

# High-Density Spin–Orbit Torque Magnetic Random Access Memory With Voltage-Controlled Magnetic Anisotropy/Spin-Transfer Torque Assist

PIYUSH KUMAR<sup>ID</sup> (Graduate Student Member, IEEE)  
and AZAD NAEEMI<sup>ID</sup> (Senior Member, IEEE)

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA

CORRESPONDING AUTHOR: P. KUMAR (pkumar315@gatech.edu)

This work was supported by the Applications and Systems driven Center for Energy-Efficient Integrated NanoTechnologies (ASCENT), one of six centers in Joint University Microelectronics Program (JUMP), an Semiconductor Research Corporation (SRC) program through Defense Advanced Research Projects Agency (DARPA).

**ABSTRACT** This article explores an area saving scheme for spin–orbit torque (SOT) magnetic random access memory (MRAM) by sharing the SOT channel and write transistor among multiple magnetic tunnel junctions (MTJs). We use two write mechanisms to selectively write the MTJs, i.e., voltage-controlled magnetic anisotropy (VCMA)-assisted write in the presence of an external magnetic field and field-free spin-transfer torque (STT)-assisted write. Using micromagnetic simulations that are augmented by the rare-event enhancement, we study various trade-offs among write current, time, and energy, write error rate (WER), and the number of MTJs on an SOT channel. We quantify the issue of IR drop on the SOT channel as a function of the SOT layer thickness and number of MTJs. Our results show having more than four MTJs on an SOT channel poses major challenges in terms of IR drop and WER. In addition, we evaluate the impact of the proposed scheme on read performance.

**INDEX TERMS** Magnetic random access memory (MRAM), spin–orbit torque (SOT), spin-transfer torque (STT), voltage-controlled magnetic anisotropy (VCMA).

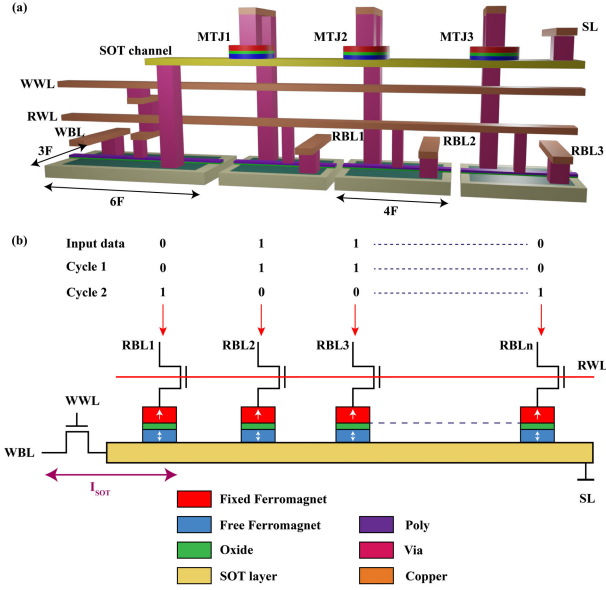
## I. INTRODUCTION

**S**PINTRONIC memories are being actively pursued for various applications, such as last-level cache [1], [2], embedded memory [3], and deep neural networks [4]. Spin-transfer torque (STT)-based and spin–orbit torque (SOT)-based magnetic random access memories (MRAMs) are two major examples of spintronic memories being explored. The STT-MRAM offers high cell density due to compact cell requiring only one transistor; however, it suffers from issues, such as low read margin, low charge to spin conversion efficiency, and oxide degradation. Moreover, large write current needed for STT-MRAM poses a challenge in terms of scaling. The SOT-MRAM is an emerging alternative for STT-MRAM. The SOT-MRAM has lower write energy while also improving the read operation by decoupling the read and write paths. There have been major advances in large-scale adoption of the SOT-MRAM technology in recent years. For example, wafer-level integration along with sub-nanosecond

magnetization switching has been demonstrated [5]. However, one key issue with SOT-MRAM is the large cell area compared with STT-MRAM, as SOT-MRAM requires two separate transistors for read and write operations.

There have been works on reducing the cell footprint of SOT-MRAM by sharing the SOT channel among multiple magnetic tunnel junctions (MTJs) with the help of STT [6] or voltage-controlled magnetic anisotropy (VCMA) effect [7], [8]. However, such schemes would require many trade-offs and a detailed evaluations of such schemes that proves low write error rates (WERs) and adequate selectivity accounting for thermal noise, variability, and the IR drop on the SOT layer are missing. Likewise, the potential impact of such schemes on write/read energy and latency as a function of cell density is also lacking.

In this article, we discuss transistor sharing schemes for SOT-MRAM with the help of VCMA effect and STT while considering the limitations in terms of WER and IR drop in



**FIGURE 1. (a)** 3-D layout of the memory cell demonstrating SOT channel sharing among multiple MTJs. The cell area is:  $6F \times 3F + N_{\text{MTJ}} (4F \times 3F)$ , with half-metal pitch ( $F$ ) of 32 nm and  $N_{\text{MTJ}}$  being the number of MTJs on the shared SOT channel. An extra  $2F$  width is required for routing sourceline (SL). **(b)** Two-cycle write operation for VCMA-assisted SOT switching.

the SOT channel. We provide detailed thickness optimization of the SOT layer while considering the effect of IR drop, write energy, and MTJ selectivity. For both the VCMA and STT-assisted write operations, we evaluate the impact of increasing the number of MTJs on an SOT channel in terms of WER and write energy. Moreover, in the case of SOT + STT scheme, we study the impact of pulse timings of the SOT and STT write currents. In addition, we evaluate the read performance of the cell as a function of oxide thickness and present the associated trade-offs in terms of read and write operations.

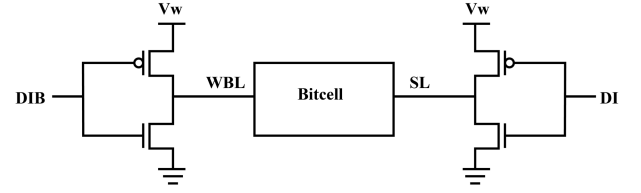
The rest of this article is organized as follows. After this introduction, Sections II and III describe the SOT + VCMA and SOT + STT schemes, respectively. In Section IV, we evaluate the read performance. Section V presents the optimization and benchmarking results for cell area and write performance, and the key findings of this article are summarized in Section VI.

## II. SOT + VCMA

The first write mechanism we discuss is to use VCMA effect to selectively write into MTJs on a shared SOT channel.

### A. CELL DESIGN AND WRITE OPERATION

The 3-D layout and schematic of the cell are shown in Fig. 1. The SOT channel is shared among multiple MTJs while having a single SOT write transistor. Selecting a specific MTJ for writing data is achieved by applying a voltage on the desired MTJ through the corresponding read/write select transistor. The write operation is based on utilizing the VCMA effect [9] to lower the thermal stability ( $\Delta$ ), thereby



**FIGURE 2. Write drivers for driving the SOT channel. DI and DIB represent the write data and its complementary value, respectively, and  $V_w$  is the write voltage.**

lowering the switching current by applying a voltage across an MTJ. The applied spin current is then selected such that it is large enough to switch MTJs with reduced thermal stability and small enough to avoid flipping nonselected MTJs. Writing to all MTJs on an SOT channel can be accomplished in two cycles, as shown in Fig. 1(b). In Cycle 1, all the 1's can be written, while all the 0's can be written in the next cycle by reversing the direction of the SOT current. The write operation requires the presence of an external magnetic field, which can be generated on-chip by using a cobalt magnetic hard mask [5]. For driving the SOT channel, the driver design described by earlier work [10] can be used. Fig. 2 shows the schematic of the write driver, which uses eight fin transistors to provide sufficient SOT current. The pitch and height of the write driver are  $8F$  and  $28F$  [11], respectively, with half metal pitch ( $F$ ) being 32 nm. The write drivers may occupy  $\approx 7\%$  of the total area for an array size of  $256 \times 128$ . Fig. 3 shows the memory array based on shared SOT channel.

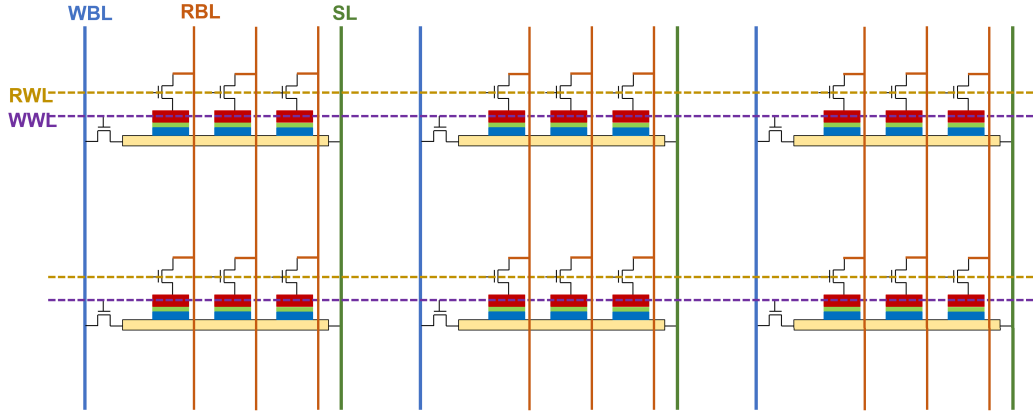
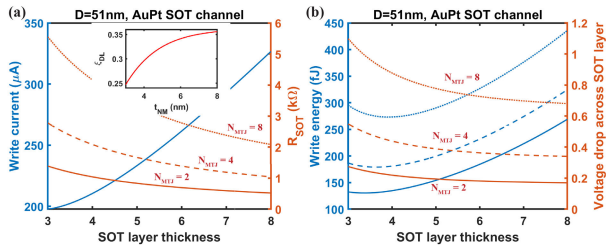
### B. IR DROP IN THE SOT LAYER

The length of the SOT layer depends on the number of MTJs ( $N_{\text{MTJ}}$ ) integrated on it. A longer SOT channel results in a higher resistance ( $R_{\text{SOT}}$ ); hence, a larger voltage drop  $V_{\text{SOT}}$  across it. A large IR drop across the SOT channel can result in larger write voltages that can pose several challenges, such as large variation in the effective VCMA voltages and the requirement for high-voltage transistors. To lower the resistance, the thickness ( $t_{\text{SOT}}$ ) of the SOT channel can be increased. However, a larger  $t_{\text{SOT}}$  may require a larger write current ( $I_w$ ) to maintain a sufficient current density ( $J_{\text{SOT}}$ ) in the SOT channel. In addition, damping-like spin-torque efficiency ( $\xi_{\text{DL}}$ ) may also change with  $t_{\text{SOT}}$ , according to the drift-diffusion model of spin generation and transport [12]

$$\xi_{\text{DL}} = \theta_{\text{SH}} \frac{G_r \tanh(t_{\text{SOT}}/2\lambda_{\text{sd}})}{\sigma_{\text{SOT}}/2\lambda_{\text{sd}} + G_r \coth(t_{\text{SOT}}/\lambda_{\text{sd}})} \quad (1)$$

where  $\theta_{\text{SH}}$  is the spin Hall angle,  $\lambda_{\text{sd}}$  is the spin diffusion length in the SOT material,  $G_r$  is the real part of the spin-mixing conductance ( $G_{\uparrow\downarrow}$ ), and  $\sigma_{\text{SOT}}$  is the conductivity of the SOT material.

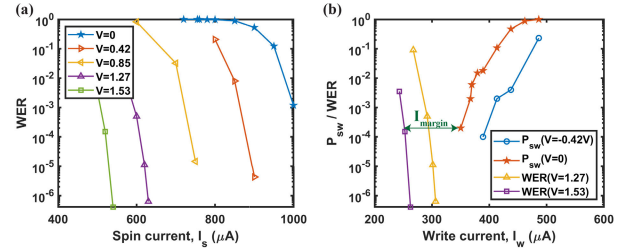
For the SOT channel, we use AuPt [13], which is a well studied SOT material with low resistivity ( $83 \mu\Omega\text{cm}$ ) and large  $\xi_{\text{DL}}$ . Fig. 4(a) shows the required  $I_w$  as well as  $R_{\text{SOT}}$  versus  $t_{\text{SOT}}$ . The inset plot in Fig. 4(a) shows the variation of  $\xi_{\text{DL}}$  with  $t_{\text{SOT}}$ . Increasing  $t_{\text{SOT}}$  results in an increased  $I_w$  despite the increase in  $\xi_{\text{DL}}$  as  $J_{\text{SOT}}$  decreases. The resistance;


**FIGURE 3.** Schematic for memory array design.

**FIGURE 4.** (a) SOT write current and SOT channel resistance versus the SOT channel thickness for AuPt. Inset: damping-like spin torque efficiency for AuPt versus thickness. (b) SOT write energy and IR drop across the SOT channel versus thickness.

however, decreases with increasing  $t_{\text{SOT}}$ , resulting in an overall reduction in  $V_{\text{SOT}}$ , as seen in Fig. 4(b). Write energy ( $E_w$ ), on the other hand, is nonmonotonous, and the lowest  $E_w$  is obtained when  $t_{\text{SOT}}$  is 3.5–4 nm.

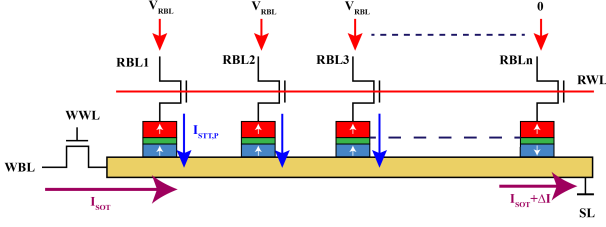
### C. DEVICE SIMULATIONS

To obtain various trade-offs among write current, write time, and WER and to evaluate the VCMA selectivity of the MTJs, we use object oriented micromagnetic framework (OOMMF [14]) simulations augmented with the rare-event enhancement [15] method. The simulation framework has already been validated with experiments [16]. We use perpendicular MTJ with a diameter of 51 nm and a free-layer thickness of 1.2 nm. The room temperature saturation magnetization ( $M_s$ ) and interface anisotropy ( $K_i$ ) are 1.257 MA/m and 1.3 mJm<sup>-2</sup>, respectively [17], which provides a room temperature  $\Delta$  of  $\approx 90$ . Required symmetry breaking for SOT switching can be achieved by applying a magnetic field of 32 mT [5]. In addition, we assume a field-like to damping-like torque ratio of 0.18 [18]. Fig. 5(a) shows the obtained WER versus applied spin current for various values of voltage applied across the MTJ ( $V_{\text{MTJ}}$ ). The duration of the write current is 1 ns. We use a VCMA coefficient of 100 fJ/V-m [19]. To quantify the VCMA selectivity, we also calculate the accidental write rate for the nonselected MTJ, as shown in Fig. 5(b). Here, accidental write rate refers to the probability of a nonselected MTJ ( $V_{\text{MTJ}} = 0$ ) getting switched.

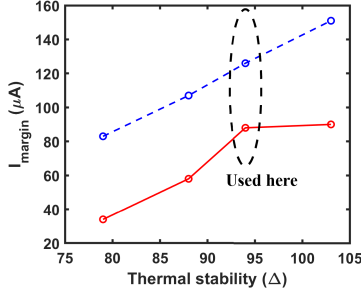

**FIGURE 5.** (a) WER versus applied spin current in the presence of VCMA effect for different MTJ voltages. (b) WER and accidental write rate versus write current for a 6-nm AuPt SOT channel. The write current is applied for 1 ns.

Here, we have ignored the effect of any STT current due to  $V_{\text{MTJ}}$ , as the design requires minimization of STT current as discussed next. In addition, field-assisted switching of perpendicular magnets usually requires larger damping coefficient [20] ( $\approx 0.1$ ), which effectively suppresses the effects of the STT current.

One key challenge with regards to the VCMA selectivity of MTJs is the current injected in the SOT channel due to the applied  $V_{\text{MTJ}}$ . Application of  $V_{\text{MTJ}}$  results in a finite amount of current being added into the SOT channel. This increases the overall SOT current in the channel. This extra current ( $\Delta I$ ) can help reduce the WER for the selected MTJs; however, it will also have the unintended effect of accidentally switching the nonselected MTJs. This extra current can be quantified as a function of  $N_{\text{MTJ}}$  and oxide thickness ( $t_{\text{ox}}$ ). The worst case, corresponding to maximum  $\Delta I$ , occurs when  $(N_{\text{MTJ}} - 1)$  consecutive MTJs are written parallel (P) to antiparallel (AP) in one cycle, while the remaining MTJ is written AP–P in another cycle, as shown in Fig. 6. In this case, we can define,  $\Delta I = (N_{\text{MTJ}} - 1)(V_{\text{MTJ}}/R_p)$ , where  $R_p$  is the resistance of the MTJs in P state. The maximum allowable value of  $\Delta I$  is determined by the available switching margin ( $I_{\text{margin}}$ ), which is defined as the difference in the write currents for the selected and nonselected MTJs corresponding to a target error rate, as shown in Fig. 5(b). It is also important to note that during the write operation, nonselected MTJs will experience negative voltage ( $V_{\text{MTJ}} < 0$ ) due to the finite potential of the SOT channel. Thus, the available  $I_{\text{margin}}$  will be higher than the depicted value in Fig. 5(b). For reliable operation,



**FIGURE 6.** Schematic representation for the worst case write scenario for the SOT + VCMA write. ( $N_{\text{MTJ}} - 1$ ) consecutive MTJs are written P-AP, while the last MTJ is written AP-P.



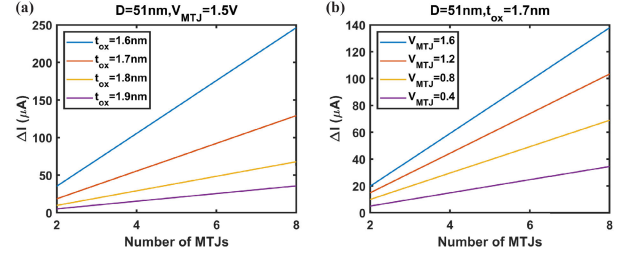
**FIGURE 7.** Available  $I_{\text{margin}}$  plotted as a function of  $\Delta$  when the voltage across the nonselected MTJ is 0 V (solid lines) and  $-0.42$  V (dashed lines), for the SOT + VCMA switching. WER is  $10^{-6}$ . The voltage across the selected MTJ is 1.53 V.

$\Delta I < I_{\text{margin}}$  is required. The obtained values of  $I_{\text{margin}}$  for a WER of  $10^{-6}$  are  $\approx 88$  and  $\approx 127$   $\mu\text{A}$ , when the voltage across the nonselected MTJ is 0 and  $-0.42$  V, respectively. In comparison, the corresponding  $I_{\text{margin}}$  values for a WER of  $10^{-4}$  are 97 and 137  $\mu\text{A}$ , respectively. Here, for the nonselected MTJ, it is not possible for us to calculate accidental write rates below  $10^{-4}$  due to the limitations imposed by the computation time. Improving the VCMA coefficient and lowering the charge to spin conversion efficiency can increase  $I_{\text{margin}}$ . In addition,  $I_{\text{margin}}$  depends on  $\Delta$ , as shown in Fig. 7. Further optimization of the magnetic parameters is required to improve  $I_{\text{margin}}$ . Reducing  $V_{\text{MTJ}}$  can lower  $\Delta I$ ; however, that would also reduce  $I_{\text{margin}}$ . The best trade-off among  $I_w$ ,  $V_{\text{SOT}}$ , and  $I_{\text{margin}}$  can be achieved by selecting  $V_{\text{MTJ}} = 1.5$  V and  $t_{\text{SOT}} = 6$  ns.

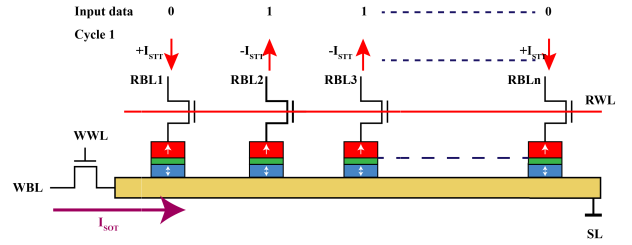
Another way to suppress  $\Delta I$  is to increase  $t_{\text{ox}}$ , which increases the MTJ resistance and lowers the current passing through it. It is shown in Fig. 8(a) where the values of  $\Delta I$  corresponding to different  $t_{\text{ox}}$  values are plotted against  $N_{\text{MTJ}}$ . Here, the MTJ resistance values are obtained from experiment [21]. Fig. 8(b) shows  $\Delta I$  versus  $N_{\text{MTJ}}$  for various values of  $V_{\text{MTJ}}$  at  $t_{\text{ox}} = 1.7$  nm. Increasing  $t_{\text{ox}}$  beyond 1.6 nm can significantly suppress  $\Delta I$ , allowing the integration of a larger number of MTJs on a single SOT channel. However, a large  $t_{\text{ox}}$  comes with a read performance penalty as discussed in Section IV.

### III. SOT + STT

Another way of sharing the SOT channel among MTJs is to use a small STT instead of the VCMA effect.



**FIGURE 8.** (a) Extra current ( $\Delta I$ ) in the SOT channel due to  $V_{\text{MTJ}}$  plotted as a function of number of MTJs at various  $t_{\text{ox}}$  values for  $V_{\text{MTJ}} = 1.5$  V. (b)  $\Delta I$  versus  $N_{\text{MTJ}}$  for  $t_{\text{ox}} = 1.7$  nm for various values of  $V_{\text{MTJ}}$ .



**FIGURE 9.** Cell design with write scheme using STT-assisted SOT.

#### A. CELL DESIGN AND WRITE OPERATION

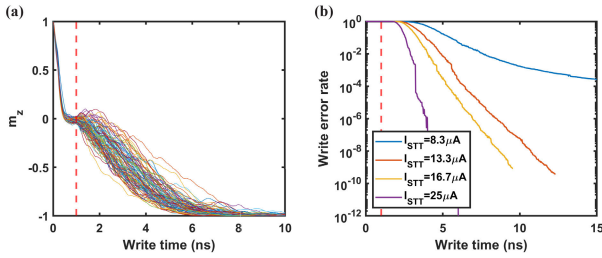
The cell design is the same as the SOT + VCMA scheme. In this case, the deterministic magnetization switching is achieved by applying a small STT current. First, an SOT current is applied to move the magnetizations of all the MTJs toward the in-plane meta-stable direction. After that, the SOT current is stopped, and a small STT current is applied through each MTJ. The direction of the STT current determines the final MTJ state. All the MTJs are written at once by applying appropriate polarities of STT currents. The write scheme is demonstrated in Fig. 9. Also, as the direction of the SOT write current remains the same, a separate driver for SL is not required.

#### B. DEVICE SIMULATIONS

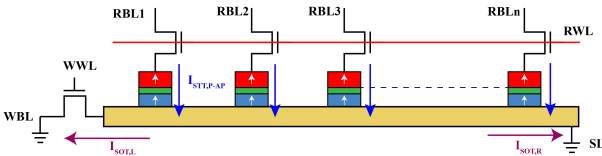
The diameter and thickness of the free-layer ferromagnet used here are 42 and 1.3 nm, respectively, giving a room temperature  $\Delta$  of  $\approx 60$ . Contrary to the SOT + VCMA case ( $\Delta \approx 90$ ),  $\Delta$  used here is lower, as the SOT + VCMA scheme requires a large  $\Delta$  to effectively suppress the accidental write rate for nonselected MTJs. The SOT + STT scheme has no such restriction, and the value of  $\Delta$  can be chosen based on the retention time requirement. Fig. 10(a) shows the magnetization switching for a single MTJ, illustrating the write scheme used. The spin current generated by the SOT is fixed at 600  $\mu\text{A}$ , which is applied for 1 ns. The magnitude and direction of the STT current are varied to obtain various WERs, as seen in Fig. 10(b). Here, we assume the STT efficiencies of 0.6 for AP-P and 0.3 for P-AP switching [22].

Similar to the VCMA-assisted write, the number of MTJs on a single SOT channel is limited by the SOT current in

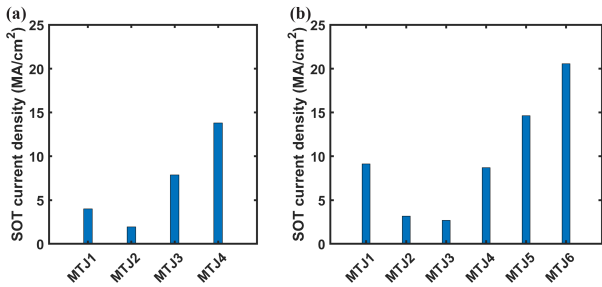




**FIGURE 10.** (a) Micromagnetic simulation results for magnetization dynamics and (b) WER versus write time for various STT currents for a single MTJ (AP-P switching) based on the SOT + STT write scheme. The spin current due to STT in (a) is  $8 \mu\text{A}$  ( $I_{\text{STT}} = 13.3 \mu\text{A}$ ).



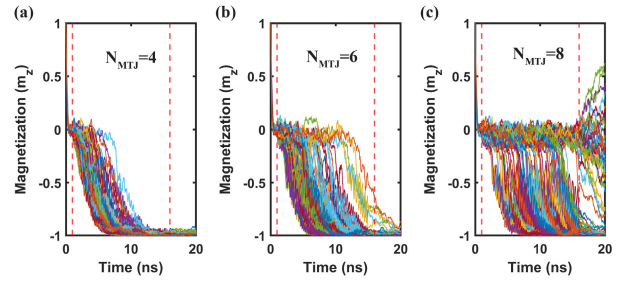
**FIGURE 11.** Worst case write operation for the SOT + STT scheme.



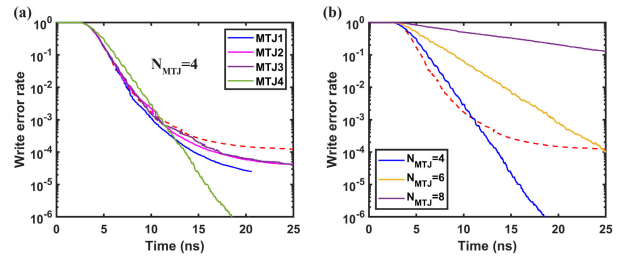
**FIGURE 12.** Current densities in the SOT layer below each MTJ calculated using finite-element simulations in the worst case when the number of MTJs on an SOT channel is (a) 4 and (b) 6.

the worst case scenario for the write operation. During the STT switching phase, there will be a finite amount of current injected into the SOT channel due to STT current. The current flowing in the SOT channel will apply an in-plane torque on the magnetization of the free layer. If this current becomes too large, it will result in the magnetization being stuck in-plane, suppressing the effect of STT. This will cause switching errors and increased WER. The worst case scenario is when all the MTJs are being switched from P to AP state, as shown in Fig. 11. To reduce the SOT current seen by MTJs, we ground both write bitline (WBL) and SL during the STT phase. This allows the current to flow in both directions within the SOT channel and lowers the voltage drop. To calculate the resulting current density in the SOT channel below each MTJ, we use COMSOL-based finite-element simulations. Fig. 12 shows the obtained SOT current densities below each MTJ for  $N_{\text{MTJ}} = 4$  and  $N_{\text{MTJ}} = 6$  cases due to the applied STT current of  $16.7 \mu\text{A}$ . The resulting current density data are used in micromagnetic simulations to calculate WER and find the limit on the number of MTJs.

Fig. 13 shows the magnetization dynamics corresponding to the worst case write operation for the MTJ seeing



**FIGURE 13.** Magnetization dynamics for the worst case write operation for the case of (a)  $N_{\text{MTJ}} = 4$ , (b)  $N_{\text{MTJ}} = 6$ , and (c)  $N_{\text{MTJ}} = 8$  for the SOT + STT scheme.

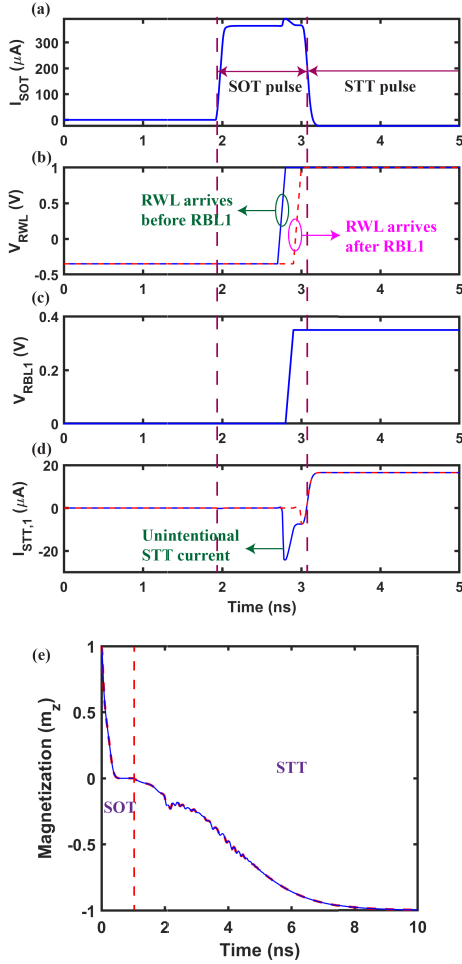


**FIGURE 14.** Simulation results for WER for the worst case write operation for the SOT + STT scheme. (a) WER compared for different MTJs sharing an SOT channel for  $N_{\text{MTJ}} = 4$ . (b) WER compared for MTJ with the largest SOT current for  $N_{\text{MTJ}} = 4, 6, \text{ and } 8$ . Dashed lines in both the panels represent the case for a single MTJ without any SOT current during the STT phase.

the most SOT current when  $N_{\text{MTJ}}$  is 4, 6, and 8. Large amount of current flowing in the SOT channel results in increased switching failures, as seen in Fig. 13(c). WER in the worst case for different MTJs on an SOT channel for  $N_{\text{MTJ}} = 4$  is shown in Fig. 14(a). Fig. 14(b) depicts WER for the MTJs experiencing the largest SOT current in the worst case write operation for  $N_{\text{MTJ}} = 4, 6, \text{ and } 8$ . The results show that increasing the number of MTJs leads to higher WER.

### C. ROBUSTNESS TO WRITE PULSE TIMING

Another key metric for the circuit is its sensitivity to the timing of SOT and STT pulses. Based on SPICE simulation results, we show that the circuit is robust with regards to any variation in the relative timings of SOT and STT pulses, as shown in Fig. 15. We apply the STT pulse 100 ps before the SOT pulse ends; assuming the uncertainty due to jitter and skew does not exceed 100 ps. This ensures that as soon as SOT ends, STT will begin to switch the magnetization in the desired direction. A delay between SOT and STT may lead to switching errors, as the magnetization remains in the meta-stable state, and thermal noise may move it in the unwanted direction. During the SOT phase, SOT channel remains at finite potential, while read bitlines (RBLs) are grounded. If read wordline (RWL) is enabled before RBLs are charged [solid lines in Fig. 15(b) and (d)], there can be STT current flowing through the MTJ from the free layer

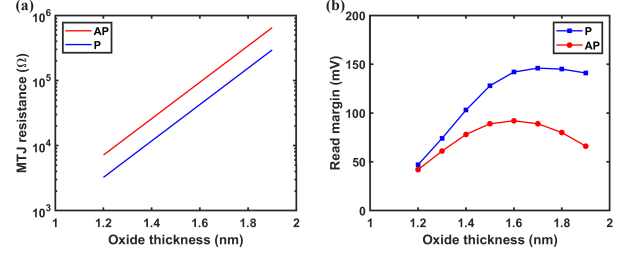


**FIGURE 15.** (a)–(d) SPICE simulation waveform for the SOT + STT write scheme showing the SOT and STT current pulses. For the STT current, two cases are considered:  $V_{RWL}$  arrives before  $V_{RBL1}$  (solid lines), and  $V_{RWL}$  arrives after  $V_{RBL1}$  (dashed lines). The results correspond to  $N_{MTJ} = 4$  with all MTJs being written P–AP. In both the cases, there is STT current in opposite direction than the intended for a short time; however, the STT current is much smaller ( $<10\%$ ) than the SOT current for that specific duration. (e) Magnetization dynamics for the two cases.

toward the fixed layer for a small amount of time. However, this will not be an issue, as this unintended STT current is much smaller ( $<10\%$ ) in magnitude than the SOT current applied on the MTJs and will not affect the magnetization dynamics, as shown in Fig. 15(e).

#### IV. READ OPERATION

The read performance is evaluated based on SPICE simulations. We use a differential sensing scheme [23], [24] for the read operation. Only one MTJ on a single SOT channel can be read at a time, as the read current path is shared among them. This is not an issue, as the number of MTJs that can be read at once is limited by the number of sense amplifiers (SAs). We assume one SA for every 64 bitlines as commonly done in STT-MRAM arrays. These 64 bitlines are



**FIGURE 16.** (a) Resistances of the MTJ in the P and AP states for an MTJ diameter of 51 nm. (b) Read margins in P and AP states plotted as a function of oxide thickness.

**TABLE 1.** Parasitic and interconnect resistance and capacitance values used in the simulations.

Quantity	Value
Gate capacitance	60 aF/fin
Wire capacitance	0.15 fF/ $\mu\text{m}$
WBL resistance	47.7 $\Omega/\mu\text{m}$
WWL resistance	20 $\Omega/\mu\text{m}$
RBL resistance	47.7 $\Omega/\mu\text{m}$
RWL resistance	47.7 $\Omega/\mu\text{m}$
SL resistance	20 $\Omega/\mu\text{m}$

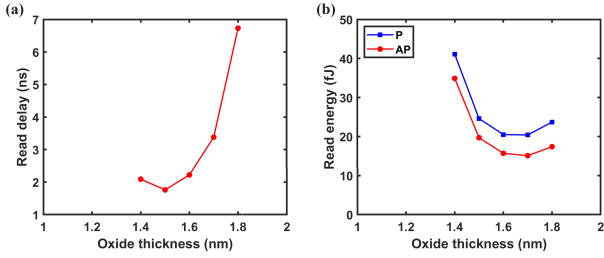
multiplexed together and then compared with the reference MTJ. The bitline voltages corresponding to the MTJ being read and the reference MTJ are compared using a double-tail latch-type voltage SA [25].

The read performance strongly depends on  $t_{ox}$  and tunnel magnetoresistance (TMR) ratio. To evaluate the read performance, we consider  $t_{ox}$  from 1.2 to 1.9 nm. The resistance area (RA) product values are obtained from experimental data [21]. We assume a constant TMR ratio of 120% [17]. Fig. 16(a) shows the resistance of the MTJ with a diameter of 51 nm in P and AP states. Read performance is evaluated using SPICE simulations for a  $256 \times 128$  array with four MTJs on each SOT channel. We use 14-nm FinFET models from the Predictive Technology Model (PTM) by Arizona State University (ASU) [26] with a half metal pitch of 32 nm. Table 1 lists the parasitic resistance and capacitance values used in the simulations. The capacitance values are obtained from prior benchmarking work [24], and the resistance values of wires are calculated based on Cu resistivity values reported in [27]. Fig. 16(b) shows the obtained read margins in P and AP states for the nominal case where the read margin is defined as the voltage difference seen at the input of the SA. The read margin reduces drastically for  $t_{ox} < 1.4$  nm and  $t_{ox} > 1.7$  nm, especially for the AP state.

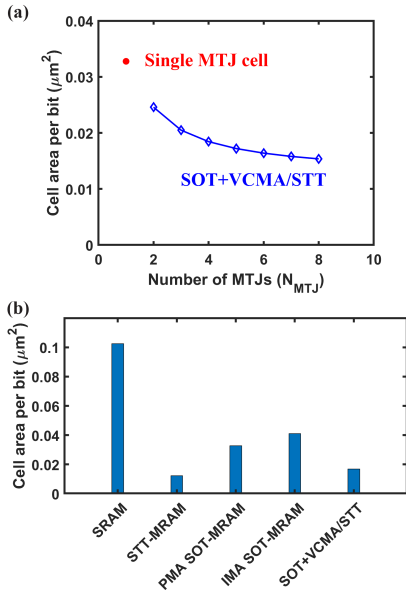
To account for variation, we use  $3\sigma$  variation of 10% in MTJ area and 10% uniform variation in the supply voltage while also accounting for thermal noise. We use a read time of  $5\sigma$  higher than the mean value to obtain read error rate below  $10^{-6}$ . The total read delay can be written as follows [24]:

$$t_{read} = 0.7R_{drive}C_{RWL} + 0.4R_{RWL}C_{RWL} + t_{sense} \quad (2)$$

where  $R_{drive}$  ( $=5$  k $\Omega$ ) and  $R_{RWL}$  are the resistances of



**FIGURE 17.** (a) Effective read delay versus oxide thickness in the AP state. (b) Read energy versus oxide thickness for reading P and AP states. Read margin used is 70 mV.

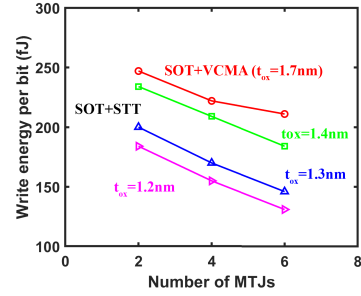


**FIGURE 18.** (a) Cell area per bit versus the number of MTJs on a shared SOT channel for SOT + VCMA/STT scheme compared against the area of an SOT-MRAM with single MTJ. (b) Cell area per bit for various magnetic memory options.

the drive transistor and RWL, respectively,  $C_{\text{RWL}}$  is the capacitance of RWL, and  $t_{\text{sense}}$  accounts for the delay to reach the required voltage margin. The effective read delay and energy, including the effects of variation, are shown in Fig. 17 for a read margin of 70 mV. For  $t_{\text{ox}} = 1.3$  nm and  $t_{\text{ox}} = 1.9$  nm, the available read margin is  $<60$  mV. Optimal read performance is observed for  $t_{\text{ox}}$  within the range of 1.4–1.6 nm. Increasing  $t_{\text{ox}}$  initially results in lower read energies because of smaller read currents; however, beyond 1.7 nm, the read energy starts to increase, as the delay goes up rapidly due to read current being too small. The choice of  $t_{\text{ox}}$  based on the reliability of write operation is different from the read performance optimization. The SOT + VCMA scheme requires a larger  $t_{\text{ox}}$  to suppress any extra current due to  $V_{\text{MTJ}}$ , while the SOT + STT scheme requires lower oxide thickness to reduce the write energy.

## V. BENCHMARKING

We benchmark this SOT + VCMA/STT scheme against other competing memories, such as SRAM, STT-MRAM,

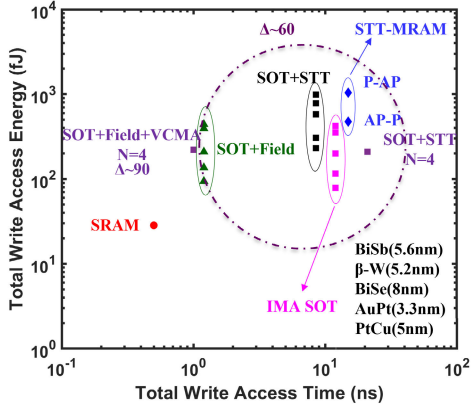


**FIGURE 19.** Write energy as a function of number of MTJs on an SOT channel for the SOT + VCMA and SOT + STT schemes. For the SOT + STT scheme, three different  $t_{\text{ox}}$  values are considered: 1.2, 1.3, and 1.4 nm. For the SOT + VCMA scheme,  $t_{\text{ox}} = 1.7$  nm.

**TABLE 2.** Write current and voltages used for the SOT + VCMA scheme.

Quantity	Value
SOT current	275 $\mu\text{A}$
$V_{\text{WWL}}$	1.2V
$V_{\text{WBL}}$	0.84V
$V_{\text{SL}}$	0.79V
$V_{\text{RWL}}$	2.4V
$V_{\text{RBL}}$	2.1V

and in-plane magnetic anisotropy (IMA) and perpendicular magnetic anisotropy (PMA)-based conventional two transistor SOT-MRAM. Fig. 18(a) shows the cell area per bit versus the number of MTJs for the SOT + VCMA/STT scheme. In Fig. 18(b), the cell area per bit of the SOT + VCMA/STT scheme with four MTJs on a shared SOT channel is compared against those of other magnetic memory options. In both plots, the 14-nm technology node ( $F = 32$  nm) and the layout rules described in prior benchmarking work [11], [28] are used to calculate the cell areas. Compared with the conventional 2T SOT-MRAM,  $\approx 2\times$  bit density can be achieved. The write energies for the SOT + VCMA and SOT + STT schemes, calculated using SPICE simulations, are shown in Fig. 19. The write voltages and current for the SOT + VCMA scheme are listed in Table 2, and the same for the SOT + STT are listed in Table 3. The write energy values are benchmarked against other memory options [16], as shown in Fig. 20. For the conventional 2T SOT-MRAM cell, the write energy results for various SOT materials are included, such as PtCu [29], AuPt [13], BiSe [30],  $\beta$ -W [31], and BiSb [32]. The SOT + VCMA scheme has a higher write energy but much lower write delay compared with the SOT + STT scheme. The higher write energy can be attributed to the large  $\Delta$  ( $\approx 90$ ) requirement as discussed in Section III-B and the large energy associated with charging RBL capacitance due to the application of  $V_{\text{MTJ}}$ . The higher thermal stability can be useful, as it will increase the data retention time. The higher write delay observed in the SOT + STT scheme with SOT channel sharing compared with the conventional 2T SOT + STT MRAM cell is due to lower STT current requirement to



**FIGURE 20.** Array-level write energy and write time for the SOT + VCMA/STT schemes compared against other memories. The array size is  $256 \times 128$  bits. The order of data points for the 2T SOT + Field, SOT + STT, and SOT + IMA cases is the same as the order of SOT materials mentioned in the plot. The numbers in brackets mention the thicknesses for the corresponding SOT materials.

**TABLE 3.** Write currents and voltages used for the SOT + STT scheme.

Quantity	Value
SOT current	$365 \mu\text{A}$
STT current	$16.6 \mu\text{A}$ (P-AP)/ $-8.3 \mu\text{A}$ (AP-P)
$V_{WWL}$	1.3V
$V_{WBL}$	1.1V
$V_{SL}$	0V
$V_{RWL}$	1V
$V_{RBL}$ ( $t_{ox}=1.2\text{nm}$ )	$\pm 0.13\text{V}$
$V_{RBL}$ ( $t_{ox}=1.3\text{nm}$ )	$\pm 0.20\text{V}$
$V_{RBL}$ ( $t_{ox}=1.4\text{nm}$ )	$\pm 0.35\text{V}$

suppress WER in the worst case write as discussed previously. Overall, both the SOT + VCMA and SOT + STT schemes discussed here provide major density advantage over the conventional SOT-MRAM while sacrificing a bit in the write performance.

One important question here is that improving which material properties would more significantly improve the array-level performance of the proposed schemes. Some key material properties, which are considered here for benchmarking, are STT efficiency, SOT efficiency, and VCMA coefficient. There are not any known approaches to improve the STT efficiency, and the current values that are commonly used (60%) are not too far from the ideal value, which is 100%. On the other hand, improving the SOT efficiency is an active area of research with many promising materials being explored. For the SOT + VCMA scheme, increasing the SOT efficiency while keeping the SOT layer thickness will reduce the available  $I_{\text{margin}}$ , resulting in higher error rates. Similarly, for the SOT + STT scheme, a higher SOT efficiency may result in an increased SOT during the STT phase and increased error rate. However, a higher available SOT efficiency may allow increasing the SOT layer thickness, which can help the IR drop issue, improving the device

performance and reliability. Also, for the SOT + VCMA scheme, improving the VCMA coefficient will have the most impact, as it will lower the write energy and increase  $I_{\text{margin}}$ , thereby lowering the WER.

## VI. CONCLUSION

This article presents a comprehensive modeling, optimization, and benchmarking of transistor sharing schemes for SOT-MRAM devices using VCMA and STT effects. Using experimentally validated micromagnetic simulations augmented with rare-event enhancement along with SPICE simulations, we demonstrate that the number of MTJs that can be put on a single SOT channel is limited by the write error induced due to the injection of current in the SOT channel through the MTJs and voltage drop on the SOT channel. For the SOT + VCMA scheme, we quantify the WER, unintentional write rate, and the current injection through MTJs as a function of the MTJ oxide thickness. For the SOT + STT scheme, finite-element simulations are used to calculate the SOT current density in the SOT channel underneath each MTJ and the resulting WER. In addition, we quantify the IR drop along the SOT layer in terms of the number of MTJs and provide a way to optimize the SOT layer thickness while considering the write energy, current, SOT channel resistance, and the voltage drop along the SOT layer. Our results indicate that having four to six MTJs on a single SOT channel provides the best trade-off among the write energy, bit density, WER, and IR drop. The SOT + VCMA/STT schemes show a  $\approx 2\times$  bit density improvement over the conventional two transistor SOT-MRAM and a  $\approx 6\times$  bit density improvement over SRAM. While the energies are slightly higher than the conventional 2T SOT-MRAM, the SOT + VCMA/STT schemes are still more energy efficient than STT-MRAM. We also quantify the read performance in terms of oxide thickness and show the read penalty associated with sharing SOT channel among MTJs. Our read simulation results show read times  $< 4$  ns for both the schemes. Moreover, since the current through the select transistors is significantly smaller than that of STT-MRAM, this approach may enable adopting SOT-MRAM to more advanced technology nodes.

While both the SOT + VCMA and SOT + STT schemes look promising, there are certain challenges that must be addressed. For the VCMA + SOT scheme, a relatively large VCMA coefficient ( $> 100$  fJ/V-m) is needed to keep the required VCMA voltage below 1.5 V. A tighter control over variation in magnetic properties is also required to ensure sufficient  $I_{\text{margin}}$ . For the SOT + STT scheme, there is an additional cost associated with the peripheral circuits that can supply both the positive and negative voltages for the write operation.

## ACKNOWLEDGMENT

The authors would like to thank Prof. D. Ralph, J.-P. Wang, and S. X. Wang, and Dr. J. Sun, C. H. Diaz, S. Dutta, W. Hwang, S.-J. Lin, N. Xu, X. Li, W. Tsai, and M. DC for insightful discussions.



## REFERENCES

- [1] Y. Seo, K.-W. Kwon, X. Fong, and K. Roy, "High performance and energy-efficient on-chip cache using dual port (1R/1W) spin-orbit torque MRAM," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 6, no. 3, pp. 293–304, Sep. 2016.
- [2] J. G. Alzate et al., "2 MB array-level demonstration of STT-MRAM process and performance towards 14 cache applications," in *IEDM Tech. Dig.*, Dec. 2019, pp. 2.4.1–2.4.4.
- [3] Y. J. Song et al., "Demonstration of highly manufacturable STT-MRAM embedded in 28 nm logic," in *IEDM Tech. Dig.*, Dec. 2018, pp. 18.2.1–18.2.4.
- [4] W. A. Borders et al., "Analogue spin-orbit torque device for artificial-neural-network-based associative memory operation," *Appl. Phys. Exp.*, vol. 10, no. 1, 2016, Art. no. 013007.
- [5] K. Garelo et al., "Manufacturable 300 mm platform solution for field-free switching SOT-MRAM," in *Proc. Symp. VLSI Technol.*, Jun. 2019, pp. T194–T195.
- [6] W. Hwang et al., "Energy efficient computing with high-density, field-free STT-assisted SOT-MRAM (SAS-MRAM)," *IEEE Trans. Magn.*, early access, Dec. 1, 2022, doi: [10.1109/TMAG.2022.3224729](https://doi.org/10.1109/TMAG.2022.3224729).
- [7] H. Yoda et al., "Voltage-control spintronics memory (VoCSM) having potentials of ultra-low energy-consumption and high-density," in *IEDM Tech. Dig.*, Dec. 2016, pp. 27.6.1–27.6.4.
- [8] K. Cai et al., "Selective operations of multi-pillar SOT-MRAM for high density and low power embedded memories," in *Proc. IEEE Symp. VLSI Technol. Circuits (VLSI Technol. Circuits)*, Jun. 2022, pp. 375–376.
- [9] M. Endo, S. Kanai, S. Ikeda, F. Matsukura, and H. Ohno, "Electric-field effects on thickness dependent magnetic anisotropy of sputtered MgO/Co<sub>40</sub>Fe<sub>40</sub>B<sub>20</sub>/Ta structures," *Appl. Phys. Lett.*, vol. 96, no. 21, May 2010, Art. no. 212503.
- [10] H. Noguchi et al., "7.2 4 Mb STT-MRAM-based cache with memory-access-aware power optimization and write-verify-write/read-modify-write scheme," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Jan. 2016, pp. 132–133.
- [11] D. E. Nikonov and I. A. Young, "Overview of beyond-CMOS devices and a uniform methodology for their benchmarking," *Proc. IEEE*, vol. 101, no. 12, pp. 2498–2533, Dec. 2013.
- [12] Y.-T. Chen et al., "Theory of spin Hall magnetoresistance," *Phys. Rev. B, Condens. Matter*, vol. 87, no. 14, 2013, Art. no. 144411.
- [13] L. Zhu, D. C. Ralph, and R. A. Buhrman, "Highly efficient spin-current generation by the spin Hall effect in Au<sub>1-x</sub>Pt<sub>x</sub>," *Phys. Rev. Appl.*, vol. 10, no. 3, 2018, Art. no. 031001.
- [14] M. Donahue and D. Porter. *OOMMF User's Guide, Version 1.0*. Accessed: Jun. 1, 2022. [Online]. Available: <http://math.nist.gov/oommf>
- [15] U. Roy, T. Pramanik, L. F. Register, and S. K. Banerjee, "Write error rate of spin-transfer-torque random access memory including micromagnetic effects using rare event enhancement," *IEEE Trans. Magn.*, vol. 52, no. 10, pp. 1–6, Oct. 2016.
- [16] P. Kumar and A. Naeemi, "Benchmarking of spin-orbit torque vs spin-transfer torque devices," *Appl. Phys. Lett.*, vol. 121, no. 11, Sep. 2022, Art. no. 112406.
- [17] S. Ikeda et al., "A perpendicular-anisotropy CoFeB–MgO magnetic tunnel junction," *Nature Mater.*, vol. 9, no. 9, pp. 721–724, Sep. 2010.
- [18] S. Shi, Y. Ou, S. V. Aradhya, D. C. Ralph, and R. A. Buhrman, "Fast low-current spin-orbit-torque switching of magnetic tunnel junctions through atomic modifications of the free-layer interfaces," *Phys. Rev. Appl.*, vol. 9, no. 1, 2018, Art. no. 0111002.
- [19] W. Skowroński et al., "Perpendicular magnetic anisotropy of Ir/CoFeB/MgO trilayer system tuned by electric fields," *Appl. Phys. Exp.*, vol. 8, no. 5, May 2015, Art. no. 053003.
- [20] Z. Wang et al., "Progresses and challenges of spin orbit torque driven magnetization switching and application (invited)," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–5.
- [21] S. Yuasa, T. Nagahama, A. Fukushima, Y. Suzuki, and K. Ando, "Giant room-temperature magnetoresistance in single-crystal Fe/MgO/Fe magnetic tunnel junctions," *Nature Mater.*, vol. 3, no. 12, pp. 868–871, Oct. 2004.
- [22] A. Jaiswal, X. Fong, and K. Roy, "Comprehensive scaling analysis of current induced switching in magnetic memories based on in-plane and perpendicular anisotropies," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 6, no. 2, pp. 120–133, Jun. 2016.
- [23] M. Jeffremow et al., "Time-differential sense amplifier for sub-80 mV bitline voltage embedded STT-MRAM in 40 nm CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2013, pp. 216–217.
- [24] C. Pan and A. Naeemi, "Nonvolatile spintronic memory array performance benchmarking based on three-terminal memory cell," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 3, pp. 10–17, 2017.
- [25] D. Schinkel, E. Mensink, E. Klumperink, E. van Tuijl, and B. Nauta, "A double-tail latch-type voltage sense amplifier with 18 ps setup+hold time," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2007, pp. 314–605.
- [26] *Predictive Technology Model*. Accessed: Jun. 1, 2022. [Online]. Available: <http://ptm.asu.edu/>
- [27] I. Ciofi et al., "Impact of wire geometry on interconnect RC and circuit delay," *IEEE Trans. Electron Devices*, vol. 63, no. 6, pp. 2488–2496, Jun. 2016.
- [28] Y.-C. Liao, C. Pan, and A. Naeemi, "Benchmarking and optimization of spintronic memory arrays," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 6, pp. 9–17, 2020.
- [29] C. Hu and C. Pai, "Benchmarking of spin-orbit torque switching efficiency in Pt alloys," *Adv. Quantum Technol.*, vol. 3, no. 8, Aug. 2020, Art. no. 2000024.
- [30] D. C. Mahendra et al., "Room-temperature high spin-orbit torque due to quantum confinement in sputtered Bi<sub>x</sub>Se<sub>1-x</sub> films," *Nature Mater.*, vol. 17, no. 9, pp. 800–807, 2018.
- [31] C.-F. Pai, L. Liu, Y. Li, H. W. Tseng, D. C. Ralph, and R. A. Buhrman, "Spin transfer torque devices utilizing the giant spin Hall effect of tungsten," *Appl. Phys. Lett.*, vol. 101, no. 12, 2012, Art. no. 122404, doi: [10.1063/1.4753947](https://doi.org/10.1063/1.4753947).
- [32] Z. Chi, Y.-C. Lau, X. Xu, T. Ohkubo, K. Hono, and M. Hayashi, "The spin Hall effect of Bi-Sb alloys driven by thermally excited dirac-like electrons," *Sci. Adv.*, vol. 6, no. 10, Mar. 2020, Art. no. eaay2324.