a bulk-silicon transistor through a metallic via. For a write current of $\sim$100 μA, the power density within the MTJ volume can be easily estimated at $\sim 10^{12}$ W/cm^3 and at the surface is $\sim 10^6$ W/cm^2. This significantly high value of the power density can lead to a localized temperature increase in the MTJ. The temperature increase is further enhanced due to the fact that MTJs are embedded within the interlayer dielectric, which is a poor conductor of heat. The metallic via conduct the heat resulting in an increased temperature in the silicon. Hence, high switching current can increase both MTJ and silicon temperature. We refer to this effect as self-heating in STTRAM. The increased temperature can degrade the operational reliability of the cell. The operational reliability of an STTRAM cell depends on the following failure mechanisms: read disturb (cell flipping during read), write failure (incorrect write operation), and detection failure (incorrect sensing of cell value). Hence, estimating the self-heating effect in STTRAM and analyzing its effect on cell reliability is important for the development of STTRAM technology. While the effects of temperature on the MTJ device has been studied [7], [8], the modeling and analysis of the self-heating effect has received limited attention in literature [9], [10].

This paper models self-heating in the STTRAM cell and analyzes its effect on the operational reliability of the cell [see Fig. 1(d)]. We have presented a detailed model based on the finite volume method (FVM) to estimate the self-heating effect in the STTRAM cell. We analyze the steady-state and transient thermal behaviors of the STTRAM cell using the FVM model. The impact of MTJ parameters such as the critical switching current, write pulsewidth, and resistance–area (R–A) product on the temperature distribution is analyzed. The effect of temperature on MTJ (the critical switching current and parallel/antiparallel resistances) and the NMOS transistor (linear, saturation, and leakage current) properties were studied. The mixed-mode device–circuit simulation is used to analyze the effect of the temperature rise on the different operational reliability metrics of the cell. Finally, the observations from

Manuscript received May 19, 2011; accepted November 21, 2011. Date of publication January 24, 2012; date of current version February 23, 2012. This work was supported in part by the National Science Foundation under Grant BS123456 and Grant ECCS-1002090. The review of this paper was arranged by Editor D. Esseni.

S. Chatterjee and S. Mukhopadhyay are with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: subho.chatterjee@gatech.edu; saibal@ece.gatech.edu).

S. Salahuddin is with the Electrical Engineering and Computer Science, University of California, Berkeley, CA 94702 USA (e-mail: sayeef@eecs.berkeley.edu).

S. Kumar is with the School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: satish.kumar@me.gatech.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/TED.2011.2180726

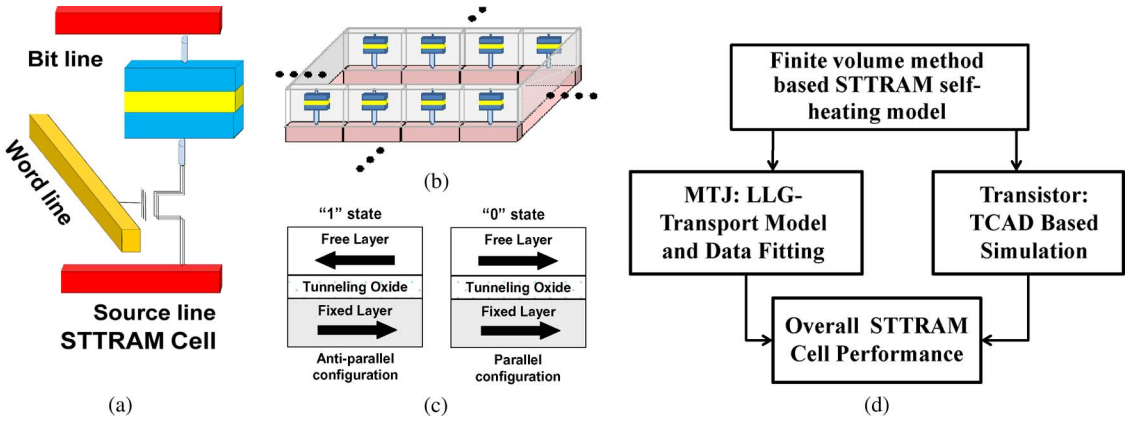

Fig. 1. (a) STTRAM cell structure with MTJ, NMOS, and controlling metal lines (BL, SL, and word line). (b) STTRAM array structure. (c) MTJ in the '1' and '0' configurations. Relative orientation of free-layer and fixed-layer determined resistance states. (d) Electrothermal cosimulation framework for STTRAM. It uses FVM-based self-heating solutions coupled with the LLG transport model for MTJ and technology computer-aided design simulations for silicon.

the above analysis were coupled to study the interaction of self-heating effect and cell reliability to estimate the effect of read/write data patterns on operational reliability of STTRAM.

The rest of this paper is organized as follows. In Section II, the impact of temperature on STTRAM properties is investigated, Section III presents the FVM-based modeling methodology to estimate self-heating, Section IV studies the effect of read–write history on the operational reliability of cell considering self-heating, and Section V summarizes this paper.

II. STTRAM: Impact of Temperature

The functional reliability of an STTRAM cell is defined by the following metrics.

- **Write margin:** The write margin is defined by the difference between the write current and the critical switching current of the MTJ. The write current depends on the MTJ resistance and the transistors. An increase in the critical switching current and/or reduction in the write current (due to increase in MTJ resistance and/or reduction in the transistor strength) degrades the write margin.
- **Read margin:** The read margin is defined as the difference between the critical switching current and the read current (current flowing through the MTJ during read operation). The read current depends on the resistances of the MTJ and the transistor. A lower critical switching current and/or higher read current degrades the read margin.
- **Detection accuracy:** Incorrect or false detection refers to the detection of a bit as '1' when the stored bit is '0' and vice versa. In STTRAM, the bit values are detected depending on the difference in the read current for cell storing '0' and '1' (i.e., MTJ in antiparallel state or parallel state). Further, during the sensing of an STTRAM cell, the current flowing through the BL is sensed. Chatterjee *et al.* [11] have pointed out that, during reading a cell in a selected column, the unselected cells in that column contributes to leakage current. This leakage acts as a circuit-induced noise to the sensed current. Hence, variations in the MTJ resistances, the transistor strength, and the transistor leakage modulate the detection accuracy.

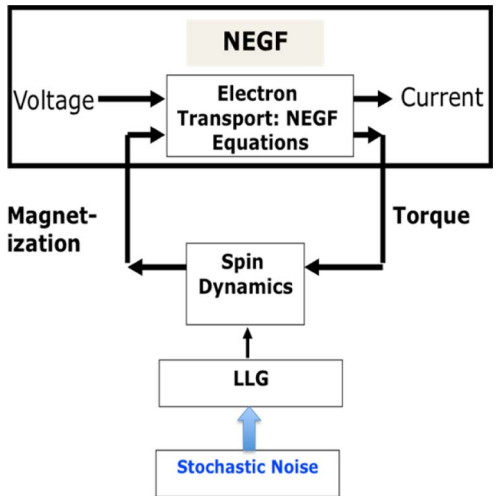

Fig. 2. MTJ device simulation framework: self-consistent solution of LLG and NEGF transport equation.

In this section, we study the effects of temperature on the properties of the MTJ, on the parameters of the NMOS device, and, finally, on reliability metrics of the cell.

A. Impact on Temperature on MTJ Properties

1) Impact on the Critical Switching Current of the MTJ: Individual STT devices were simulated using a coupled quantum transport–magnetization dynamics framework. The effect of temperature was included through a stochastic integration of the magnetization dynamics. Fig. 2 shows the schematic of the simulation methodology. Essentially, the spin polarized current flowing in the device exerts a torque on the magnet that is calculated from the nonequilibrium Green's function (NEGF) formalism and then used as a source term in the magnetization dynamics modeled using the Landau–Lifshitz–Gilbert (LLG) equation [12]. This changes the relative orientation of the magnetization. In turn, the specific orientation of the magnetization changes the way the spin polarized current flows through the device. Thus, NEGF–LLG needs to be solved self-consistently. Detailed discussions on this self-consistent framework may be found in [13]–[15]. The transport simulation can reproduce

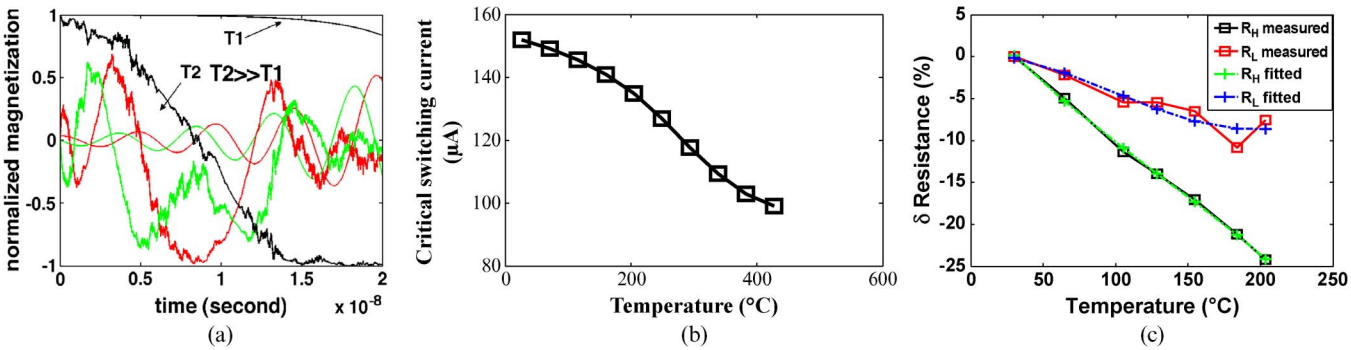

Fig. 3. (a) Faster but chaotic switching of the normalized magnetization at higher temperature ($T1 = 0$ K and $T2 = 300$ K). (b) Reduction in the critical switching current with temperature. (c) Effect of temperature variation on the MTJ resistance [R_L is the resistance of the parallel state (or low resistance state), and R_H is the resistance of the antiparallel state (or high resistance state)]. The results are based on the curve obtained from the data reported in [17].

experimental data quite well, as has been shown in [16], at $T = 300$ K. For this paper, switching dynamics was simulated by self-consistent NEGF–stochastic-LLG simulations and assuming 80% flipping as the switching threshold. In Fig. 3(a), $T1$ and $T2$ are two different temperatures such that $T2 \gg T1$. In Fig. 3(a), $T1 = 0$ K, and $T2 = 300$ K. The traces are shown to illustrate the chaotic nature of switching when temperature is properly accounted for in the simulation. The two sets of plots show the traces of magnetization in time with at two different temperatures. The black color shows the z component, the green color shows the x component, and the red shows the y component. For the same component, e.g., the black traces, the noisy evolution of dynamics at $T = T2$ is visible compared with the smooth trace at $T = T1$.We observe that an elevated temperature results in a faster but chaotic switching [see Fig. 3(a)] [13], [14]. Consequently, the critical switching current reduces at higher temperature [see Fig. 3(b)]. The required amplitude for threshold current goes down with temperature due to the fact that the magnetic spins now have a larger thermal energy to cross over the barrier. This comes through the stochastic part of the LLG equation. Hence, at a larger temperature switching current, the requirement falls off at a faster rate with time. This is particularly observed at lower pulsewidths and leads to an increased variance and a reduced mean. However, in this paper, we concentrate on the effect of temperature on the mean switching current.

2) Impact on MTJ Resistance: Majumder *et al.* have experimentally demonstrated the effect of temperature on MTJ resistance [17]. It was observed that the resistances in the parallel and antiparallel modes for MgO-based MTJs reduce at a higher temperature [see Fig. 3(c)]. Another critical observation is that the antiparallel resistance R_H reduces at a faster rate with increasing temperature compared with the parallel resistance R_L. Consequently, the tunnel magneto-resistance (TMR) of the MTJ defined as $(R_H - R_L)/R_L$ also degrades with increasing temperature. This is in conjunction with the model assuming conductance having two components, i.e., spin dependent and spin independent. The spin-dependent component is expected to follow the empirical dependence of $T^{-\alpha}$ [8]. This reduction in resistance would increase write and read currents through the cell. We fit the observed data to a polynomial curve and use that to analyze the effect of self-heating.

TABLE I
PROPERTIES OF THE NMOS ACCESS DEVICE

DIBL	44.44 mV/V
I_{ON}	532μA
I_{OFF}	4.36nA
I_{ON}/I_{OFF}	1.22e5
$V_{THRESHOLD}$	0.26 V
Gate Length	65nm
Work Function	4.2eV
Oxide + High K Thickness	3nm

B. Impact of Temperature on the NMOS Access Device

In this section, we perform the study considering a 65-nm high-k metal-gate transistor using a drift–diffusion-based mixed-mode device simulator (i.e., Medici [18]). In the following subsections, the transistor characteristics and their temperature dependence are characterized to evaluate the STTRAM properties. Table I shows the simulated transistor dimensions and properties. There is a good agreement between the Medici-simulated 65-nm device characteristics and the measurement data [19]. The I_d–V_g characteristics for the transistor are shown in Fig. 4(a).

Note that, during read, the NMOS is in the linear mode. This is also the situation for write '0' when current flows from BL to SL. For write '1,' however, the current flows from SL to BL, and the transistor is in the saturation region with the MTJ resistance in its source. Therefore, we study the effect of temperature on the linear mode resistance (for read and write '0'), as well as the saturation current (for write '1') of the transistor. The transistor resistance is found to increase by 0.5 kΩ for about 100 °C rise in temperature in the linear mode [see Fig. 4(b)]. Likewise, we observe that saturation current reduces at a higher temperature [see Fig. 4(c)]. Hence, it is expected that read current and both forms of write current (write '0' and write '1') will reduce at higher temperature. This can negatively impact the read/write margin and detection accuracy. On the other hand, the subthreshold current of the transistor increases exponentially with an increase in temperature, as shown in Fig. 4(d). The subthreshold leakage is defined as the drain-to-source current of a MOSFET before the inversion, i.e., when gate-to-source voltage is less than the threshold voltage

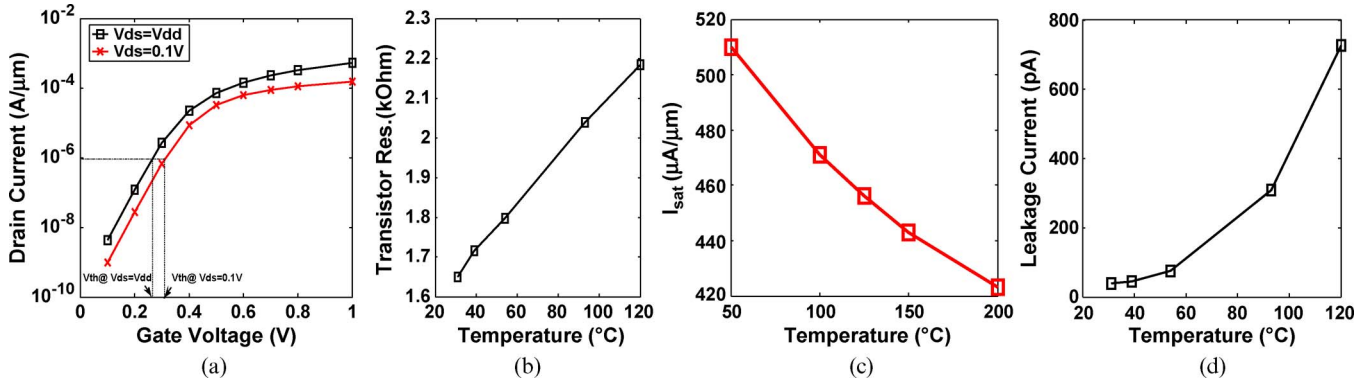

Fig. 4. (a) Simulated transistor I_d–V_g characteristics. A current density of 10^6 A/μm is used to find threshold voltage across saturation and linear modes. This is used to evaluate drain induced barrier lowering (DIBL). (b) Temperature dependence of transistor resistance in linear mode. (c) $I_{\rm sat}$ variation with temperature. (d) Temperature dependence of leakage across a transistor of 1-μm width.

TABLE II
MTJ PROPERTIES AT 27 °C

R_L-A	32.9 Ω-μm^2
TMR	42.8%
R_H and R_L	4.7kΩ and 3.29kΩ
Switching Current Density (J_C)	1x10^{6}A/cm^2
Area (A)	100x100 nm^2
Height (h)	10nm

$(V_{\rm gs} < V_{\rm th})$. The subthreshold leakage of a transistor is due to the diffusion of the minority carriers before the inversion of the channel (i.e., electrons for the NMOS). In other words, the subthreshold current is due to the thermoionic emission of electrons over the source-to-channel potential barrier that exists before inversion. As the probability of thermoionic emission over a barrier depends exponentially on the temperature, the subthreshold leakage depends exponentially on temperature. This is also evident from the model of the subthreshold leakage current as follows:

$$I_{\rm leakage} = I_{D0} e^{\frac{V_{\rm gs}-V_{\rm th}}{\eta V_T}}$$

where I_{D0} is the leakage current at gate-to-source bias $V_{\rm gs}$ equal to the threshold voltage $V_{\rm th}$, η is a constant, and $V_T = {\rm kT/q}$.

C. STTRAM Cell Level Impact

For the STTRAM cell, we simulate the read and write conditions in HSPICE and measure the currents. For the simulation of the STTRAM cell, we consider the MTJ properties shown in the Table II. To understand the variation of the STTRAM performance metrics with temperature, we have to simulate the MTJ and NMOS temperature dependence in conjunction. We study the combined effect using mixed-mode device simulation (using [18]). We model the STTRAM cell using the NMOS device discussed in previous section and resistances to represent the MTJ. To study the effect of temperature variation on the cell parameters, we vary the "simulation temperature," which modifies the NMOS properties. On the other hand, the effect of temperature on the MTJ is captured by modifying the value of the MTJ resistances [R_H and R_L as appropriate and by following Fig. 3(c)].

First, we consider the values of MTJ resistances R_H and R_L and the critical switching current at room temperature. The width of the NMOS was chosen to ensure correct write operation at room temperature. This is achieved when the current flowing through the cell in both directions (i.e., BL to SL and SL to BL) is higher than the MTJ switching current at room temperature. We consider the write pulsewidth of 100 ns. For read operation, we consider the current-based sensing operation [4]. We consider a BL voltage of 0.35 V and estimate the read currents for high (R_H) and low (R_L) MTJ resistances. The NMOS transistor in both conditions uses the same gate voltage and hence contributes the same resistance. Further, we account for the leakage current of the unselected cells in the selected column. The read pulsewidth is assumed to be the same as that of the write pulsewidth (∼100 ns). As previously mentioned, the sensed current includes the read current of one selected cell plus leakage of unselected cells. This is defined as the array-level read current. The reference current is considered to be the average of the array-level read currents estimated with the MTJ of the selected cell at R_H (bit 1) and R_L (bit 0). We quantify the detection accuracy as the array-level TMR (ATMR) $[= (I_{\rm cell0} - I_{\rm cell1})/(I_{\rm cell0} + I_{\rm leakage})]$, where $I_{\rm leakage}$ is the total leakage of all unselected cells [11].

1) Read Disturb: Read disturb occurs when the read current is larger than the switching current and, hence, flips the bit content. Given the direction of read, only the bit flip from '1' to '0' is likely. Therefore, the read margin can be defined as the difference between MTJ switching current and read '1' current. We plot the behavior of read '0' and '1' currents with temperature and compare it with the switching current at different temperature [see Fig. 5(a)]. Note that, at higher temperature, MTJ resistances in both parallel (read '0,' R_L) and antiparallel modes (read '1,' R_H) reduces, but R_H reduces at a much higher rate [see Fig. 3(c)]. The transistor resistance increases with temperature [see Fig. 4(b)]. Results indicate that the read '0' current reduces with temperature as the increase in transistor resistance overshadows reduction in R_L. However, read '1' current can even increase at higher temperature as large reduction in R_H can mask the increase in transistor resistance. The increase in read '1' current [see Fig. 5(a)] and decrease in the critical switching current of MTJ [see Fig. 3(b)] results in a reduction in the read margin [see Fig. 5(a)] at high temperature.

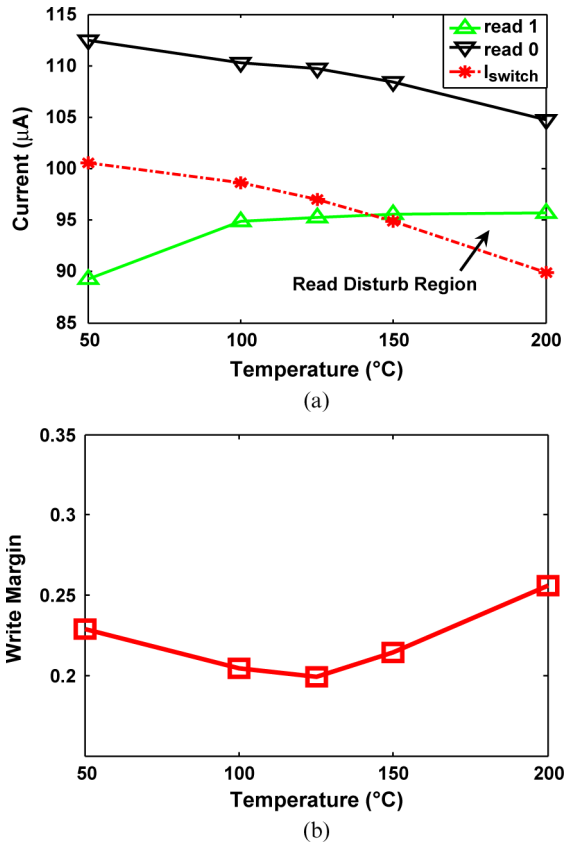

Fig. 5. Evaluation of the impact of temperature on (a) read disturb and (b) write margin.

2) Write Failure: Write failure occurs when the write current falls below the critical switching current. For a '0' to'1' flip, the circuit is in source degenerate mode (MTJ resistance present at the source). For a '1' to '0' flip, MTJ acts as a resistive load at the NMOS drain. Thus, for the same $V_{\rm dd}$ applied to both cases, write current is lesser among two flip conditions for '0' to '1.' Hence, we consider the write '0' to '1' as the more probable switching condition for write failure. At higher temperature, switching current requirement reduces, which tends to increase the write margin. The effect of increased temperature on write current is determined by two contrasting factors. A reduction on the current of transistors at higher temperature tends to reduce write current [see Fig. 4(c)], whereas a lower MTJ resistance [see Fig. 3(c)] ($\sim R_L$ before MTJ switching) helps increase write current. The net effect of the above three factors determine the sensitivity of the write margin with temperature. We observe that the combined effect makes the write margin less sensitive to temperature [see Fig. 5(b)].

3) False Read and Detection Accuracy: A higher NMOS resistance reduces the ratio between cell resistance (MTJ resistance + NMOS resistance) for bit '0' and '1' with respect to the average cell resistance. This can increase the probability of false detection. This is shown in Fig. 6(a), which plots the ratio of read current during reading '1' and reading '0' at higher temperature. Moreover, as explained earlier, leakage from the unselected cells of the selected column reduces the detection accuracy further. As the leakage increases with temperature, ATMR degrades with temperature [see Fig. 6(b)].

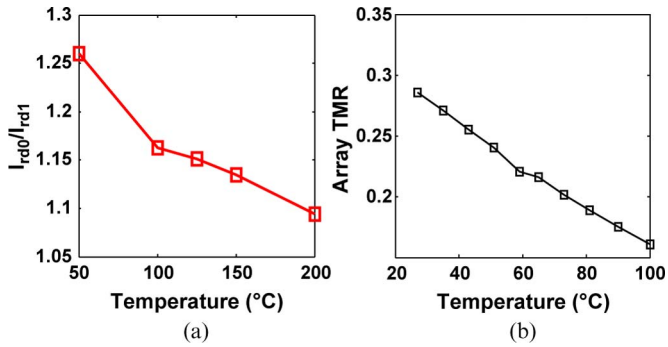

Fig. 6. Impact of temperature on (a) Read '0' current/Read '1' current (b) Array-level distinguish-ability metric [11].

III. Modeling of Self-Heating in STTRAM

We perform an FVM-based heat transport analysis in the STTRAM structure to characterize the self-heating effect.

A. FVM-Based Model

FVMs have been widely used to obtain temperature distribution in thermal systems. In this method, Fourier conduction equations are integrated over each control volume (grid cell) to get algebraic equations for each cell. There is a tradeoff between the meshing resolution and solution time. For this paper, we use Gambit for the generation of the STTRAM cell mesh [see Fig. 7(a)]. Fig. 7(b) shows a cross section of the implemented cell structure, whereas Fig. 7(c) shows the top view of the cell layout. We used nonuniform meshing across the STTRAM structure. This is to maintain a balance between the small mesh resolutions required for representing certain portions (e.g. MTJ) while a coarse resolution is maintained for other parts to restrict the memory requirement and computation time. We use Fluent finite volume solver for the thermal simulations. A convection boundary condition is applied at the Si surface such that a current of 10 μA flowing through an equivalent resistance of 1 kΩ across a transistor of length $=$ 200 nm, width $=$ 700 nm, and junction depth $=$ 10nm gives a temperature rise of 52 °C across the bulk silicon [19]. Next, we consider an MTJ with an area of 100×100 nm^2. The $R{-}A$ product of such an MTJ is 32.9 $\Omega \cdot \mu$m^2.

B. Results of FVM Analysis: Steady State

Considering a critical switching current density of 1×10^6A/cm^2, the critical switching current requirement for the MTJ is evaluated at 100 μA. For an MTJ with a height of 10 nm, the power density of 2.5×10^{17} W/m^3 results across the MTJ, and the power density of 5×10^{15} W/m^3 results across silicon. If current flows continuously through the structure, the final temperature distribution reaches the steady state. Fig. 7(d) shows the thermal distribution across an isolated cell at a steady-state condition for the FVM simulation conducted using Fluent. The cell consists of the bulk silicon (temperature in blue), the MTJ (temperature in red), and the metallic via (temperature in yellow and green). The bulk and MTJ show an 11 °C temperature difference. A steady-state solution estimates a final temperature of 112 °C inside the MTJ.

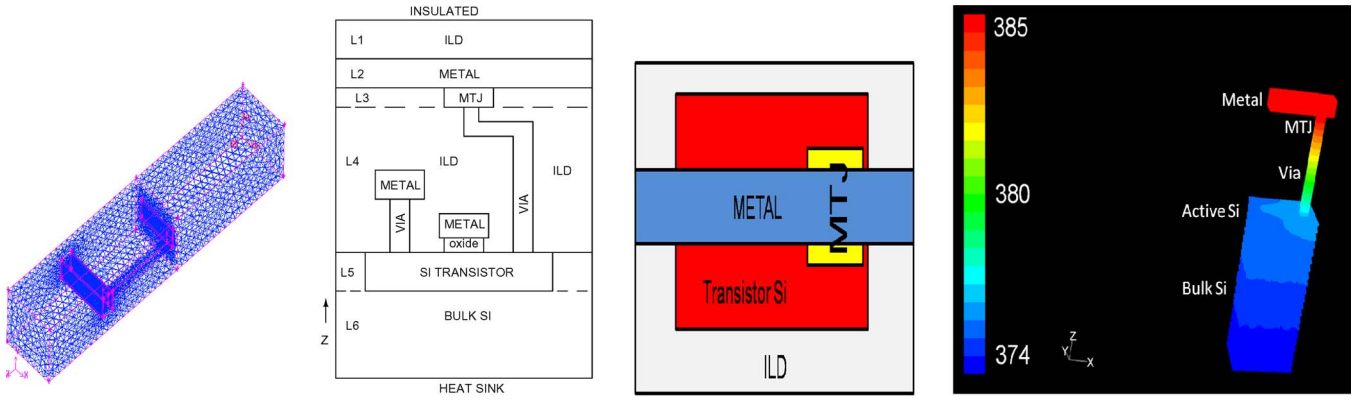

Fig. 7. (a) Constructed mesh for the cell in Gambit. (b) Cross-sectional view of a STTRAM cell. (c)Top view of the cell. (d) Temperature distribution across simulated FVM model for a STTRAM cell. The STTRAM cell distinctly shows the silicon (bulk and active) via the MTJ and metal showing a temperature difference of 11 °C.

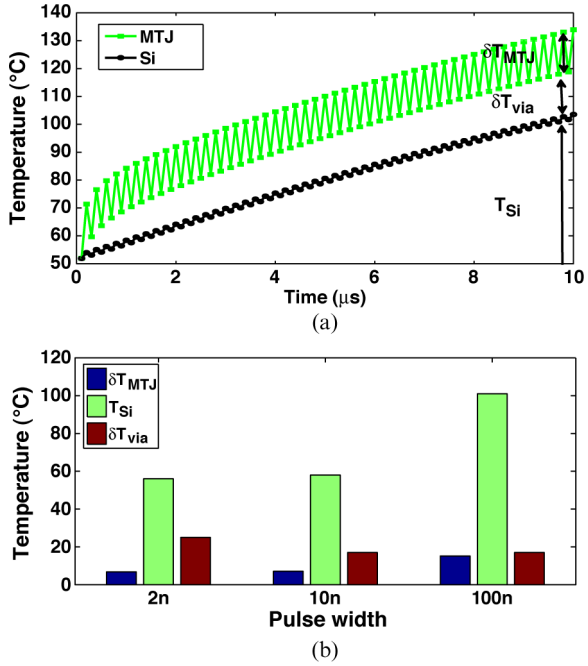

Fig. 8. (a) Temperature rise for applied write pulse across MTJ and bulk. (b) Effect of different pulsewidths on temperature (δT is temperature difference across the material).

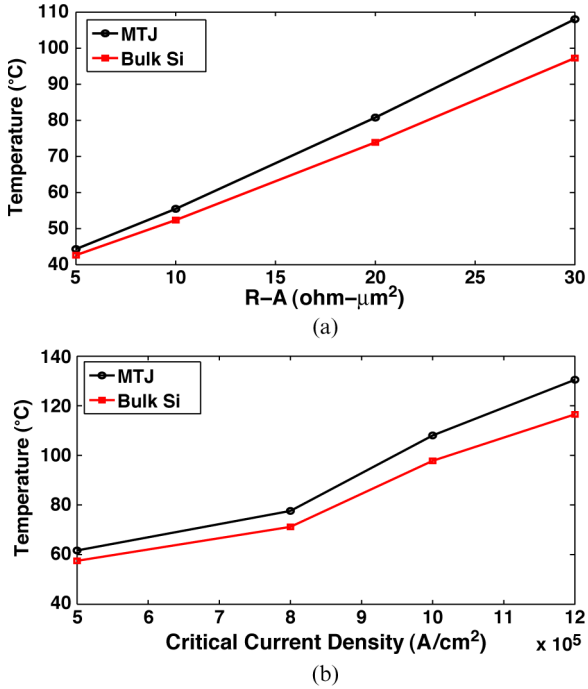

Fig. 9. Temperature variation with (a) R–A product and (b) critical switching current density

C. Results of FVM Analysis: Transient Simulation

We are interested in the transient thermal response of the MTJ and Si. The FVM model is subjected to current pulses of 100 μA with a 200-ns time period and a 50% duty cycle, and the temperature is observed. The result is shown in Fig. 8(a). Note that the MTJ temperature rises at a much faster rate than the Si. This is because of the higher thermal capacitance of the bulk Si compared with the MTJ. In addition, the MTJ shows a fluctuation in temperature by 7 °C within a time period between the on and off cycles. We consider this along with the temperature drop across the metal via and silicon temperature to characterize the thermal transient. We consider the thermal distribution across the STTRAM for cases of write pulsewidths of 2, 10, and 100 ns considering the switching currents of 480, 300, and 160 μA (obtained from [4]), respectively. The results for performing 100 consecutive write cycles are shown in Fig. 8(b). Note that the difference across the metal via, i.e., the MTJ, and the silicon is approximately 9 °C–10 °C and is the highest for 2 ns. The overall temperature rise is greater for the 100-ns case.

D. Results of FVM Analysis: Effects of MTJ Properties

The MTJ material and geometry decides the R–A product. The current R–A products for MgO-based devices stand near 30 $\Omega\mu$m^2. We simulate the cell temperature distribution for a range of R–A products from 5 to 30 $\Omega\mu$m^2 to understand how variation in the R–A product modifies temperature distribution. For the same area and switching current, increasing resistance by a factor of 2 will result in a two times increase in the power dissipation. This translates to a proportional temperature increase, as shown in Fig. 9(a). Similar deductions can be made regarding the critical switching current density. With choices of anisotropy, a free-layer damping factor, and geometry, there

is a constant effort to reduce the switching threshold current density J_{c0}. Recent reported values of J_{c0} usually lie in the range of 1–3 MA/cm^2 [5]. The critical switching current density J_c depends on the choice of J_{c0} and the pulsewidth. We study the impact of critical switching current density on temperature in Fig. 9(b). A reduction in the critical switching current density implies a reduced current and, hence, reduced temperature for the same area and the R–A product. In summary, we observe that reducing the critical current density and the R–A product are desirable to reduce the self-heating effect and the temperature rise within the cell.

IV. Impact of Self-Heating on STTRAM

In this section, we apply the self-heating results obtained in the previous section to evaluate the operational reliability metrics of the STTRAM cell. The NMOS device and MTJ shown in Tables I and II are used for thermal and electrical analysis. We begin with several read and write currents estimated from circuit simulations (i.e., applying a constant BL, SL, and wordline voltage depending on the operating condition). The MTJ and silicon temperature will be different for a given read or write condition. The estimated set of temperatures is used in the mixed-mode circuit simulations with the NMOS device and the MTJ resistance to reestimate the read and write currents. We perform the above analysis considering self-heating for different read and write currents depending on the bit values and patterns. For a given read/write condition, we consider the steady-state MTJ and NMOS temperature for cell reliability analysis (i.e., the pulsewidth is large or the same operation is repeated continuously for a large number of cycles) to evaluate the worst-case cell reliability.

A. Different Write and Read Patterns

Considering a read voltage of 0.35 V between BL and SL, we measure the read currents for reading '0' and '1.' Fig. 10(a) shows the current for different read and write conditions. Fig. 10(b) shows the corresponding steady-state temperatures. We observe that, for the same pulsewidth, the temperature is less by 25 °C–30 °C for read operations, compared with the write operations. As read and write occur to the same cell, the temperature distribution of a cell is a function of the read–write access pattern and the worst case arises due to consecutive write operations. Further, there can be write operations for flipping (0-1 or 1-0) and redundant write operations (0-0 or 1-1). Each write condition offers a certain initial MTJ resistance (R_H ['1'] or R_L ['0']) and circuit configuration (write 0: BL− >high SL− >low, and write 1: BL $->$ low and SL $->$ high). Simulation results show that the write 0-0 and write 1-0 gives rise to the maximum temperature [see Fig. 10(b)] as the currents are much higher compared with other cases [see Fig. 10(a)].

B. Effect of Past Access History

The operational reliability of the STTRAM cell will depend on the read–write history of the cell. In this section, we study the impact of past operations on the reliability of a current read or write operation.

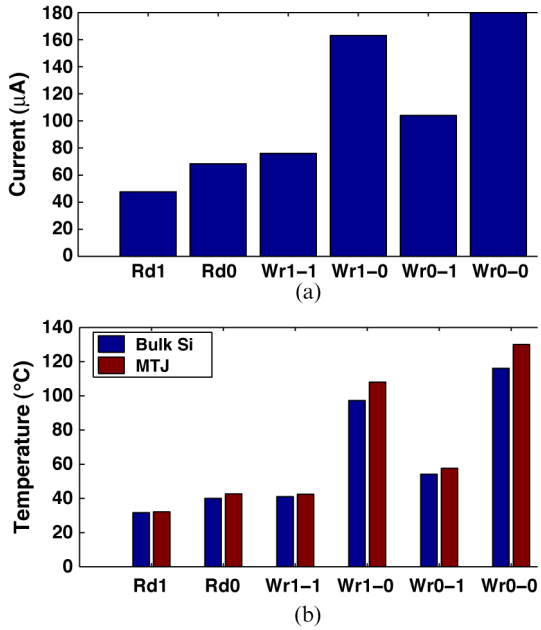

Fig. 10. (a) Read–write currents. (b) Temperatures for the cases.

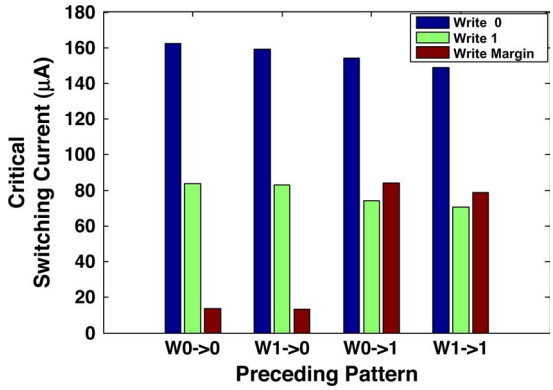

Fig. 11. Access pattern history dependence of write current

Write: In this section, we analyze the effect of previous access patterns on write disturb. We first evaluate the MTJ and NMOS temperature at the beginning of a current write operation considering self-heating due to different past read/write operations. The estimated temperature is used to estimate the write current (using mixed-mode simulations) and the MTJ critical switching current [using Fig. 3(b)] at the beginning of current operation. In Fig. 11, the x-axis indicates the previous set of operations executed. Write failure can occur when there is a write with flip requirement ('0' to '1' or '1' to '0'). The first set of bars with x-axis label w0 $\rightarrow$ 0 represents the write currents for writing '0' (blue), writing '1' (green), and write margin (red). The label w0 $\rightarrow$ 0 represents that the previous operation was writing '0' to '0.' When the previous history is w0 $\rightarrow$ 0, the next write '0' represents a redundant write. The write '0' in this case, is therefore important only for energy analysis. The write margin (red bar) in this case is computed considering write '1' current.

We observe that, for all cases of past operations, the write '0' current is always larger than the write '1' current. This is attributed to the bidirectional switching condition in STTRAM.

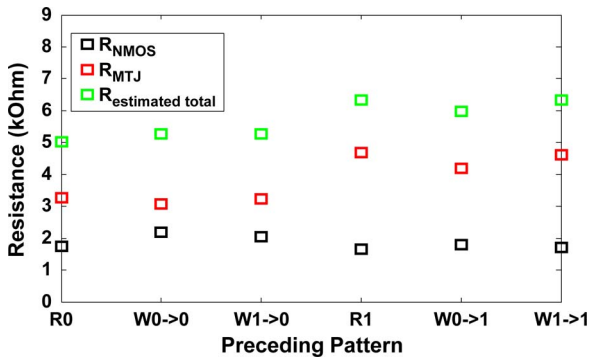

Fig. 12. Access history dependence of resistance

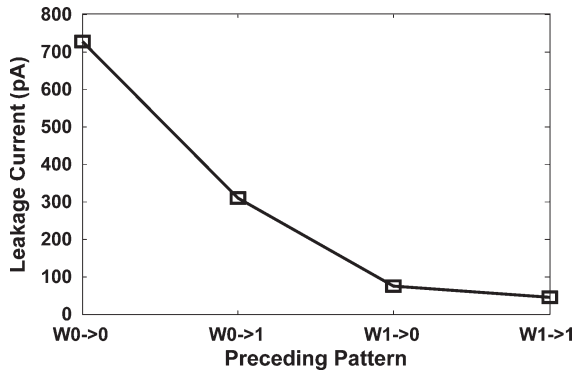

Fig. 13. Access history dependence of leakage.

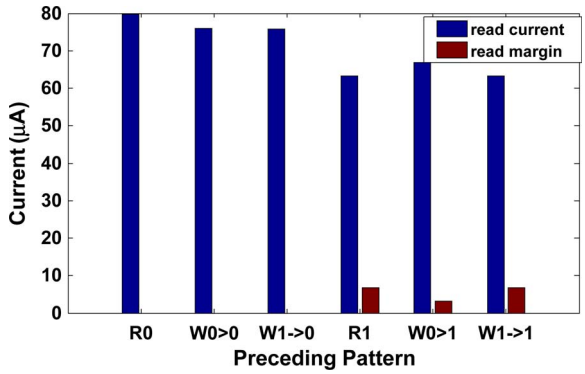

Fig. 14. Access history dependence of read disturb.

The NMOS size is determined considering write '1' (i.e., when MTJ appears in source of NMOS), which leads to larger than the required size for the write '0' condition (the MTJ acts as a load at the drain of NMOS). Write failure while writing '0' can only occur when past conditions are w0 → 1 or w1 → 1. We observe that write '0' current is minimum when previous pattern is w1 → 1 (see Fig. 11). This is because w1 → 1 leads to the minimum initial temperature and, hence, the maximum value of MTJ resistance in the antiparallel (R_H) condition and higher MTJ switching current [see Fig. 3(b)]. Hence, Fig. 11 shows that self-heating and the past history result in a 4% reduction in the write margin for write '0.' Likewise, we observe that the write margin for write '1' is 3% smaller when the past operation is w1 → 0, compared with w0 → 0. This is because w0 → 0 results in higher initial temperature due to self-heating.

Read: Next, we study the effect of the preceding pattern on the detection reliability (see Fig. 12). Fig. 12 plots the MTJ, NMOS, and combined resistances. The read '0' cases are the ones with the prior history of R0 (read '0') and/or w0 → 0 and/or w1 → 0. The two resistances (i.e., NMOS and MTJ) are closest to each other following a redundant write '0' (w0 → 0). This is due to the maximum initial temperature for both MTJ and NMOS (see Fig. 10), which results in the lower MTJ resistance R_L [see Fig. 3(c)] and higher device resistance [see Fig. 4(b)]. The maximum difference between MTJ and NMOS resistance during read '0' is observed for the previous pattern of R0 (the smaller MTJ and NMOS temperature in Fig. 10). Cell resistance for read '0' can vary by 5.4% due to prior history and self-heating. For read '1,' the previous history of w0 → 1 results in the least cell resistance and the minimum difference between the NMOS and MTJ resistance R_H. This is because the prior history of w0 → 1 results in the maximum initial temperature for read '1' condition (see Fig. 10) and, hence, smaller MTJ resistance and higher NMOS resistance. The prior history of read '1' R1 leads to maximum cell resistance and maximum difference between MTJ and NMOS resistance (due to smaller temperature considering self-heating). Therefore, we observe that cell resistance for read '1' can vary by 6% due to self-heating. It is imperative that detection accuracy degrades when the difference between MTJ resistance and NMOS resistance reduces. Hence, false detection for read '0' and read '1' are most probable with the past history of w0 → 0 and w0 → 1, respectively.

The cell-level distinguishing ability depends on the ratio of cell current while reading '0' and '1' (i.e., ratio of cell resistances). The ratio of cell resistance (MTJ + NMOS) while reading '1' and reading '0' can vary from 1.26 to 1.13 (see Fig. 12). The distinguish ability can be further varied due to the variation in the leakage currents of the unselected cells. Depending on the prior read/write history of the unselected cells, we may have different leakage currents, as shown in Fig. 13. The maximum leakage is expected for redundant '0' write as the temperature is found to be maximum for this case.

The self-heating and the prior read/write history also modulates the read margin. We note that read disturb is only possible when reading '1' as flipping '0' is not possible by current flowing from BL to SL. Fig. 14 shows the read currents and read margins with the x-axis indicating the previous set of operations. We observe that the read margin is minimum when read '1' follows w0 → 1. This is attributed to the fact that w0 → 1 results in higher temperature compared with the other two cases of prior operations (i.e., R1 and w1 → 1). The higher temperature results in lower cell resistance and, hence, higher read current [see Figs. 12 and 14]. This is coupled with lower switching current at higher temperature [see Fig. 3(b)]. Therefore, the read margin is smallest among different cases.

V. Conclusion

We have analyzed the self-heating effect in the STTRAM cell and its impact on operational reliability. We have evaluated temperature distribution within the STTRAM cell using FVM-based thermal simulations. Our analysis shows that the high

write current and the small volume of the MTJ results in appreciable MTJ and silicon temperature due to self-heating. Finally, due to the self-heating and dependence of cell parameters on MTJ and silicon, there exist a correlation between read–write history of a cell and its operational reliability. Our analysis at the cell level suggests that self-heating can have strong impact on the reliable STTRAM operation and, hence, needs careful considerations.

REFERENCES

[1] A. Driskill-Smith, "STTRAM – A new spin on universal memory," Future Fab Int., Mill Valley, CA, Future Fab Int. Rep., Sep. 2008.

[2] T. Kawahara, R. Takemura, K. Miura, J. Hayakawa, S. Ikeda, Y. M. Lee, R. Sasaki, Y. Goto, K. Ito, T. Meguro, F. Matsukura, H. Takahashi, H. Matsuoka, and H. Ohno, "2 Mb SPRAM (Spin-Transfer Torque RAM) with bit-by-bit bi-directional current write and parallelizing-direction current read," *IEEE J. Solid-State Circuits*, vol. 43, no. 1, pp. 109–120, Jan. 2008.

[3] D. Halupka, S. Huda, W. Song, A. Sheikholeslami, K. Tsunoda, C. Yoshida, and M. Aoki, "Negative resistance based read and write scheme for STT-MRAM," in *Proc. IEEE ISSCC*, Feb. 2010, pp. 256–257.

[4] M. Hosomi, H. Yamagishi, T. Yamamoto, K. Bessho, Y. Higo, K. Yamane, H. Yamada, M. Shoji, H. Hachino, C. Fukumoto, H. Nagao, and H. Kano, "A novel nonvolatile memory with spin torque transfer magnetization switching: Spin-RAM," in *IEDM Tech. Dig.*, pp. 459–462.

[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 *IEDM Tech. Dig.*, 2009, pp. 1–4.

[6] X. Wang, Y. Chen, H. Li, D. Dimitrov, and H. Liu, "Spin torque random access memory down to 22 nm technology," *IEEE Trans. Magn.*, vol. 44, no. 11, pp. 2479–2482, Nov. 2008.

[7] D. Apalkov, Z. Diao, A. Panchula, S. Wang, Y. Huai, and K. Kawabata, "Temperature dependence of spin transfer switching in nanosecond regime," *IEEE Trans. Magn.*, vol. 42, no. 10, pp. 2685–2687, Oct. 2006.

[8] X. Liu, D. Mazumdar, W. Shen, B. D. Schrag, and G. Xiao, "Thermal stability of magnetic tunneling junctions with MgO barriers for high temperature spintronics," *Appl. Phys. Lett.*, vol. 89, no. 2, pp. 023504–023504-3, Jul. 2006.

[9] S. Chatterjee, S. Salahuddin, S. Kumar, and S. Mukhopadhyay, "Modeling of the self-heating in STTRAM and analysis of its impact on reliable memory operations," in *Proc. IEEE NVMTS*, 2010, pp. 86–89.

[10] S. Chatterjee, S. Salahuddin, S. Kumar, and S. Mukhopadhyay, "Analysis of thermal behaviors of spin-torque-transfer RAM: A simulation study," in *Proc. ACM/IEEE ISLPED2010*, pp. 13–18.

[11] S. Chatterjee, M. Rasquinha, S. Yalamanchili, and S. Mukhopadhyay, "A methodology for robust, energy efficient design of spin-torque-transfer RAM arrays at scaled technologies," in *Proc. ICCAD*, 2009, pp. 474–477.

[12] T. Gilbert, "A phenomenological theory of damping in ferromagnetic materials," *IEEE Trans. Magn.*, vol. 40, no. 6, pp. 3443–3449, Nov. 2004.

[13] S. Salahuddin, D. Datta, P. Srivastava, and S. Datta, "Quantum transport simulation of tunneling based Spin Torque Transfer (STT) devices: Design tradeoffs and torque efficiency," in *IEDM Tech. Dig.*, 2007, pp. 121–124.

[14] S. Salahuddin and S. Datta, "Self-consistent simulation of quantum transport and magnetization dynamics in spin-torque based devices," *Appl. Phys. Lett.*, vol. 89, no. 15, pp. 153504–153504-3, Oct. 2006.

[15] S. Salahuddin and S. Datta, "Electrical detection of spin excitations," *Phys. Rev. B, Condens. Matter Mater. Phys.*, vol. 73, no. 8, pp. 081301–081301-4, Feb. 2006.

[16] D. Datta, B. Behin-Aein, S. Salahuddin, and S. Datta, "Quantitative model for TMR and spin-transfer torque in MTJ devices," in *IEDM Tech. Dig.*, 2010, pp. 22.8.1–22.8.4.

[17] D. Mazumder, X. Liu, B. D. Schrag, W. Shen, M. Carter, and G. Xiao, "Thermal stability, sensitivity and noise characteristics of MgO based magnetic tunnel junctions," *J. Appl. Phys.*, vol. 101, no. 9, pp. 09B 502–09B 502-6, May 2007.

[18] [Online]. Available: http://www.synopsys.com/TOOLS/TCAD/DEVICESIMULATION/Pages/TaurusMedici.aspx

[19] B. Murmann, P. Nikaeen, D. J. Connelly, and R. W. Dutton, "Impact of scaling on analog performance and associated modeling needs," *IEEE Trans. Electron Devices*, vol. 53, no. 9, pp. 2160–2167, Sep. 2006.

Subho Chatterjee received the B.E. degree in electronics and telecommunication engineering from Jadavpur University, Calcutta, India, in 2004; the M.S. degree in electrical engineering from the Indian Institute of Technology, Kharagpur, India, in 2009; and the M.S.E.E degree from Georgia Institute of Technology, Atlanta, in 2010. He is currently working toward the Ph.D. degree with the School of Electrical and Computer Engineering, Georgia Institute of Technology.

In 2010, He worked as a Research Intern with the AMD Research and Advanced Development Laboratory. His research interests include nonvolatile memory devices, and low-power and variation-tolerant design techniques.

Sayeef Salahuddin received the B.Sc. degree in electrical and electronic engineering from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 2003 and Ph.D. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, in 2007.

In 2008, he joined the Faculty of Electrical Engineering and Computer Science, University of California, Berkeley. His research interests include the interdisciplinary field of electronic transport in nanostructures currently focusing on novel electronic and spintronic devices for low-power logic and memory applications. His research interests include the concept of using 'interacting systems' for switching and showing the fundamental advantage of such systems over the conventional devices in terms of power dissipation.

Dr. Salahuddin was the recipient of the Kintarul Islam Gold Medal from BUET in 2003, the Meissner fellowship from Purdue University for 2003–2004, an IBM Ph.D. Fellowship for 2007–2008, a MARCO/FCRP Inventor Recognition Award in 2007, a UC Regents Junior Faculty Fellowship in 2009, and a Hellman Faculty Fellowship in 2010.

Satish Kumar received the B.Tech. degree in mechanical engineering from the Indian Institute of Technology, Guwahati, India, in 2001; the M.S. degree in mechanical engineering from Louisiana State University, Baton Rouge, in 2003; and the M.S. degree in electrical and computer engineering and the Ph.D. degree in mechanical engineering from Purdue University, West Lafayette, IN, in 2007.

In 2009, he joined the George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology (Georgia Tech), Atlanta, as an Assistant Professor. Prior to joining Georgia Tech, he was with IBM Corporation, Austin, TX, where he was responsible for the thermal management of electronics devices. He is the author or co-author of over 50 journal and conference publications. His current research interests include thermal management, atomistic transport models for nanostructures, flexible electronics, and thermoelectric coolers.

Dr. Kumar was a recipient of the Purdue Research Foundation Fellowship in 2005.

Saibal Mukhopadhyay (S'99–M'07) received the B.E. degree in electronics and telecommunication engineering from Jadavpur University, Calcutta, India, in 2000 and the Ph.D. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, in 2006.

He is currently with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, as an Assistant Professor. Prior to joining Georgia Institute of Technology, he was with the IBM T. J. Watson Research Center, Yorktown Heights, NY, as a Research Staff Member. He is the author or coauthor of over 100 papers in refereed journals and conferences. He is a holder of five U.S. patents.His research interests include analysis and design of low-power and robust circuits in nanometer technologies and 3-D circuits and systems.

Dr. Mukhopadhyay was the recipient of the Best Paper Awards at the 2003 IEEE NANO and the 2004 International Conference on Computer Design, the IBM Ph.D. Fellowship Award for 2004–2005, the Best in Session Award at the SRC TECHCON and the SRC Technical Excellence Award in 2005, the SRC Inventor Recognition Award in 2008, the IBM Faculty Partnership Award in 2009 and 2010, and the National Science Foundation CAREER Award in 2011.