Fokker–Planck Study of Parameter Dependence on Write Error Slope in Spin-Torque Switching

Yunkun Xie, Behtash Behin-Aein, and Avik W. Ghosh, *Senior Member, IEEE*

***Abstract*— This paper analyzes write errors in spin-torque switching due to thermal fluctuations in a system with perpendicular magnetic anisotropy. Prior analytical and numerical methods are summarized; a physics-based general 2-D Fokker–Planck equation (FPE) is solved numerically. Due to its computational efficiency and broad applicability to all switching regimes and system symmetries, the 2-D FPE has been used to study the relation between write error slope and material parameters as well as some emerging switching schemes.**

***Index Terms*— 2-D Fokker–Planck (FP), perpendicular magnetic anisotropy (PMA), spin-transfer torque (STT), write error rate (WER).**

I. Introduction

OVER the past decade, emerging spintronics and nanomagnetic devices have attracted a lot of attention due to their versatility, scalability, and energy efficiency. Part of the excitement stems from the discovery and experimental demonstration of spin-transfer-torque (STT) effect [1], [2]. Together with the tunneling magnetoresistance effect, they provide the means of write and read operations for memory applications. Owing to the nonvolatility of nanomagnets enabling persistent binary states, STT-based magnetic random access memories (STT-MRAMs) can be used both as working memory and for code storage purposes. Most major semiconductor companies are developing STT-MRAM for embedded or standalone applications. A fundamental issue accompanying magnetic switching is its susceptibility to thermal noise. At room temperature, the magnetic switching under STT reacts to thermal fluctuations and often results in a distribution of switching currents and delays. Some proposed applications even explicitly utilize these thermal fluctuations, such as random number generators in spin dice [3] or stochastic simulation of neuromorphic behavior [4]. In the case of write operation in STT-MRAM, increasing the write-voltage and/or pulsewidth can reduce write error rate (WER) but both quantities are limited by reliability, endurance, and overall performance metrics.

This paper first generalizes the Fokker–Planck (FP) formalism to a 2-D form and compares with several analytical models widely used for parameter extraction in experimental measurements for their relative strengths and weaknesses. It then uses FP to show how temperature, device, and material parameters affect WER slopes. It ends with some discussion on some proposed switching schemes with potential lower WER.

II. Switching Regions and Models

Discussions on thermal effects in spin-torque switching usually fall into two switching regions set by the ratio between injected current I and critical current I_c. The supercritical $I \gg I_c$ regime is called current dominated region while the subcritical regime $I \ll I_c$ is referred to as thermal region. This division allows approximate analytical models to be built in the two limits. However, practical magnetic tunnel junction (MTJ) applications mostly work in the intermediate regime at the moment because of three reasons—large MTJ resistance, relatively large critical switching current, and need to avoid time-dependent dielectric breakdown. Most analytical models do not have a "smooth" transition between supercritical/subcritical regimes and often encounter mathematical singularities in the transition. Recently there have been efforts at formulating a brute force mathematical transition between the analytical equations [5]. Such a scheme offers a simple fix to the discontinuity of the analytical equations but lacks physical insights into the switching behavior at transition. The alternative is to use numerical solution to avoid mathematical approximations and consider physical parameters at the same time. Numerical methods are quite universal and not limited to a specific region but are, in general, much less computationally efficient than analytical approaches. Among those numerical methods, we will show the advantages of FP approach in Section III. Note that all equations to be discussed are based on macrospin approximation. In reality, STT switching can involve complications, such as subvolume effects [6] or edge effects [7]. While it is crucial to understand those effects, accounting for both nonmacrospin effects and thermal effects can be computationally challenging in simulations. Thus, macrospin model is still valuable, because one can typically approximate those complicated effects with effective parameters that are good enough to draw physical insights and at the same time interface with practical device or circuit simulations.

Manuscript received June 30, 2016; revised September 23, 2016; October 26, 2016; November 15, 2016; accepted November 21, 2016. Date of current version December 24, 2016. This work was supported in part by NSF under Award NEB1124714 and in part by NSF under Grant CCF1514219. The review of this paper was arranged by Editor M. M. Cahay.

Y. Xie and A. W. Ghosh are with the Charles L. Brown Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22904 USA (e-mail: yx3ga@virginia.edu; ag7rq@virginia.edu).

B. Behin-Aein is with GLOBALFOUNDRIES, US, Inc., Santa Clara, CA 95054 USA (e-mail: behtash.behin-aein@globalfoundries.com).

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.2016.2632438

III. General Fokker–Planck Equation

The phenomenological Landau–Lifshitz–Gilbert (LLG) equation describes the dynamics of the normalized magnetization $\mathbf{m} = \mathbf{M}/M_s$ determined by the torque from effective magnetic field $\mathbf{H}_{\text{eff}}$, magnetic damping, and Slonczewski spin torque $\mathbf{L}_{\text{STT}}$ in the case of STT switching

$$\begin{aligned}\frac{\partial \mathbf{m}}{\partial t} &= \mathbf{L}_{\text{prec}} + \mathbf{L}_{\text{damp}} + \mathbf{L}_{\text{STT}}\\ \mathbf{L}_{\text{prec}} &= -\mu_0\gamma\, \mathbf{m}\times\mathbf{H}_{\text{eff}}/(1+\alpha^2)\\ \mathbf{L}_{\text{damp}} &= -\alpha\mu_0\gamma\, \mathbf{m}\times(\mathbf{m}\times\mathbf{H}_{\text{eff}})/(1+\alpha^2)\\ \mathbf{L}_{\text{STT}} &= -\frac{\mu_B I\eta}{q\Omega M_s}\mathbf{m}\times(\mathbf{m}\times\mathbf{I}_s)/(1+\alpha^2).\end{aligned} \tag{1}$$

M_s is the saturation magnetization. μ_0 is the permeability constant and γ is the gyromagnetic ratio. α is the magnetic damping coefficient. Ω is the volume of the magnet. μ_B is the Bohr magneton. q is the single electron charge. I is the charge current and $\mathbf{I}_s$ is the unit vector along the injected spin orientation. η is the spin polarization of the injected current. To include the thermal effect, $\mathbf{H}_{\text{eff}}$ is replaced by $\mathbf{H}_{\text{eff}} + \mathbf{F}_{\text{th}}$ where $\mathbf{F}_{\text{th}}$ is a random torque from thermal noise.

An alternative way to quantify the statistical nature of STT switching under thermal fluctuations is to solve the corresponding FP equation (FPE). The FP method has been applied to describe thermally agitated magnet by Brown [8]. The method can be generalized to include STT. We start from the LLG equation described earlier and assume a Gaussian distribution of random thermal noise: $\mathbf{F}_{\text{th}}$

$$\Pi(\mathbf{F}_{\text{th}}) = \frac{1}{\sqrt{8\pi^3 D}}\exp\left(\frac{-|\mathbf{F}_{\text{th}}|^2}{2D}\right). \tag{2}$$

Instead of keeping track of the random trajectory of $\mathbf{m}$, the FPE solves for the probability distribution of magnetization

$$\rho(\mathbf{m};t) = \int \Pi(\mathbf{F}_{\text{th}})\delta(\mathbf{m}-\mathbf{m}_{\mathbf{F}})d\mathbf{F}_{\text{th}}. \tag{3}$$

Taking the derivative of (3) with respect to time gives the general FPE for nanomagnet in the form of a convection-diffusion equation on a 2-D spherical surface

$$\frac{\partial\rho}{\partial t} = -\nabla\cdot(\mathbf{L}\rho) + D\nabla^2\rho \tag{4}$$

where $\rho(\theta,\phi;t)$ is the probability density of the magnetization in spherical coordinate. $\mathbf{L}(\mathbf{m}) = \mathbf{L}_{\text{prec}} + \mathbf{L}_{\text{damp}} + \mathbf{L}_{\text{STT}}$ includes all the deterministic torque terms in (1). The effective "diffusion" constant describing the thermal effect is defined as

$$D = \frac{\alpha\gamma k_B T}{(1+\alpha^2)\mu_0 M_s\Omega} \tag{5}$$

where T is the temperature and k_B is the Boltzmann constant.

We solve (4) through finite-element method with triangular meshes generated on a unit spherical surface [9], [10]. The differential operator is discretized with Galerkin's method and the time evolution is calculated through Crank Nicolson's method. The solution is a 2-D probability density $\rho(\theta,\phi;t)$ evolving on the surface of a unit sphere. The WER (or non-switching probability) is then evaluated as the total integrated probability lingering on the upper hemisphere (assuming the initial magnetization points at the $+z$ direction)

$$\text{WER} = \int_0^{\pi/2}\int_0^{2\pi}\rho(\theta,\phi;t)d\phi d\theta. \tag{6}$$

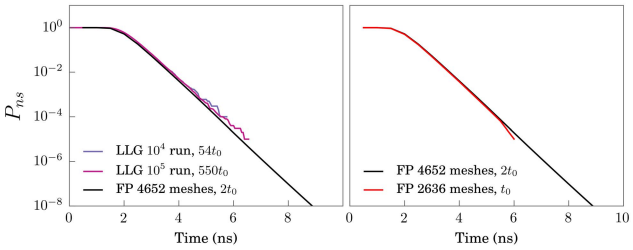

Fig. 1. Nonswitching probability P_{ns} as a function of switching time for $I = 6I_c$, where I_c is the critical current. Left: compare FP result and LLG simulations with different numbers of trials. Right: FP simulations with different numbers of grid meshes. All simulation times are normalized by the shortest one t_0 for FP simulation with 2636 meshes. The initial magnetization distribution of our LLG simulations is sampled from a Boltzmann distribution. A very fine sampling grid is needed to capture the initial magnetization close to the easy axis.

A. Computational Efficiency

To accurately describe switching distribution, (1) needs to be solved for a large number of trials, which is computationally expensive. FPE only needs to be solved once. Even though a single run of FPE takes longer than a single run of LLG, FPE simulation shows much higher efficiency and accuracy than LLG approach at low WER. Fig. 1 shows the comparison between the LLG approach and the 2-D FP approach in terms of accuracy and running time. Capturing low probability events with high accuracy in LLG method requires enormous number of runs. In comparison, FPE approach can achieve high accuracy with reasonable number of grid meshes. In a magnetic system with rotational symmetry, (4) can be easily reduced to a 1-D differential equation [11], which can be solved even more efficiently.

B. Supercritical/Subcritical Regimes

The other advantage of FPE is that it works from subcritical to supercritical regimes. At those two limits, the FPE can even be solved quasi-analytically in some cases [11]. Here, we compare the FPE with widely used analytical models and experiments.

In the supercritical regime, Sun's approach [12] is very popular due to its simplicity. In Sun's equation, the switching probability is a double exponential function of current and pulsewidth

$$P_{\text{sw}} = \exp\{-4\Delta\exp[-2\tau(i-1)]\} \tag{7}$$

where i, τ, and Δ are scaled quantities defined as

$$\begin{aligned}i &= \frac{I}{I_c},\quad \tau = \frac{t}{\tau_D},\quad \Delta = \frac{\mu_0 H_k M_s\Omega}{2k_B T}\\ I_c &= \frac{2\alpha q}{\eta\hbar}\mu_0 H_k M_s\Omega,\quad \tau_D = \frac{1+\alpha^2}{\alpha\gamma\,\mu_0 H_k},\quad H_k = \frac{2K_u}{\mu_0 M_s}\end{aligned} \tag{8}$$

with I_c the critical current and Δ the thermal stability factor. K_u and H_k are anisotropy constant and anisotropy field,

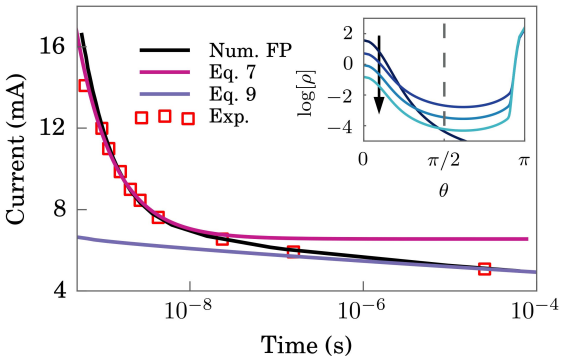

Fig. 2. Average switching current (P_{sw} = 0.5) as a function of pulsewidth in a 100-nm spin-valve nanopillar. The experimental data are extracted from [15, Fig. 7(a)]. Inset: time evolution of probability distribution at t = 0.1,5.0,10, and 15 ns with current I = 6.55 mA. Arrow: increasing time.

respectively. The physical picture described by Sun's equation is a thermally disturbed magnetization that is switched by an overdrive current in a *deterministic* way. In other words, Sun's equation considers the equilibrium thermal distribution of initial magnetization before switching, but neglects the thermal fluctuation during magnetization switching. An overdrive current $I \gg I_c$ is also a necessary condition for the switching time approximation implicit in Sun's equation [12]. These assumptions work well in the supercritical regime, because a large driven spin torque will diminish any thermal effect during the switching process. On the other hand, Sun's equation cannot describe the intermediate or subcritical regime where the applied current is close to or below the critical current.

In the subcritical regime, switching happens due to thermal agitation. The switching probability can be described by an empirical thermal transition model

$$P_{sw} = 1 - \exp\{-tf_0 \exp[-\Delta(1-i)^{\beta}]\} \tag{9}$$

where f_0 is an empirical attempt frequency set between $10^9 \sim 10^{10}$ Hz to describe experimental data in most magnets. The term $\Delta(1-i)^{\beta}$ represents an effective scaled energy barrier between the two stable states, in this case the two orientations along the easy axis. Parameter $\beta = 1$ is for in-plane systems [13] and $\beta = 2$ is for perpendicular magnetic anisotropies (PMAs) [11]. In the thermal regime, the current contributes to switching by reducing the effective barrier.

One example to show the applicability of FPE in all regimes is to plot the average switching current as a function of pulsewidth. The average switching current is defined when the switching probability equals half: $P_{sw}(I_{sw}) = 0.5$ and this average is usually plotted as a function of pulsewidth. Fig. 2 shows the $I_{sw} - t$ relation for a 100-nm-diameter metallic spin valve measured in [14] and [15]. The experimental data are fitted by (7) in the large current regime, (9) in the small current regime, and FP in all regimes. The physical parameters fitted here are the thermal stability factor Δ, critical current I_c, and τ_D. In the case of (9) in the thermal regime, f_0 is set to a typical value of 10^9 Hz. β is also treated as a fitting parameter to account for nonidealities. Fitting with (7) and (9) gives $\Delta = 63$, $I_c = 6.55$ mA, $\tau_D = 0.26$ ns, and $\beta = 1.2$. FP fitting gives $\Delta = 80$, $I_c = 8.3$ mA, and $\tau_D = 0.25$ ns. Given the differences discussed before, this discrepancy is reasonable. Although the switching behavior is not ideal macrospin (see detailed discussions in [15]), the match is enough to show the applicability of FPE method throughout different regimes.

IV. Write Error Rate

In developing energy-efficient STTRAMs, a great deal of effort has been devoted into reducing the critical current I_c, which is a primary performance factor. In the meantime, practical STT devices demand fairly reliable switching with error rates set by the specific applications. We focus here on WER defined earlier. The WER is usually plotted as a function of voltage/current for a given pulsewidth. Such a plot contains multiple messages. First, the plot indicates the onset of switching, which can indicate the average switching current. The plot also sets the approximate boundary for read current, since a switching event in read operation causes read error. Second, the lowest error rate achieved at certain current limit will determine the size of the memory array. Third, when the error rate is plotted in logarithmic scale, the slope of the "error tail" shows how fast the error rate goes down with increasing current, which characterizes the error margin for the write operation. In general, WER has a complex dependence on current, pulsewidth, and material properties. Current and pulsewidth are often determined by the application. Therefore, we will focus our discussion on a few other related parameters: temperature, anisotropy and damping, saturation magnetization, and current polarization. Some parameters are prone to change and hard to control (e.g., temperature rise due to joule heating) while others are often tuned by experimentalists to achieve better performance. We hope to shed some light on how those parameters affect the WER. For the convenience of discussion, we have chosen a set of experimental data from [16] and constructed a "reference" device—a perpendicular CoFeB MTJ. The physical parameters are shown in Table I. All the following discussions assume a fixed pulsewidth of $\tau_{pw} = 10$ ns. Before going into the discussion, it is worth mentioning that in general, specific WER targets need to be met by the memory array or various applications to build a viable product. The obvious way to achieve low WER is to increase the voltage but that comes with a power consumption penalty and more importantly the probability of dielectric breakdown over time. Switching efficiency is another aspect that can help but so far the highest efficiency is achieved in MTJs with smaller diameter [17], which is accompanied by very high resistance. In this paper, we focus on the effect of material parameters on the write error and write margin. The actual WER engineering requires diligent handing of all such factors. Nevertheless, the approach is quite general and can be easily applied to other discussions.

Let us now discuss the parameter dependences on the MTJ WERs. In the following WER-V plots, the voltages have been scaled by the average switching voltage $V_{sw}^{ref}(\text{WER} = 0.5)$ of the reference device to deemphasize their exact values but to focus on the general trend.

A. Temperature

Practical applications of STT are inevitably tied to stochastic switching at finite temperature. Indeed, STTs are prone

TABLE I
MTJ Parameters

parameter	value	remark
d	40 nm	free layer diameter
t	1 nm	free layer thickness
M_s	1.23×10^6A/m	saturation magnetization
α	0.027	magnetic damping
Δ	43	thermal stability
η	0.6	spin polarization
RA	18 $\Omega \cdot \mu$m^2	resistance area product

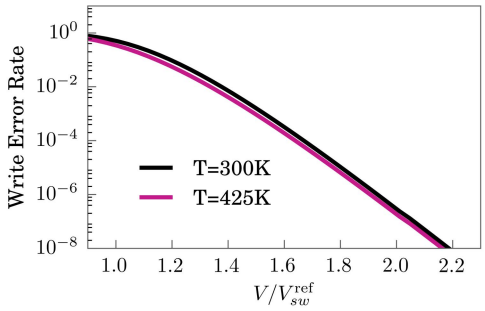

Fig. 3. WER as a function of junction bias for different temperatures.

to environmental change and joule heating during the writing process. The impact of thermal fluctuations on WERs is critical to the operation of STT-based devices, the main effects being a change in the initial magnetization distribution and a change in the energy barrier $\Delta = E_b/k_BT$. Fig. 3 shows that the impact on WER of a drastic change in temperature is quite minimal. More complicated current-driven temperature change could arise due to self-heating effects that are beyond the scope of this paper [18]. The weak dependence of WERs on temperature variation can be understood by recognizing that these simulations are performed in the fast switching regime (10 ns), during which the chance of thermal switching is very low. Therefore, temperature variation only affects the initial magnetization distribution. This effect can be estimated from the Sun's equation (7). One can show that the temperature difference from 300 to 425 K only changes log[WER] by about log[425 K/300 K] from the Taylor series expansion of (7) at WER $\to$ 0 (or $P_{\text{sw}} \to 1$) limit.

B. Anisotropy, Damping, and Saturation Magnetization

Besides the physical dimensions of the magnet, anisotropy, magnetic damping, and saturation magnetization are three important material parameters experimentalists often work with. Improvements in fabrication techniques as well as material modeling can give better control over these parameters over time. Fig. 4 shows that magnetic damping α and anisotropy field H_k have nearly identical effects on the WER, assuming M_s is held constant. Reducing either parameter would reduce the switching current. This is expected, since these are the physical forces that oppose magnetic switching [19]. Fig. 4 also shows that reducing α or H_k does not change the slope of the WER-V curve, meaning the noise margin remains the same even when the average switching current is changed.

Fig. 5 shows how saturation magnetization M_s affects the WER. Since M_s and H_k are related through (8), two different scenarios emerge. In the left plot of Fig. 5, H_k is kept

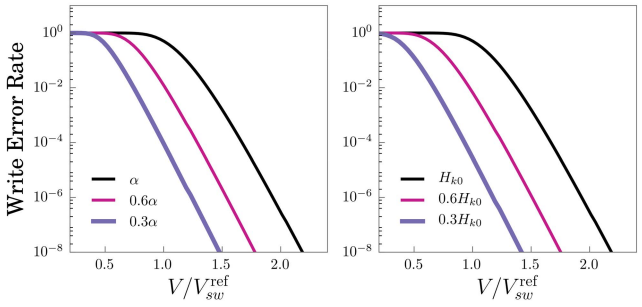

Fig. 4. WER-V plot for various magnetic dampings and anisotropies. Left: WER-V for different magnetic dampings α. Right: WER-V for different magnetic anisotropy fields H_k.

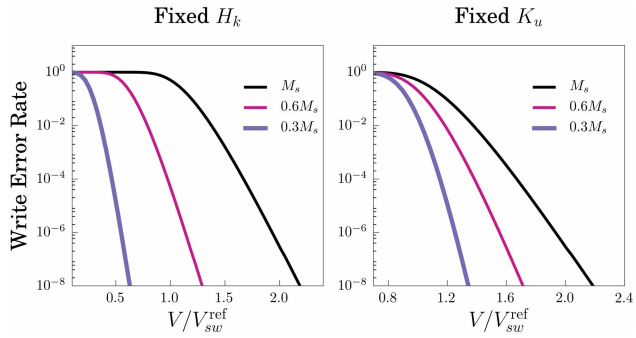

Fig. 5. WER-V plot for various saturation magnetizations. Left: WER-V with fixed H_k. Right: WER-V with fixed K_u.

fixed. This means the overall magnetic anisotropy constant K_u changes with M_s [see (8)], a reduction in M_s reduces the switching current and increases the slope. In the right plot, K_u is kept unchanged while H_k changes with M_s accordingly. The average switching current is altered moderately while the slope of WER-V curve changes in a similar way as the left plot. It is important to notice that in the left scenario, the change of saturation magnetization M_s would change the thermal stability Δ as well but in the right scenario (fixed K_u), Δ remains unchanged for different M_s values according to (8).

C. Spin Polarization

The degree of spin polarization is also a critical determinant of most spintronic applications, although affected by many factors and hard to control. In the case of MTJ, it is spin filtering by the fixed magnet that imposes a torque on a noncollinear free magnet. Typically, the spin polarization is determined by the materials used but can be largely affected by the interfacial configurations, such as symmetry filtering oxides [20], [21], defects, and strain [22], [23]. Fig. 6 plots the WER-V for various spin polarizations. Understandably, increased polarization lowers the critical current [from (8)] as well as the average switching current (WER = 0.5). At the same time, the slope of WER-V increases with spin polarization, resulting in a narrower write margin. This improvement of performance saturates with $\eta \to 1$.

In the real world, it is hard to change one parameter at a time. Indeed, many of the parameters discussed earlier are, in fact, correlated. However, it is not difficult to comprehend their combined effect on the write process. Notice that the average switching current I_{sw} is mostly determined by the

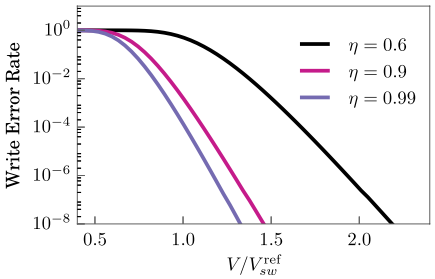

Fig. 6. WER as a function of junction bias for various polarizations. Higher polarization allows lower critical current and a sharper slope.

TABLE II

WER Slope for Different Parameters

	parameter value	decades per 100 mV		ratio of FPE to $\mathcal{S}_0$
		FPE*	eq. 10	
	reference	1.26	1.56 ($\mathcal{S}_0$)	0.99
Fig. 3	$T = 425$ K	1.26	1.56	0.99
Fig. 4. left	0.6α	1.33	1.56	1.05
Fig. 4. left	0.3α	1.33	1.56	1.05
Fig. 4. right	$0.6H_k$	1.29	1.56	1.02
Fig. 4. right	$0.3H_k$	1.32	1.56	1.05
Fig. 5	$0.6M_s$	2.09	2.61	1.66
Fig. 5	$0.3M_s$	4.10	5.22	3.25
Fig. 6	$\eta = 0.9$	1.92	2.35	1.52
Fig. 6	$\eta = 0.99$	2.10	2.58	1.67

*slope extracted from numerical FPE simulations.

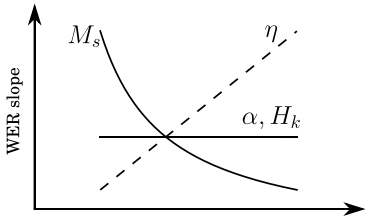

Fig. 7. Schematic for the WER slope dependence on various physical parameters.

intrinsic critical current I_{c0}. From (8), it is easy to see that $\alpha, \eta, M_s H_k \propto K_u$ are the factors determining the critical current. On the other hand, The slope of WER-V curve is a function of voltage/current but its asymptotic value at small WER can be approximated from Sun's equation (7) in the limit WER$\rightarrow$ 0 where we get

$$\mathcal{S} := -\frac{d\log[\mathrm{WER}]}{dV} \approx \frac{\mu_B}{M_s \Omega}\frac{\eta}{1+\alpha^2}\frac{t}{qR} \tag{10}$$

with R the junction resistance. Table II shows that (10) can be used to approximate the WER slope and its dependence on the physical parameters, which is summarized qualitatively in Fig. 7.

V. 2-D Fokker–Planck for Nonsymmetric Systems

The key advantage of the general 2-D FPE approach is that WER can be well captured without the need for large number of stochastic LLG runs. It also allows the study of a plethora of MTJ designs that break the rotational symmetry of PMA

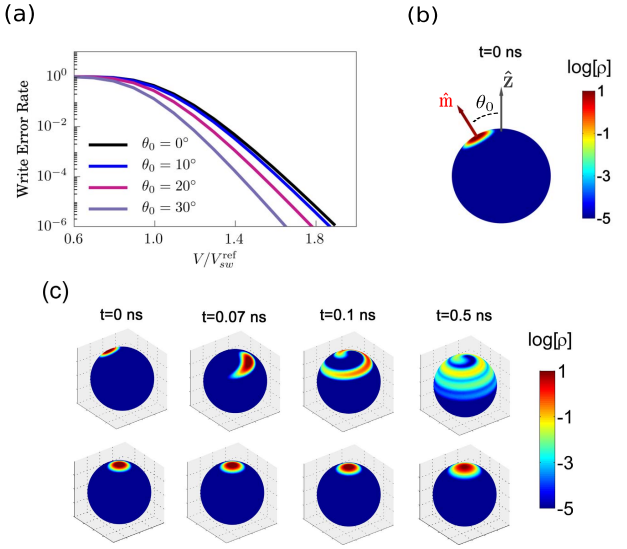

Fig. 8. (a)WER of magnetization switching from an initial angle. (b) Initial distribution of magnetization probability $\rho(\theta, \phi)$ on a unit sphere. (c) Time evolution of the probability density for canted case (top) with $\theta_0 = 30°$ and collinear case (bottom) with $\theta_0 = 0°$ at 1.25 V_{sw}^{ref} junction bias.

to achieve low WER. In those cases, the existing 1-D FPE approach cannot be used. One example is to use extra forces to nudge the initial magnetization out of its easy axis to facilitate switching. These proposed schemes include a thermal torque applied to excite the magnetization to a large angle [24], as well as an in-plane spin-orbital torque arising from a spin Hall effect to disturb the initial magnetization [25]. To capture those dynamics, we have to go beyond 1-D FPE. We leave the details on magnetization switching and subsequent dynamics to future studies and focus here on the numerical outcomes of the 2-D FPE. Fig. 8 shows the WER of our reference system but with a "tilted" initial angle. The assumption is that the initial magnetization still obeys a Boltzmann distribution except the maximum probability density is inclined at an angle θ_0 with respect to the z-axis (easy axis) shown in Fig. 8(b). Fig. 8(a) shows that the initial excitation (usually fast, $<$1 ns, so any overhead delays can be neglected) helps reduce both the critical write current and write noise margin. Fig. 8(c) further shows the different types of switching behavior in a canted system versus a collinear system. Whereas in the collinear case, the initial magnetization distribution remains almost unchanged for some time and starts to "diffuse" to the $-z$-direction; in the canted case, the magnetization starts to process and move to the $-z$-direction almost immediately after the current is applied. Therefore, the initial delay in the collinear case is avoided in a canted system. We should caution that our assumption of the Boltzmann distribution of the tilted initial angle is likely oversimplified. Any nonsymmetric force, such as the stray field, could break the symmetry and result in a non-Boltzmann distribution. In that case, a separate simulation can be performed to obtain the initial distribution before switching. Including different torques from external magnetic fields or spin-orbital couplings is quite straightforward in the general 2-D FPE. The purpose here is to show a fast numerical result based on the 2-D FPE with a simplified assumption, relegating further details to future publications.

VI. Conclusion

In summary, we have discussed thermal noise-induced WERs in PMA systems with a numerical simulation of the 2-D FPE. The effects of several material parameters on write error are investigated. Among them, spin polarization and saturation magnetization change both average switching current and write error margin, while magnetic damping and anisotropy field only affect the switching current. It is worth mentioning that some experimental works on MTJ show better WERs than the prediction from macrospin model. We believe it could be due to the subvolume excitation effect [6], which is not captured in our current simulation. One possible direction for future work is to modify the current model to have multiple FPE simulations averaged over several different volumes to make up the magnet volume that might capture subvolume excitations. In other hybrid switching schemes with broken cylindrical symmetry, 2-D FP can be a very useful and efficient tool to study the WERs.

Acknowledgment

The authors would like to thank Prof. W. H. Butler and P. Visscher from the University of Alabama for discussions on the Fokker–Planck methods. They would also like to thank Dr. K. Lee for valuable discussions on spin-transfer torque-based magnetic random access memory write error.

References

[1] J. C. Slonczewski, "Current-driven excitation of magnetic multilayers," *J. Magn. Magn. Mater.*, vol. 159, nos. 1–2, pp. L1–L7, 1996.
[2] L. Berger, "Emission of spin waves by a magnetic multilayer traversed by a current," *Phys. Rev. B*, vol. 54, no. 13, p. 9353, 1996.
[3] A. Fukushima *et al.*, "Spin dice: A scalable truly random number generator based on spintronics," *Appl. Phys. Exp.*, vol. 7, no. 8, pp. 3–7, 2014.
[4] A. F. Vincent *et al.*, "Spin-transfer torque magnetic memory as a stochastic memristive synapse for neuromorphic systems," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 2, pp. 166–174, Apr. 2015.
[5] A. F. Vincent, N. Locatelli, J. O. Klein, W. S. Zhao, S. Galdin-Retailleau, and D. Querlioz, "Analytical macrospin modeling of the stochastic switching time of spin-transfer torque devices," *IEEE Trans. Electron Devices*, vol. 62, no. 1, pp. 164–170, Jan. 2015.
[6] J. Z. Sun *et al.*, "Effect of subvolume excitation and spin-torque efficiency on magnetic switching," *Phys. Rev. B*, vol. 84, no. 6, p. 064413, 2011.
[7] K. Song and K. J. Lee, "Spin-transfer-torque efficiency enhanced by edge-damage of perpendicular magnetic random access memories," *J. Appl. Phys.*, vol. 118, no. 5, pp. 17–21, 2015.
[8] W. F. Brown, Jr., "Thermal fluctuations of a single-domain particle," *Phys. Rev.*, vol. 130, no. 5, pp. 1677–1686, Jun. 1963.
[9] K. Zhang and D. R. Fredkin, "Thermal switching of a stoner-wohlfarth particle: Numerical solution of Fokker-Planck equation without symmetry," Ph.D. dissertation, Dept. Phys., Univ. California, San Diego, CA, USA, 2001.
[10] P.-O. Persson and G. Strang, "A simple mesh generator in MATLAB," *SIAM Rev.*, vol. 46, no. 2, pp. 329–345, 2004.
[11] W. H. Butler *et al.*, "Switching distributions for perpendicular spin-torque devices within the macrospin approximation," *IEEE Trans. Magn.*, vol. 48, no. 12, pp. 4684–4700, Dec. 2012.
[12] J. Z. Sun, T. S. Kuan, J. A. Katine, and R. H. Koch, "Spin angular momentum transfer in a current-perpendicular spin-valve nanomagnet," *Proc. SPIE*, vol. 5359, no. 1, pp. 445–455, Jul. 2004.
[13] Z. Li and S. Zhang, "Thermally assisted magnetization reversal in the presence of a spin-transfer torque," *Phys. Rev. B*, vol. 69, no. 13, pp. 1–6, 2004.
[14] D. Bedau *et al.*, "Ultrafast spin-transfer switching in spin valve nanopillars with perpendicular anisotropy," *Appl. Phys. Lett.*, vol. 96, no. 2, p. 022514, 2010.
[15] H. Liu *et al.*, "Dynamics of spin torque switching in all-perpendicular spin valve nanopillars," *J. Magn. Magn. Mater.*, vols. 358–359, pp. 233–258, May 2014.
[16] S. Ikeda *et al.*, "A perpendicular-anisotropy CoFeB–MgO magnetic tunnel junction," *Nature Mater.*, vol. 9, pp. 721–724, Jul. 2010.
[17] M. Gajek *et al.*, "Spin torque switching of 20 nm magnetic tunnel junctions with perpendicular anisotropy," *Appl. Phys. Lett.*, vol. 100, no. 13, p. 132408, 2012.
[18] R. C. Sousa, "Tunneling hot spots and heating in magnetic tunnel junctions," *J. Appl. Phys.*, vol. 95, no. 11, p. 6783, 2004.
[19] J. Z. Sun, "Spin-current interaction with a monodomain magnetic body: A model study," *Phys. Rev. B*, vol. 62, no. 1, pp. 570–578, 2000.
[20] W. H. Butler, X.-G. Zhang, T. C. Schulthess, and J. M. MacLaren, "Spin-dependent tunneling conductance of Fe| MgO| Fe sandwiches," *Phys. Rev. B*, vol. 63, no. 5, p. 054416, Jan. 2001.
[21] Y. Xie, I. Rungger, K. Munira, M. Stamenova, S. Sanvito, and A. W. Ghosh, "Spin transfer torque: A multiscale picture," in *Nanomagnetic Spintronic Devices for Energy-Efficient Memory Computing*, J. Atulasimha and S. Bandyopadhyay, Eds. Hoboken, NJ, USA: Wiley, 2016, pp. 91–132.
[22] P. G. Mather, J. C. Read, and R. A. Buhrman, "Disorder, defects, and band gaps in ultrathin (001) MgO tunnel barrier layers," *Phys. Rev. B*, vol. 73, no. 20, p. 205412, May 2006.
[23] G. X. Miao, Y. J. Park, J. S. Moodera, M. Seibt, G. Eilers, and M. Münzenberg, "Disturbance of tunneling coherence by oxygen vacancy in epitaxial Fe/MgO/Fe magnetic tunnel junctions," *Phys. Rev. Lett.*, vol. 100, no. 24, p. 246803, 2008.
[24] N. N. Mojumder, D. W. Abraham, K. Roy, and D. C. Worledge, "Magnonic spin-transfer torque MRAM with low power, high speed, and error-free switching," *IEEE Trans. Magn.*, vol. 48, no. 6, pp. 2016–2024, Jun. 2012.
[25] A. van den Brink *et al.*, "Spin-Hall-assisted magnetic random access memory," *Appl. Phys. Lett.*, vol. 104, no. 1, p. 012403, 2014.

Yunkun Xie received the B.S. degree in physics from Peking University, Beijing, China, in 2012. He is currently pursuing the Ph.D. degree in electrical engineering with the University of Virginia, Charlottesville, VA, USA.

His current research interests include the first principle material modeling of transport on magnetic tunnel junctions, the device simulation of nano-magnetic memory/logic including spin-transfer torque Random Access Memory and multiferroics.

Behtash Behin-Aein received the Ph.D. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, USA, in 2010.

He is currently with GLOBALFOUNDRIES, US, Inc., Santa Clara, CA, USA. He has been a Lead Industry Liaison for Semiconductor Research Corporation on spin-based devices, a Lead Principal Investigator for the Department of Energys Advanced Scientific Computing Research Leadership Computing Challenge Award, a Distinguished Industry Associate for the center for spintronics materials interfaces, and novel architectures and an Invited Author for Material Research Society bulletin.

Avik W. Ghosh (SM'11) is Professor of ECE and Physics at University of Virginia. He has a PhD in physics, Postdoctoral Fellowship in Electrical Engineering, over 90 refereed publications, 2 upcoming books and is Fellow of the Institute of Physics, and has received the IBM Faculty Award and the NSF CAREER Award.