Yield Optimization of 1T-1MTJ STT-RAM Memory Cells

Richard Dorrance and Yuta Toriyama
{dorrance, yuta}@ee.ucla.edu

***Abstract*—We present a design-space feasibility region, as a function of magnetic tunnel junction (MTJ) characteristics and target memory specifications, to explore the design margin of a one-transistor-one-MTJ (1T-1MTJ) memory cell for spin-transfer torque random access memories (STT-RAMs). Data from measured devices is used to model the statistical variation of an MTJ's critical switching current and resistance. The sensitivity of the design space, to different design parameters, is also analyzed for scaling of both the MTJ and the underlying transistor technology. A design flow, leveraging a sensitivity-based analysis and a Landau-Lifshitz-Gilbert equation (LLGE) based MTJ switching model, is proposed to optimize design margins for gigabit-scale memories. Design points for improved yield, density, and memory performance are extracted from MTJ-compatible CMOS technologies for 90nm, 65nm, 45nm, and 32nm processes.**

***Index Terms*—Magnetic Tunnel Junction (MTJ), Magnetoresistive Random Access Memory (MRAM), Spin-Transfer Torque (STT), Spin-Transfer Torque Random Access Memory (STT-RAM), Variability, PVT.**

I. Introduction

MAGNETORESISTIVE Random Access Memories (MRAMs) have attracted a significant amount of interest as a commercially viable universal memory technology. With the density of DRAM, the speed of SRAM, and the non-volatility of Flash it is easy to see why [1]. MRAMs require zero standby power and boast a nearly unlimited programming endurance ($> 10^{15}$ cycles) [2]. Such a memory would eliminate the need for multiple application-specific memories, improving system performance and reliability, while also lowering cost and power consumption in everything from mobile devices to datacenters [3] (see Fig. 1).

The non-volatile storage element of an MRAM is the Magnetic Tunnel Junction (MTJ). Structurally, an MTJ is a pair of ferromagnets separated by a thin insulating layer. Data storage is achieved by exploiting the magnetic orientation of these ferromagnetic layers [4]. Only two magnetic states are stable: the parallel combination (Fig. 2(a)) and the antiparallel combination (Fig. 2(b)). The parallel configuration leads to a low resistive state (R_P), while the antiparallel configuration leads to a high resistive state (R_{AP}). Tunnel magnetoresistance (TMR), the ratio of the difference between R_P and R_{AP}, is a metric for determining the efficiency of the spintronic operation of an MTJ [5]. TMR is defined as:

$$TMR = \frac{R_{AP} - R_P}{R_P}. \quad (1)$$

Spin-transfer torque (STT) based switching has been the primary method to exploit the magnetic hysteresis of MTJs.

	SRAM	DRAM	Flash (NOR)	Flash (NAND)	MRAM	STT-RAM
Non-volatile	No	No	Yes	Yes	Yes	Yes
Cell Size [F^2]	50-120	6-10	10	5	16-40	6-20
Read Time [ns]	1-100	30	10	50	3-20	2-20
Write/Erase Time [ns]	1-100	15	1µs/1ms	1ms/0.1ms	3-20	2-20
Endurance	10^{16}	10^{16}	10^5	10^5	$>10^{15}$	$>10^{15}$
Write Power	Low	Low	Very High	Very High	High	Low
Other Power Consumption	Leakage	Refresh	None	None	None	None
High Voltage Required	No	3V	6-8V	16-20V	3V	<1.5V

Fig. 1. Performance of various memory technologies (Source: Wolf et al. [3])

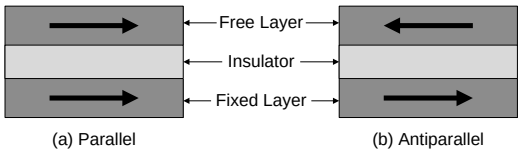

Fig. 2. MTJ ferromagnetic layers in (a) parallel and (b) antiparallel configurations.

Rather than using an indirect current to generate a magnetic field, STT uses a spin-polarized current through the MTJ to accomplish device switching [6]. Toggling of the MTJ is roughly determined by the current density [7]. As the area of the MTJ device decreases, so does the writing current. Spin-Transfer Torque Random Access Memories (STT-RAMs) have the added benefit of being architecturally much simpler than conventional MRAMs [8]. The simplest of STT-RAM architectures uses the one-transistor-one-MTJ (1T-1MTJ) structure.

Despite the importance of the 1T-1MTJ structure for the future success of STT-RAM, very little comprehensive analysis has been done on the subject. Analysis in work by Raychowdhury et al. [9], [10] considers MTJs, but not the underlying transistor technology. In fact, the design of the MTJ and the access transistor are intertwined. A given CMOS technology constrains the design space of the MTJ due to the overhead and impact of the access transistor in each memory cell. This, in turn, affects the performance of the MTJ, which further impacts the design of the access transistor. Ono et

TABLE I
MEASURED DEVICE STATISTICS

	X	Y	Z
TMR [%]	105.7	107.3	105.3
σ_{TMR} [%]	4.7	2.7	4.6
RA [Ω·μm^2]	4.88	5.51	5.22
σ_{RA} [Ω·μm^2]	0.342	0.297	0.311

al. [11] used a stochastic MTJ model, later verified with on-chip measurements, to optimize the design of a 32Mbit test chip in the presence of asymmetric access transistor behavior. Similarly, Chen et al. [12] discuss how a statistical model for the MTJ, which ignores the role of the access device, produces a suboptimal memory cell in both area and yield. Furthermore, the feasibility and yield of the memory depend on the design space and the variation of the MTJs [13].

In this work, we present a comprehensive analysis of the design space of a 1T-1MTJ memory cell for STT-RAMs. We use a precessional-based switching model, modified to include thermally-activated switching, to capture the dynamic nature of the MTJ. The effects of both CMOS and MTJ device variability across Process-Voltage-Temperature (PVT), which is notably absent in prior works, are demonstrated with our analysis.

II. MODELING MTJ VARIABILITY AND SCALING

This section describes the MTJ model and characteristics that are used in the subsequent sections to explore the design space for several scaled CMOS technologies.

A. MTJ Device Variability

While statistical variation of CMOS is generally well understood, similar characteristics for MTJs have not been well documented. This work uses a combination of fundamental equations and measured device characteristics to model the statistical behavior of MTJs.

1) Resistance: Variations in MTJ resistance and TMR are due to the small geometric differences between fabricated nanopillars. These typically arise from a combination of lithographic variations in the physical dimensions of the nanopillar, as well as minute fluctuations in the thicknesses of the up to 20 different layers in state-of-the-art MTJ processes [14]. Fig. 3(a) contains a plot of measured R_P vs. R_{AP} for 105 MTJ nanopillars of varying size and target resistance-area (RA) products. The cumulative effects of random geometric variation on MTJ resistance can be condensed into random Gaussian variation on RA and TMR [15]. Fig. 3(b) and Table I show the calculated statistics for our MTJ nanopillars. Variation on TMR is on the order of 3-5% and variation in RA is on the order of 0.3Ωμm^2.

2) Switching Current: Variation in the MTJ critical switching current is the result of two different mechanisms. The first is thermal agitation, which leads to probabilistic switching in MTJ nanopillars at finite temperatures [16]. An example of this probabilistic switching behavior can be seen in Fig. 4(a)

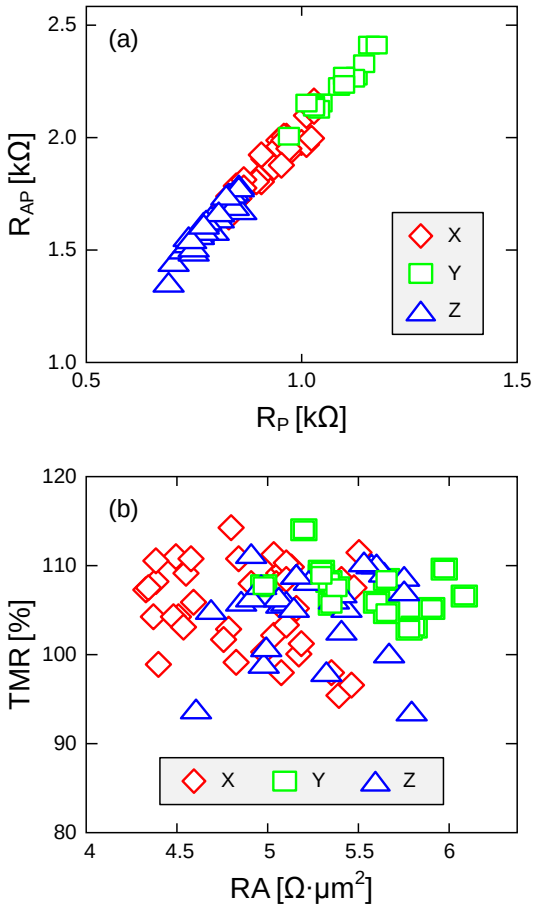

Fig. 3. Measured (a) R_{AP} vs. R_P and (b) TMR vs. RA for MTJ nanopillars measuring 150 × 45nm^2 (X), 130 × 50nm^2 (Y), and 170 × 45nm^2 (Z) at room temperature (300K).

for several MgO-based MTJ nanopillars. The second cause of switching current variation in MTJs is due to process related geometric variation [17]. The effects of geometric variation are clearly evident in Fig. 4(a) as the varying offsets between the probability of switching curves for each MTJ. The general shape of the probability of switching curve for an MTJ has been shown, both theoretically and experimentally, to depend upon the thermal stability (Δ) of the MTJ [18], [19].

To study the effects that geometric variation has on the critical switching current, measurements were obtained from MTJ nanopillars that have purposefully been fabricated with large geometric variation. In these measurements, several critical layers in the MTJ were deposited as a wedge, with their thickness systematically varying by several nanometers from chip edge to chip edge. The resulting induced geometric variation is more than 10 times greater than typical random process variation. A strong correlation ($\rho = -0.929$) was found to exist between the RA and the switching current of each device (see Fig. 4(b)). This allowed us to use fewer devices to measure the statistical variation of the critical switching current. Based on device measurements, the σ of the switching current due to geometric variation was estimated to be 7μA, or about 2% of the critical switching current. This is in good agreement with measurements from Driskill-Smith et

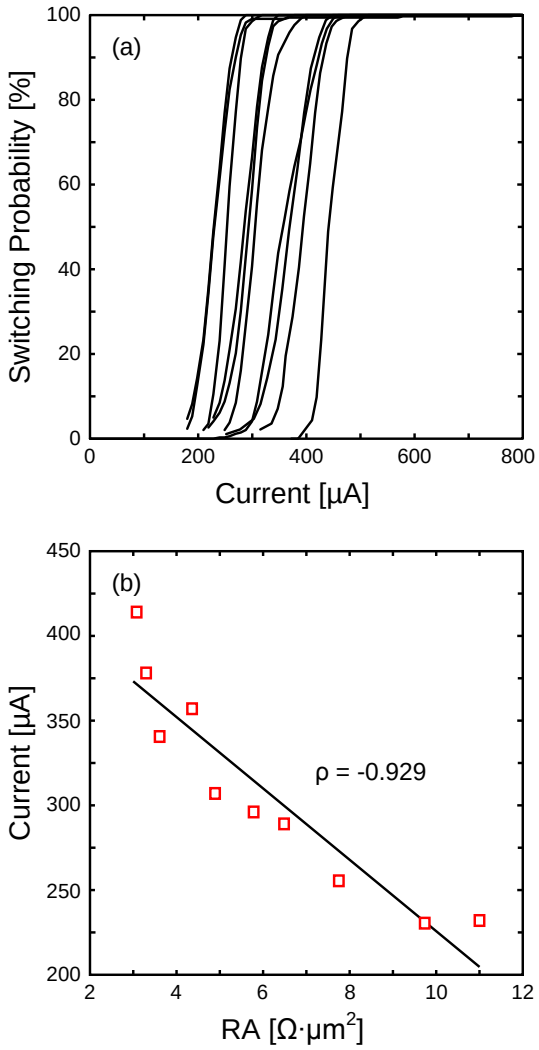

Fig. 4. Measurements of (a) probability of switching vs. current and (b) RA vs. current (at 50% switching probability) for MTJ nanopillars measuring $135 \times 65\text{nm}^2$.

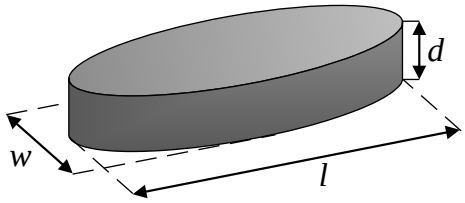

Fig. 5. MTJ free layer dimensions.

al. (3% variation) [20], Huai et al. (3% variation) [16], and Pakala et al. (3.5% variation) [18].

B. Scaling of MTJ Current and Resistance

The resistance and switching current can be modeled using a precessional-based switching model, modified to include thermally-activated switching [21]. The switching current of an MTJ in the precessional region, for a constant pulse of duration τ, is given by:

$$I_C = I_{C0}\left[1 - \frac{\ln(\tau/\tau_0)}{\Delta}\right], \tag{2}$$

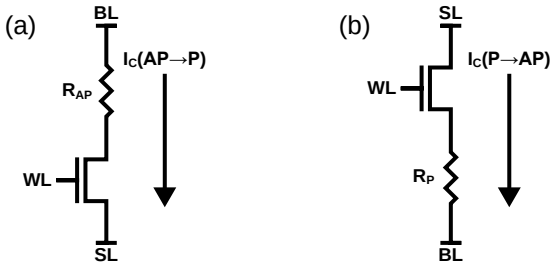

Fig. 6. 1T-1MTJ cell architecture showing MTJ switching current for (a) AP to P and (b) P to AP.

where τ_0 is the natural time constant and I_{C0} is the critical switching current. This critical switching current [22] is given by:

$$I_{C0} = \frac{\alpha 4\pi e}{\eta\hbar} M_S^2 V, \tag{3}$$

where α is the Gilbert damping constant, η is the factor of spin polarization, $\hbar$ is the reduced Planck's constant, e is the elemental charge of an electron, M_S is the magnetization saturation of the free layer, and V is the volume of the free layer.

For an MTJ with free layer dimensions $l > w >> d$ [23], [24], as shown in Fig. 5, the thermal stability of an MTJ is approximately:

$$\Delta = \frac{E}{k_B T} = \frac{H_K M_S}{2k_B T} V \approx d\left(\frac{1}{w} - \frac{1}{l}\right)\frac{M_S^2}{k_B T} V, \tag{4}$$

where k_B is Boltzmann's constant, T is the absolute temperature in Kelvin, H_K is the out-of-plane uniaxial anisotropy, and E is the energy of anisotropy [25], [26].

In this work, dimensional scaling is performed to maintain a constant Δ in order to ensure the long-term non-volatility of the STT-RAM. Therefore, dimensions l and w of the MTJ are scaled by a factor λ to manipulate I_{C0} and $R_{P/AP}$, then to keep Δ constant, d must scale by $\lambda^{-1/2}$. This results in $I_{C0} \propto lwd \rightarrow \lambda^{3/2}$ and $R_{P/AP} \propto l^{-1}w^{-1} \rightarrow \lambda^{-2}$.

III. Design Space

The analysis in this work uses a conventional 1T-1MTJ cell architecture as shown in Fig. 6. The minimum writing currents, to ensure a target write error rate (WER), for flipping the cell resistance are defined as $I_C(P \rightarrow AP)$ and $I_C(AP \rightarrow P)$. The design space of a single STT-RAM memory cell can be illustrated using an R_{AP} vs. R_P plot as shown in Fig. 7. The feasibility region is indicated by the shaded region. It contains all points (R_P, R_{AP}) in the design space so that a memory cell made with such an MTJ is functional. In the design space, the two lower bounds are set by the read margin of the cell, while the two upper bounds are set by the write margin of the cell.

The lower bound $R_{P,MIN}$ is dependent on the implementation of the sense amplifier, and represents the minimum resistance required for reliable circuit operation. Parasitic resistances from the access transistor and column-mux, as well as the bit- and source-lines, all contribute to $R_{P,MIN}$. Additionally, $R_{AP,MIN}$ is determined by TMR_{MIN} (Fig.

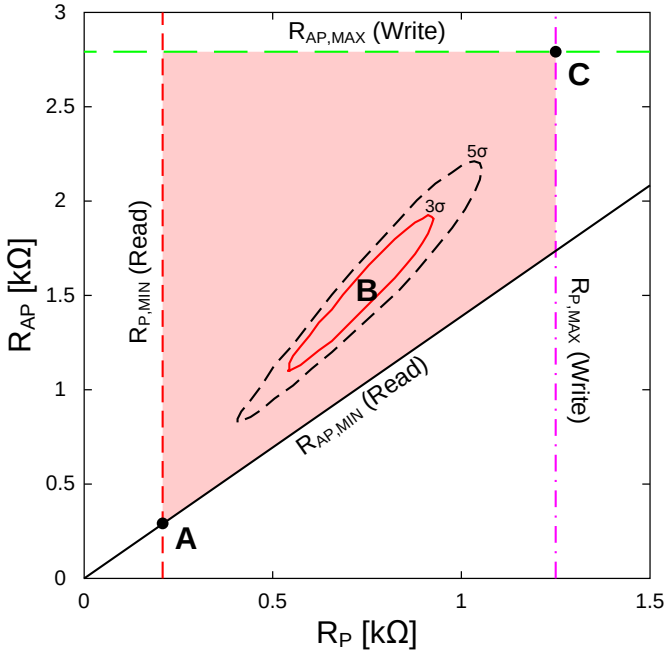

Fig. 7. Design space, in a 65nm process, for $W_N = 2.0\mu m$, $I_C(P \rightarrow AP) = 500\mu A$, $I_C(AP \rightarrow P) = 375\mu A$, with an overlay of device X from Table I.

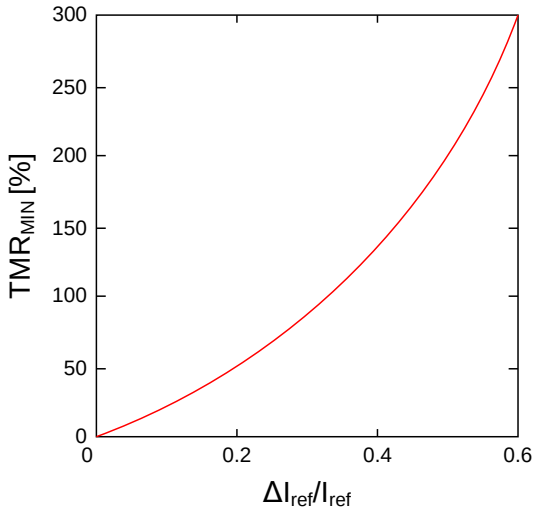

Fig. 8. Design space lower bound TMR_{MIN} vs. $\Delta I_{ref}/I_{ref}$ for a current-sensing read circuit with ideal reference resistance $2(R_P \parallel R_{AP})$.

8), the minimum TMR required for the read amplifier to differentiate between R_P and R_{AP}. Regardless of the specifics of the implementation, all sense amplifiers are either a voltage- or current-sensing topology. For a generic current-sensing read circuit, a read margin of ΔI_{ref} results in:

$$\text{(Current) } TMR_{MIN} = \frac{2\Delta I_{ref}/I_{ref}}{1 - \Delta I_{ref}/I_{ref}}. \quad (5)$$

For I_{ref} flowing through the reference resistance R_{ref}, $I_{ref} + \Delta I_{ref,1}$ flows through R_P and $I_{ref} - \Delta I_{ref,2}$ through R_{AP}. When $\Delta I_{ref,1} = \Delta I_{ref,2} = \Delta I_{ref}$, TMR_{MIN} is minimized. Under this condition, $R_{ref} = 2(R_P \parallel R_{AP})$ and we can express TMR_{MIN} as a function of the normalized

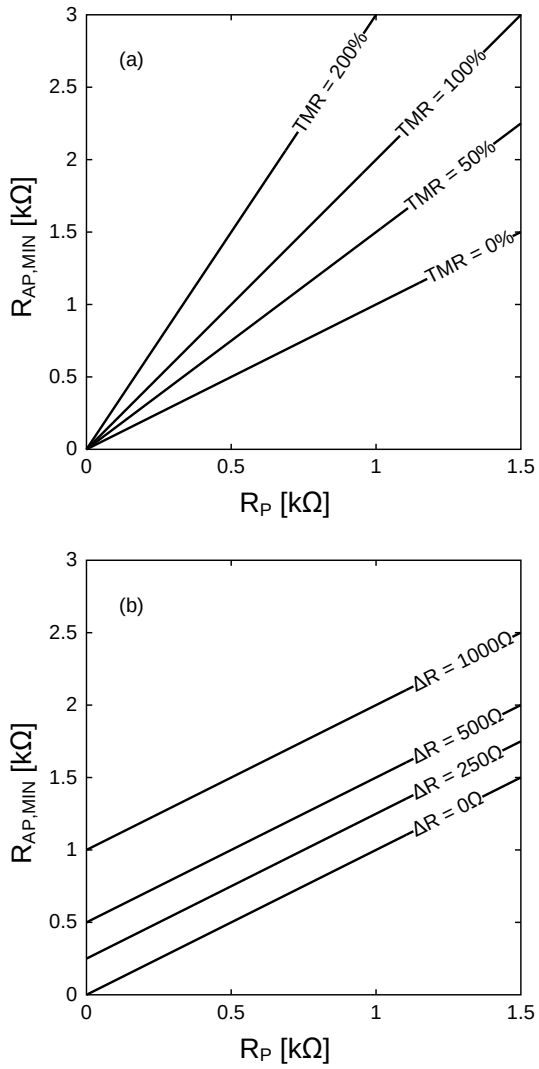

Fig. 9. Difference in $R_{AP,MIN}$ between (a) current sensing and (b) voltage sensing.

fractional sensing current ($\Delta I_{ref}/I_{ref}$). In Eq. 5, ΔI_{ref} must be chosen so that the read amplifier correctly evaluates across all transistor PVT variations.

Similarly, TMR_{MIN} for a generic voltage-sensing topology is:

$$\text{(Voltage) } TMR_{MIN} = \frac{2\Delta V}{R_P I_{ref}} = \frac{\Delta R}{R_P}, \quad (6)$$

where ΔV, the voltage reading margin, is the minimum difference in sensing voltage between the MTJ and the reference resistance, and $\Delta R = 2\Delta V/I_{ref}$ is the minimum difference in resistance between R_P and R_{AP}. The difference between voltage- or current-sensing topologies can be seen in Fig. 9. Voltage-sensing is better suited for devices with larger RAs, where a small TMR can still translate into a large resistance difference. Alternatively, current-sensing topologies are better able differentiate low RA MTJs. It should be noted that the lower bounds $R_{P,MIN}$ and $R_{AP,MIN}$, while critical to the readability of the cell, are almost completely independent of the MTJs used. The only requirement is that sensing time and

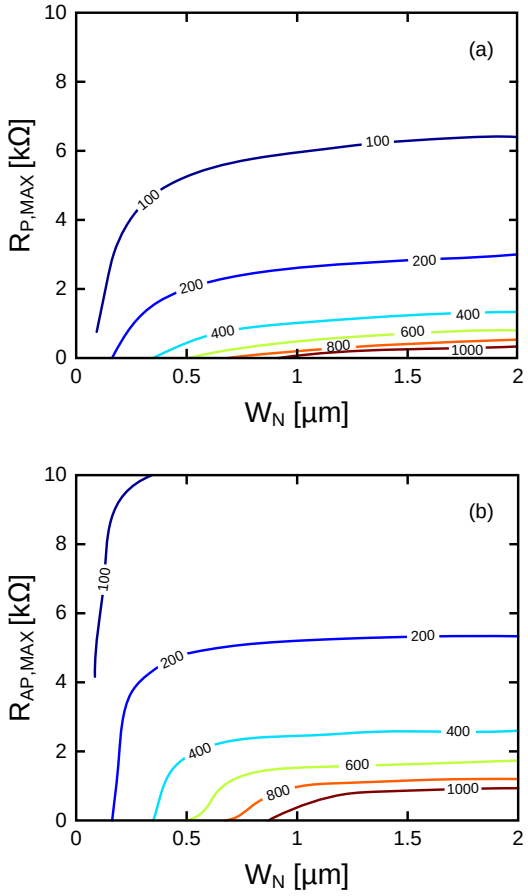

Fig. 10. (a) $R_{P,MAX}$ and (b) $R_{AP,MAX}$ at nominal V_{DD} for a 65nm process (I_C contours are in µA).

current (I_{ref}) be small enough so as not to disturb the cell during the read operation.

The upper bounds, $R_{P,MAX}$ and $R_{AP,MAX}$, are the maximum allowable resistances such that the access transistor, in a 1T-1MTJ configuration, is still able to provide the minimum critical writing currents $I_C(P \rightarrow AP)$ and $I_C(AP \rightarrow P)$. These upper bounds are subsequently very sensitive to the specific characteristics of the MTJ device used. As such, to ensure a sufficiently low WER, the effects of stochastic thermal fluctuations [27], self-induced heating [28], and backhopping [29] on the probability of switching should not be overlooked. Transistor-level simulations are used to determine the relationship between $R_{P/AP,MAX}$, I_C, and cell size (transistor width W_N) for a technology. Fig. 10 shows an example of such a simulation in a 65nm process. Using the conventional configuration from Fig. 6, W_N is swept along with R_{MAX}. The contours of the simulated current are shown.

Fig. 7 shows a specific MTJ cell and its associated statistical variation (the concentric ovals around point B) overlaid on the design space. The design-space margin can be defined as the number of σ's of MTJ variation before crossing any of the previously defined bounds. Defining design-space margin (DSM) in terms of σ simplifies feasibility characterization to a single variable and thus allows yield to be quickly calculated. To a first order, 3σ, 4σ, 5σ, and 6σ of design margin roughly

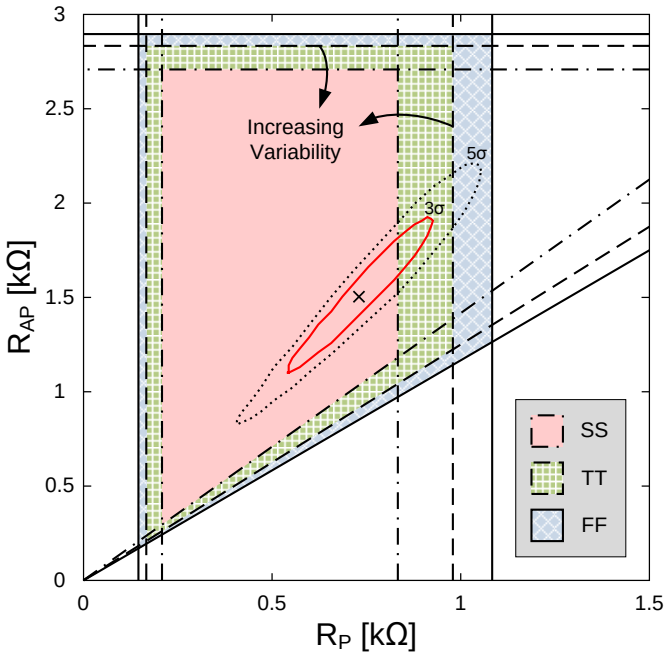

Fig. 11. Design space, in a 65nm process, for $W_N = 750nm$, $I_C(P \rightarrow AP) = 300$µA, $I_C(AP \rightarrow P) = 300$µA, with an overlay of device X from Table I, for SS, TT, and FF corners.

correspond to being able to reliably produce 1kbit, 32kbit, 4Mbit, and 1Gbit memory arrays.

Fig. 11 highlights the effects of CMOS variability on the design space bounds. To illustrate the effects more clearly, a $35F^2$ cell in a 65nm process is chosen. As expected, the more stringent constraints of the SS corner causes the design space to shrink. This shift is caused by an increase in the threshold voltage of the access transistor. Environmental variables, like temperature, also have a significant effect on the design space. A consumer grade STT-RAM is expected to operate over a range of more than 100°C in which TMR can drop by more than 30% [30], degrading the DSM for readability. These sources of technological and environmental variability must also be considered in the design process.

IV. 1T-1MTJ Cell Optimization Using a Sensitivity Analysis

Many variables, at both the circuit and device levels, affect the design space. In order to optimize all variables for a target memory specification, we must determine how each variable impacts the design space. This section introduces a design-space sensitivity (DSS) as a metric to quantify the behavior of the change in design space as a function of various design parameters (V_{DD}, λ, J_C, RA, TMR, W_N, etc.). We then present a sensitivity-based design flow that leverages DSS to optimize the DSM of a 1T-1MTJ memory cell. A short design example using a 65nm CMOS technology is provided.

A. Design-Space Sensitivity Analysis

First consider the points A, B, and C in Fig. 7. Points A and C correspond to the corner values of R_P and R_{AP} in the

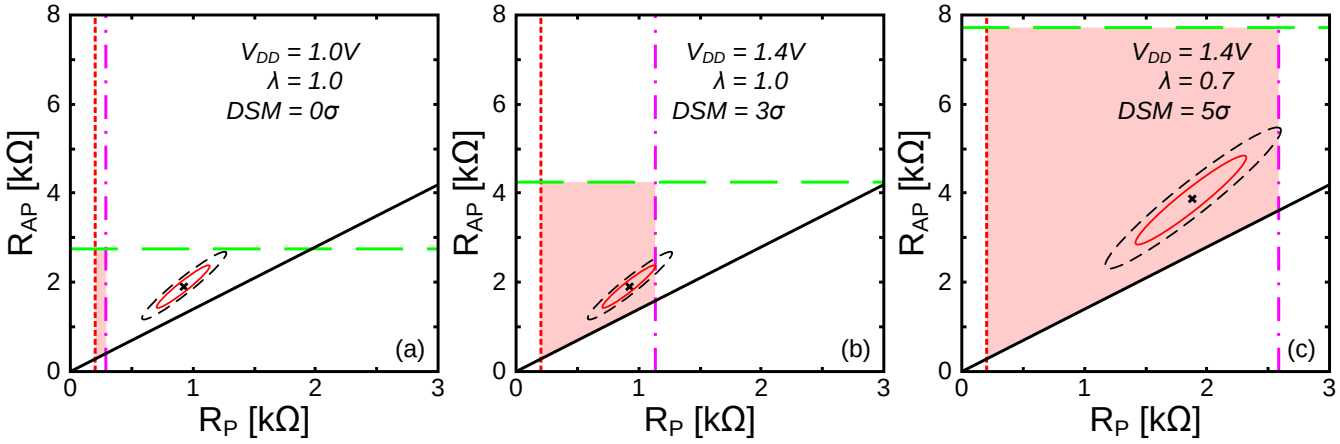

Fig. 12. Design space, in a 65nm process, for a $30F^2$ cell ($W_N = 0.65$µm) for device X from Table I: $I_C(P \rightarrow AP) = 450$µA, $I_C(AP \rightarrow P) = 300$µA. Inner red oval represents 3σ of MTJ device variation. Dashed, black oval corresponds to 5σ of MTJ variation.

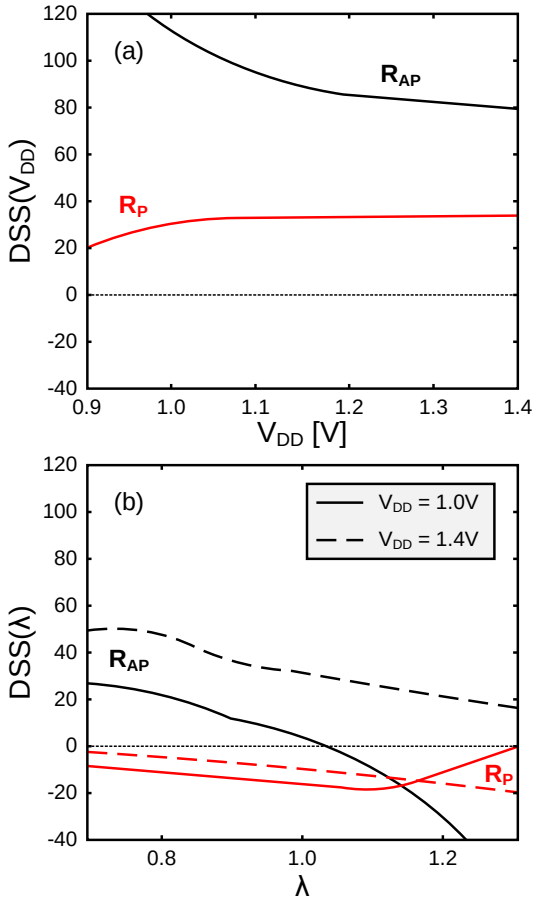

Fig. 13. Design-space sensitivity of parameters (a) V_{DD} and (b) λ in a 65nm technology.

feasible design space. Point B represents the nominal MTJ at the center of the MTJ device distribution. For a positive design margin to exist, point B must fall somewhere between points A and C.

A "better" design space can be achieved from altering a design parameter, when a larger distribution of the MTJs (the number of σ) falls within the feasible region. Note that the improved design space is not simply increasing the area of the feasibility region, but increasing the number of sigma enclosed by the feasible region. Recall that point A depends only slightly on the MTJ parameters. Therefore, the improvement (or deterioration) of the design space depends mostly on the change in DSM between points B and C as a function of a particular design variable.

Therefore, we define the design-space sensitivity to the parameter X as:

$$DSS(X) = \frac{\partial(\frac{R_C - R_B}{\sigma})_{P/AP}}{\partial X}, \tag{7}$$

where R_B and R_C are taken as either R_P or R_{AP} at points B and C, thus defining the DSS along each dimension of the design space. $\frac{R_C - R_B}{\sigma}$ is the normalized distance between points B and C in the design space along the $R_{P/AP}$ dimension. Intuitively, the $DSS(X)$ describes the instantaneous rate of change in DSM to a particular design parameter X. The derivative loses positional information, and so we used the DSS in conjunction with the original plot of the design space to determine the benefit of tuning the design parameter X. For both the R_P and R_{AP} dimensions, if $DSS(X) > 0$, then the DSM is improved by increasing X, and if $DSS(X) < 0$, then DSM is improved by decreasing X. When the design-space sensitivities for the two dimensions conflict, the DSM in each dimension should then be taken into account.

B. Design Example

In this section we use the sensitivity analysis to design a 4Mbit STT-MRAM with a $30F^2$ cell size (comparable to eDRAM) in a 65nm technology. Device X from Table I with $I_C(P \rightarrow AP) = 450$µA and $I_C(AP \rightarrow P) = 300$µA is the nominal MTJ and can be scaled by λ. Also, approximately 5σ of design margin is required for reasonable yield.

Fig. 12(a) shows the design space for a nominal $V_{DD} = 1.0V$ and $\lambda = 1.0$. The inner red oval is the 3σ variation of the MTJ, while the dashed, black oval represents the 5σ variation of the MTJ. Clearly, with nominal V_{DD} and λ, the memory is not functional. Fig. 13 shows that the design space

is much more sensitive to V_{DD} than it is to λ. Therefore, we choose to scale V_{DD} to $1.4V$. It should be noted that at $1.4V$, most of the voltage is dropped across the MTJ, leaving the V_{GS} and V_{DS} of the access transistor below $1V$. Fig. 12(b) shows the new design space, with the 3σ bound at the edge of the design boundary.

Scaling V_{DD} alone proves insufficient to meet the 5σ design margin required, and so we simultaneously scale λ. Fig. 13(b) shows that scaling λ results in conflicting DSS. The R_{AP} margin improves more by scaling λ up, while the R_P margin improves by scaling λ down. However, Fig. 12(b) indicates that R_{AP} dimension has considerable margin and we can trade off some of that margin for improved margin in R_P. Therefore, we choose to scale λ down to 0.7. As we can see in Fig. 12(c), the desired 5σ bound on MTJ variation is essentially enclosed within the design space.

V. Conclusion

We have introduced a modeling methodology to accurately model the device behavior of MTJs, which forms the basis of our work. In this work we have shown that the joint optimization of multiple design parameters is essential in the design of an STT-RAM memory array. We have also derived the necessary framework to allow for such a systematic design procedure.

Acknowledgements

We would like to thank Pedram Khalili and Juan G. Alzate of UCLA for providing measured device data for fabricated MTJ nanopillars.

References

[1] B. Cockburn, "The Emergence of High-Density Semiconductor-Compatible Spintronic Memory," in *Proc. Int. Conf. MEMS, NANO and Smart Systems 2003*, July 2003, pp. 321–326.

[2] S. Tehrani *et al.*, "Magnetoresistive Random Access Memory Using Magnetic Tunnel Junctions," *Proc. IEEE*, vol. 91, no. 5, pp. 703–714, May 2003.

[3] S. A. Wolf *et al.*, "The Promise of Nanomagnetics and Spintronics for Future Logic and Universal Memory," *Proc. IEEE*, vol. 98, no. 12, pp. 2155 –2168, Dec. 2010.

[4] M. E. Flatte, "Spintronics," *IEEE Trans. Electron Devices*, vol. 54, no. 5, pp. 907–920, May 2007.

[5] J. Z. Sun, "Spin Angular Momentum Transfer in Current-Perpendicular Nanomagnetic Junctions," *IBM J. Res. Dev.*, vol. 50, no. 1, pp. 81–100, Jan. 2006.

[6] J. C. Slonczewski, "Current-Driven Excitation of Magnetic Multilayers," *J. Magn. Magn. Mater.*, vol. 159, no. 1-2, pp. L1–L7, 1996.

[7] E. Chen *et al.*, "Advances and Future Prospects of Spin-Transfer Torque Random Access Memory," *IEEE Trans. Magn.*, vol. 46, no. 6, pp. 1873–1878, Jun. 2010.

[8] M. Hosomi *et al.*, "A Novel Nonvolatile Memory with Spin Torque Transfer Magnetization Switching: Spin-RAM," in *IEDM 2005*, Dec. 2005, pp. 459–462.

[9] A. Raychowdhury *et al.*, "Design Space and Scalability Exploration of 1T-1STT MTJ Memory Arrays in the Presence of Variability and Disturbances," in *IEDM 2009*, Dec. 2009, pp. 1–4.

[10] A. Raychowdhury *et al.*, "Modeling and Analysis of Read (RD) Disturb in 1T-1STT MTJ Memory Bits," in *Device Res. Conf. (DRC), 2010*, Jun. 2010, pp. 43–44.

[11] K. Ono *et al.*, "A Disturbance-Free Read Scheme and a Compact Stochastic-Spin-Dynamics-Based MTJ Circuit Model for Gb-scale SPRAM," in *Electron Devices Meeting (IEDM), 2009 IEEE International*, Dec. 2009, pp. 1–4.

[12] Y. Chen *et al.*, "Design Margin Exploration of Spin-Transfer Torque RAM (STT-RAM) in Scaled Technologies," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 18, no. 12, pp. 1724–1734, Dec. 2010.

[13] M. Motoyoshi *et al.*, "A study for 0.18 μm high-density MRAM," in *2004 Symp. VLSI Tech. (VLSIT)*, Jun. 2004, pp. 22–23.

[14] S. R. Min *et al.*, "Etch Characteristics of Magnetic Tunnel Junction Stack with Nanometer-Sized Patterns for Magnetic Random Access Memory," *DPS 2006*, vol. 516, no. 11, pp. 3507–3511, Nov. 2008.

[15] R. Beach *et al.*, "A Statistical Study of Magnetic Tunnel Junctions for High-Density Spin Torque Transfer-MRAM (STT-MRAM)," in *IEDM 2008*, Dec. 2008, pp. 1–4.

[16] Y. Huai *et al.*, "Spin-Transfer Switching Current Distribution and Reduction in Magnetic Tunneling Junction-based Structures," *IEEE Trans. Magn.*, vol. 41, no. 10, pp. 2621–2626, Oct. 2005.

[17] Y. Katoh *et al.*, "Analysis of MTJ Edge Deformation Influence on Switching Current Distribution for Next-Generation High-Speed MRAMs," *IEEE Trans. Magn.*, vol. 45, no. 10, pp. 3804–3807, Oct. 2009.

[18] M. Pakala *et al.*, "Critical Current Distribution in Spin-Transfer-Switched Magnetic Tunnel Junctions," *J. Appl. Phys.*, vol. 98, no. 5, p. 056107, 2005.

[19] A. Driskill-Smith *et al.*, "Non-volatile Spin-Transfer Torque RAM (STT-RAM): Data, Analysis and Design Requirements for Thermal Stability," in *2010 Symp. VLSI Tech. (VLSIT)*, Jun. 2010, pp. 51–52.

[20] A. Driskill-Smith *et al.*, "Non-volatile Spin-Transfer Torque RAM (STT-RAM): An Analysis of Chip Data, Thermal Stability and Scalability," in *IEEE Int. Memory Workshop (IMW) 2010*, May 2010, pp. 1–3.

[21] T. Moriyama *et al.*, "Tunnel Magnetoresistance and Spin Torque Switching in MgO-based Magnetic Tunnel Junctions with a Co/Ni Multilayer Electrode," *Appl. Phys. Lett.*, vol. 97, no. 7, p. 072513, 2010.

[22] J. Z. Sun, "Spin-Current Interaction with a Monodomain Magnetic Body: A Model Study," *Phys. Rev. B*, vol. 62, no. 1, pp. 570–578, July 2000.

[23] H. Chang and J. Burns, "Demagnetizing and Stray Fields of Elliptical Films," *J. Appl. Phys.*, vol. 37, no. 8, pp. 3240–3245, July 1966.

[24] J. A. Osborn, "Demagnetizing Factors of the General Ellipsoid," *Phys. Rev.*, vol. 67, no. 11-12, pp. 351–357, Jun. 1945.

[25] V. Korenivski and R. Leuschner, "Thermally Activated Switching in Nanoscale Magnetic Tunnel Junctions," *IEEE Trans. Magn.*, vol. 46, no. 6, pp. 2101–2103, Jun. 2010.

[26] J. Sun and D. Ralph, "Magnetoresistance and Spin-Transfer Torque in Magnetic Tunnel Junctions," *J. Magn. Magn. Mater.*, vol. 320, no. 7, pp. 1227–1237, 2008.

[27] T. Devolder *et al.*, "Single-Shot Time-Resolved Measurements of Nanosecond-Scale Spin-Transfer Induced Switching: Stochastic Versus Deterministic Aspects," *Phys. Rev. Lett.*, vol. 100, p. 057206, Feb. 2008.

[28] Y. Zhang *et al.*, "Micromagnetic Study of Hotspot and Thermal Effects on Spin-Transfer Switching in Magnetic Tunnel Junctions," *J. Appl. Phys.*, vol. 101, no. 10, p. 103905, 2007.

[29] J. Z. Sun *et al.*, "High-Bias Backhopping in Nanosecond Time-Domain Spin-Torque Switches of MgO-based Magnetic Tunnel Junctions," *J. Appl. Phys.*, vol. 105, no. 7, p. 07D109, 2009.

[30] K. Lee and S. Kang, "Design Consideration of Magnetic Tunnel Junctions for Reliable High-Temperature Operation of STT-MRAM," *IEEE Trans. Magn.*, vol. 46, no. 6, pp. 1537–1540, Jun. 2010.