Switching Time of Spin-Torque-Driven Magnetization in Biaxial Ferromagnets

Ankit Shukla[1,*] Arun Parthasarathy,[2,†] and Shaloo Rakheja[1,‡]

[1]*Holonyak Micro and Nanotechnology Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*
[2]*Department of Electrical and Computer Engineering, New York University, Brooklyn, New York 11201, USA*

We analytically model the magnetization switching time of a biaxial ferromagnet driven by an anti-dampinglike spin torque. The macrospin magnetization dynamics is mapped to an energy-flow equation, wherein a rational-function approximation for the elliptic integrals for moderate spin current and small damping results in a closed-form expression for the switching time. The randomness in the initial angle of magnetization gives the distribution function of the switching time. The analytical model conforms to results obtained from Monte Carlo simulation for a broad range of material parameters. Our results can improve the design and benchmarking of in-plane spin-torque magnetic memory by obviating the need for expensive numerical computation.

I. INTRODUCTION

Current-induced spin phenomena, such as spin-transfer torque (STT) and spin-orbit torque (SOT), allow electrical manipulation of magnetic order, and form the basis of emerging spintronic technologies such as non-volatile memory [1], magnonic interconnects [2], and radio-frequency oscillators [3]. A spin current can transfer angular momentum to a magnetic layer and reorient its magnetization, similarly to how an electric current can transfer charge to a capacitor and modulate its voltage. As shown in Fig. 1(a), an electric current flowing orthogonal to the plane of a spin valve becomes spin polarized in a direction parallel to the magnetization of the fixed layer. This spin-polarized current affects the magnetization of the free layer due to STT [4,5]. On the other hand, an in-plane electric current flowing through a nonmagnetic (NM) material with spin-orbit coupling is spin polarized in the plane of the nonmagnetic material but spin polarized transverse to the electric current, due to the spin Hall effect [6]. This in-plane-polarized spin current can exert a SOT [7] on the free-layer magnetization of the spin valve, as shown in Fig. 1(b).

Thin-film magnets—the path toward miniaturized spintronics—are subject to epitaxial strain from substrate and finite-size effects, which can elicit an in-plane or a perpendicular spin orientation [8]. The symmetries in the energy landscape of thin films up to quadratic (lowest-order) terms in the magnetization components are characterized by a biaxial anisotropy, consisting of an axis of minimum energy that is "easy" for spins to orient along and an orthogonal axis of maximum energy that is "hard." Uniaxial anisotropy is a special case of biaxial anisotropy, where the hard axis is absent. For perpendicularly magnetized films, a perpendicular easy axis can be used to approximate the anisotropy, assuming symmetry in the plane. For in-plane-magnetized films, an in-plane easy axis and a perpendicular hard axis offer the correct description.

In STT memory, the perpendicular free-layer configuration is superior to the in-plane configuration due to its lower switching current, faster speeds, and higher density [1]. A key problem in writing to STT memory is its vulnerability to dielectric breakdown of the tunnel barrier. This is addressed in SOT memory, where the writing occurs with an in-plane current that need not traverse the tunnel barrier. Three-terminal SOT memory separates the read and write paths, improving memory endurance at the cost of cell size. However, deterministic switching of the perpendicular free layer in SOT memory requires either a biasing magnetic field [9] or additional layers in the device stack, adding to its complexity of fabrication [10]. The in-plane free-layer configuration of SOT memory is preferred due to its simplicity of fabrication, magnetic-field-free switching, and lower switching currents [11], although its writing speed is inferior to that of perpendicular SOT memory.

When the spin polarization of the injected spin current is antiparallel to the stable orientation of the free-layer magnetization, the spin torque is antidampinglike—it competes with the intrinsic damping so as to raise the macrospin energy—until halfway in the magnetization-reversal process, when it becomes dampinglike—it contributes to damping so as to cause dissipation of macrospin energy.

---

*ankits4@illinois.edu
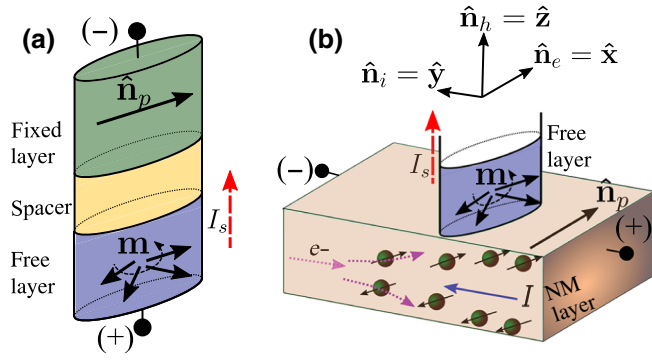†arun.parth@nyu.edu
‡rakheja@illinois.edu

FIG. 1. Injection of a spin current I_s (dashed red line) into a free magnetic layer produces (a) STT in a spin valve and (b) SOT in a spin Hall structure. **m** represents the direction of free-layer magnetization, $\hat{\mathbf{n}}_p$ represents the direction of spin polarization, and $\hat{\mathbf{n}}_e$, $\hat{\mathbf{n}}_i$, and $\hat{\mathbf{n}}_h$ are the unit vectors along the easy, intermediate, and hard anisotropy axes, respectively. The flow of electrons in the NM layer is opposite to that of the injected charge current I.

The switching process is characterized by the time required to reverse the orientation of the free layer as a function of the input spin current. A closed-form expression for the switching time is useful for designing and optimizing the performance of spin-torque memory. Previous studies [7,12–14] have derived expressions for the switching time for uniaxial anisotropy, but not for the more general biaxial anisotropy presented in this paper.

Here, we treat the magnetization of the free-layer ferromagnet as a single domain or a macrospin [15], which switches coherently due to spin torque. The macrospin approximation is valid if the dimensions of the ferromagnet are smaller than the characteristic domain-wall width, typically about 50 nm for Co, Fe, and Ni [16,17]. The macrospin model has been applied successfully to explain measured probability distributions of the switching dynamics of ferromagnetic thin films with dimensions on the order of 50 nm [18]. Inhomogeneous switching dynamics, such as that due to domain-wall nucleation and propagation, has been experimentally measured in large ferromagnetic samples [19,20]. However, consideration of multidomain switching of ferromagnets is outside the scope of this paper.

Within the macrospin approximation, there are three equivalent approaches to analyzing the magnetization dynamics: the perturbative approach [12], constructing a Fokker-Planck representation [21,22], or using a constant-energy stochastic equation [23–26]. In this paper, we adopt constant-energy orbit averaging (CEOA) to study magnetization reversal in a biaxial magnetic system because this approach simplifies a coupled three-dimensional (3D) (two-dimensional if the magnetization has a constant magnitude) stochastic problem to a tractable one-dimensional (1D) problem at low temperature and for low to moderate applied spin currents [24–26].

The switching dynamics is modeled as a slow perturbation of the rapid constant-energy gyration around the easy axis (Sec. II). In the deterministic limit, a closed-form expression for the switching time as a function of the input spin current, initial magnetization energy, and material parameters is obtained. The average switching time and the probability distribution of the switching time follow for an initial Boltzmann-distributed ensemble of spins (Sec. III). The analytical results are shown to conform to results obtained from numerical Monte Carlo simulations. A brief discussion of the stochastic dynamics and its applications is included in Sec. IV.

II. THEORY

The dynamics of a magnetization subject to an effective magnetic field, intrinsic damping, and spin torque is described by the Landau-Lifshitz-Gilbert (LLG) equation. Using the dimensionless forms of the physical parameters listed in Table I, the LLG equation is [24,27]

$$\frac{\partial \mathbf{m}}{\partial \tau} = -(\mathbf{m} \times \mathbf{h}_{\text{eff}}) - \alpha \mathbf{m} \times (\mathbf{m} \times \mathbf{h}_{\text{eff}})$$
$$- I_s \mathbf{m} \times (\mathbf{m} \times \hat{\mathbf{n}}_p) + \alpha I_s (\mathbf{m} \times \hat{\mathbf{n}}_p), \quad (1)$$

where **m** is the normalized free-layer magnetization, α is the Gilbert damping constant, I_s is the input spin current, and $\hat{\mathbf{n}}_p$ is the unit vector along the direction of spin polarization. The effective magnetic field $\mathbf{h}_{\text{eff}}$ includes contributions from an internal field produced by the magnetic anisotropy, an externally applied magnetic field ($\mathbf{H}_a$), and a thermal field. The sum of the internal and external magnetic fields, normalized to the Stoner-Wohlfarth field H_k, is given by

$$\mathbf{h} = -\frac{1}{2} \nabla_{\mathbf{m}} [g_L(\mathbf{m}, \mathbf{H}_a)], \quad (2)$$

where g_L is the free energy of the macrospin normalized to its uniaxial energy $K_u V$, where $K_u = \mu_0 M_s H_k / 2$ (μ_0 is the free-space permeability and M_s is the saturation magnetization of the free layer). Neglecting higher-order anisotropy terms,

$$g_L(\mathbf{m}, \mathbf{H}_a) = D_e (\mathbf{m} \cdot \hat{\mathbf{n}}_e)^2 + D_h (\mathbf{m} \cdot \hat{\mathbf{n}}_h)^2 - \frac{2}{H_k} \mathbf{m} \cdot \mathbf{H}_a,$$
$$(3)$$

where D_e and D_h are the effective anisotropy coefficients along the easy ($\hat{\mathbf{n}}_e$) and hard ($\hat{\mathbf{n}}_h$) axes, respectively. For thin-film magnets, the energy landscape is characterized by biaxial anisotropy with $D_e = -1$ and $D_h = M_s / H_k$. Without loss of generality, $\hat{\mathbf{n}}_e$ and $\hat{\mathbf{n}}_h$ are assumed to coincide with the $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ axes, respectively. Using Eq. (3) in Eq. (2), we obtain $\mathbf{h} = m_x \hat{\mathbf{x}} - R m_z \hat{\mathbf{z}} + \mathbf{H}_a / H_k$, where $R = D_h = M_s / H_k$.

TABLE I.   Symbols in Eq. (1) and their definitions. M_s, H_k, and V are the saturation magnetization, Stoner-Wohlfarth field, and volume, respectively, of the free layer. μ_0 is the free-space permeability, γ is the gyromagnetic ratio, and α is the Gilbert damping constant.

Physical parameter	Dimensionless form
Magnetization **M**	$\mathbf{m} = \mathbf{M}/M_s$
Time t	$\tau = \gamma\mu_0 H_k t/(1+\alpha^2)$
Effective magnetic field $\mathbf{H}_{\text{eff}}$	$\mathbf{h}_{\text{eff}} = \mathbf{H}_{\text{eff}}/H_k$
Spin current β_s	$I_s = \beta_s/(\mu_0 M_s H_k V)$

The thermal field is a Langevin field [28] that is spatially isotropic and uncorrelated in space and time; therefore,

$$\langle \mathbf{h}_T(t) \rangle = 0, \tag{4a}$$

$$\langle \mathbf{h}_{T,p}(t_1)\mathbf{h}_{T,r}(t_2) \rangle = D\delta_{p,r}\delta(t_1 - t_2), \tag{4b}$$

where p, r represent Cartesian coordinates, $\delta_{p,r}$ is the Kronecker delta, and $\delta(t)$ is the Dirac delta. Assuming the macrospin to be in thermal equilibrium with a thermal bath and neglecting Joule heating, the diffusion coefficient is given by

$$D = \frac{\alpha k_B T}{(1+\alpha^2) K_u V} = \frac{\alpha}{(1+\alpha^2)\Delta_0}, \tag{5}$$

where k_B is the Boltzmann constant, T is the temperature of the bath, and Δ_0 is the barrier height of a uniaxial-anisotropy magnet and measures the thermal stability of the macrospin [28]. Here, Δ_0 is normalized to the thermal energy $k_B T$.

A. Constant-energy orbit averaging

The energy of a macrospin is conserved when damping, a thermal field (noise), an external magnetic field, and spin torque are absent. In this case, **m** precesses around the easy axis on the unit-magnetization sphere with a fixed macrospin energy g_L (< 0). Trajectories of conserved motion, illustrated in Fig. 2(a), are obtained by solving Eqs. (1) and (3) with α, $\mathbf{H}_a$, and I_s set equal to zero, and $\|\mathbf{m}\| = 1$. However, with finite damping, **m** loses energy, eventually relaxing to a stable equilibrium state ($m_x = \pm 1$). A nonzero spin torque can pump energy to the macrospin and act against its inherent damping, causing the magnetization to deviate from its equilibrium position.

In the case of a small to moderate input spin torque, small damping, and low temperature, two distinct timescales of magnetization dynamics emerge: (i) a fast timescale associated with constant-energy gyration around the easy axis and (ii) a slow timescale corresponding to perpendicular diffusion of magnetization from one
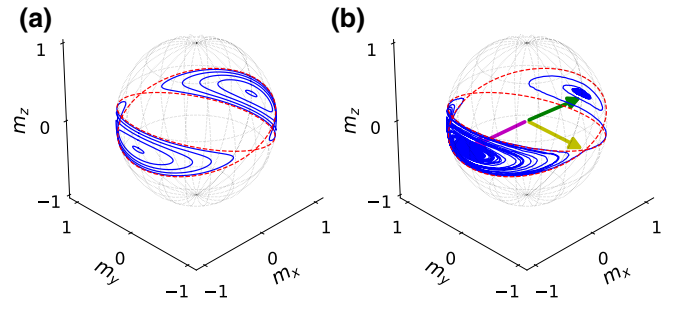

FIG. 2.   (a) Constant-energy curves for $g_L < 0$ in both antiparallel and parallel wells for $R = 15$. (b) Switching trajectory of a magnetization initially in the antiparallel well for $R = 15$ and a large input spin current. The external field is assumed to be absent. The arrows shown in the figure mark the initial (magenta), intermediate (yellow), and final (green) positions of the magnetization, in a counterclockwise sense. The dashed red curves in both parts correspond to a zero-energy separatrix.

constant-energy orbit to another as a result of damping, spin torque, and the thermal field. Figure 2(b) shows one such trajectory, where the magnetization switches from an antiparallel well to a parallel well under the simultaneous effects of damping, an input spin current, and a thermal field.

The rate of change of the macrospin energy due to the nonconservative torques is given by

$$\frac{\partial g_L}{\partial \tau} = \nabla_{\mathbf{m}} g_L \cdot \frac{\partial \mathbf{m}}{\partial \tau} = 2\left(Rm_z\frac{\partial m_z}{\partial \tau} - m_x\frac{\partial m_x}{\partial \tau}\right). \tag{6}$$

Averaging the above equation over one time period of the undamped motion reduces the coupled stochastic dynamics of Eq. (1) to a 1D stochastic dynamics as follows [24,26] (see Appendix A):

$$\left\langle\frac{\partial g_L}{\partial \tau}\right\rangle = \frac{\pi\alpha}{K(R,g_L)}\sqrt{\frac{R-g_L}{1+R}}$$
$$\times\left\{\frac{I_s}{\alpha}(1+g_L) - \frac{2}{\pi}\sqrt{(1+R)(R-g_L)}\right.$$
$$\times[E(R,g_L) + g_L K(R,g_L)]\Bigg\} + 2\sqrt{\frac{\alpha}{\Delta_0}}$$
$$\times\sqrt{\frac{(R-g_L)}{K(R,g_L)}}\sqrt{E(R,g_L) + g_L K(R,g_L)} \circ \dot{W}_{g_L}. \tag{7}$$

Here, $\circ$ denotes multiplication of thermal noise in the Stratonovich sense [24,28], while $K(R,g_L)$ and $E(R,g_L)$ are the complete elliptic integrals of the first and second kind, respectively. Time averaging of $\partial g_L/\partial\tau$ over a period of precessional motion enables us to study the dynamics due to slow diffusion of energy with respect to fast periodic oscillations [24,26].

Assuming deterministic dynamics and using Eq. (7) to evaluate this dynamics for zero energy flow at $g_L = 0$ and $g_L = -1$ leads to two different threshold currents [12,22,24]: the minimum current required to push the magnetization over the energy barrier into the adjoining basin, $I_s^{\text{th0}} = \alpha \left[(2/\pi)\sqrt{R(1+R)} \right]$, and the minimum current required to move the magnetization away from stable equilibrium, $I_s^{\text{th1}} = \alpha[R/2 + 1]$. R_c denotes the critical value of R for which I_s^{th0} equals I_s^{th1}. The threshold currents demarcate regions of deterministic switching from those that require thermal assistance, as shown in Fig. 3(a). Figures 3(b) and 3(c) show the rate of change of the energy for deterministic dynamics. This rate is positive for the complete range of macrospin energy only for $I_s^{\text{th}m} = \max\left(I_s^{\text{th0}}, I_s^{\text{th1}}\right) < I_s$, which is consistent with Fig. 3(a). For other values of the current, thermal assistance is required for switching. CEOA is valid when the variation in the macrospin energy over one precessional cycle is small, i.e., $|T\langle \partial g_L / \partial \tau \rangle| \ll \max[|g_L|] = 1$. For this constraint on the variation of macrospin energy to be satisfied, the maximum spin current is $I_s^{\text{th}M} = I_s^{\text{th0}}\left[1 + 1/\left(8\alpha\sqrt{R}\right)\right]$.

III. ANALYTICAL SWITCHING-TIME MODEL AND SWITCHING-TIME DISTRIBUTIONS

The switching time due to spin torque is defined as the time required for the macrospin energy to change from an initial value g_{L_i} to its final value g_{L_f}. The analytical solutions for the switching time presented in this section rely on one or more of the following rational approximations [29] for the elliptic integrals in Eq. (7):

$$\frac{2}{\pi}\left[E(R, g_L) + g_L K(R, g_L)\right] = A(R)g_L^2 + B(R)g_L + C(R),$$
(8a)

$$K(x) = \frac{\pi}{2}\left[\frac{x-4}{2x-4}\right],$$
(8b)

$$E(x) = \frac{\pi}{2}\left[1 - \frac{x}{4}\frac{x^2 - 28x + 64}{4x^2 - 40x + 64}\right],$$
(8c)

where $x = R[(1 + g_L)/(R - g_L)]$. Defining $\tilde{I}_s = I_s/\alpha$ and using Eqs. (8a) and (8b) together in Eq. (7), we obtain the magnetization switching time as follows (see Appendix B for derivation):

$$\tau_s = \frac{1}{4\alpha}\int_{g_{L_i}}^{g_{L_f}}\left[\frac{3R - g_L(R+4)}{R - g_L(R+2)}\right]\sqrt{\frac{1+R}{R - g_L}}$$

$$\times \frac{\partial g_L}{\left[\tilde{I}_s(1 + g_L) - \sqrt{1+R}\sqrt{R - g_L}\left(Ag_L^2 + Bg_L + C\right)\right]}.$$
(9)

TABLE II. The parameters A, B, and C in Eq. (9) are approximated as $k_1 + k_2 R^{k_3}$, where k_1, k_2, and k_3 depend on R.

R		k_1	k_2	k_3
[1, 3]	A	0.35661	−0.51244	−0.38689
	B	1.05148	−0.55504	−0.28598
	C	0.61670	0.03018	−1.00153
[3, 50]	A	0.20223	−0.38439	−0.68424
	B	0.81746	−0.34729	−0.63939
	C	0.61765	0.02994	−1.08243
[50, 100]	A	0.17370	−0.51992	−0.97986
	B	0.78501	−0.48295	−0.97726
	C	0.61755	0.02625	−1.01442

In the above equation, the parameters A, B, and C are functions of R, given by $k_1 + k_2 R^{k_3}$, where the values of k_1, k_2, and k_3 are chosen for different intervals of R to reduce the error in approximating the elliptic integrals. See Table II for details. Note that the values of A, B, and C are independent of the device geometry and depend only on $R = M_s/H_k$. Equation (9) can be simplified and integrated using partial fractions to arrive at the following closed-form expression for the switching time:

$$\tau_s = \frac{1}{2\alpha(R+2)}\left[\sum_{i=1}^{5}\frac{N}{D}\log\left[\frac{(R - g_{L_f}) - \lambda_i}{(R - g_{L_i}) - \lambda_i}\right]\right.$$

$$+ \sqrt{\frac{R(1+R)}{R+2}}\left(\frac{\log\left[\frac{\sqrt{(R - g_{L_f})(R+2)} - \sqrt{R(1+R)}}{\sqrt{(R - g_{L_i})(R+2)} - \sqrt{R(1+R)}}\right]}{A\prod_{n=1}^{5}\left[\frac{R(1+R)}{R+2} - \lambda_i\right]}\right.$$

$$\left.\left.+ \frac{\log\left[\frac{\sqrt{(R - g_{L_f})(R+2)} + \sqrt{R(1+R)}}{\sqrt{(R - g_{L_i})(R+2)} + \sqrt{R(1+R)}}\right]}{A\prod_{n=1}^{5}\left[\frac{R(1+R)}{R+2} + \lambda_i\right]}\right)\right],$$
(10)

where the λ_i are the roots of a fifth-degree polynomial $x^5 - (B/A + 2R)x^3 + (\tilde{I}_s/A\sqrt{1+R})x^2 + [R^2 + (B/A)R + C/A]x - \tilde{I}_s\sqrt{1+R}/A$, and $N = (R+4) + \{2R(1+R)/[(R+2)\lambda_i^2 - R(1+R)]\}$ and $D = 5A\lambda_i^4 - 3(B + 2AR)\lambda_i^2 + 2(\tilde{I}_s/\sqrt{1+R})\lambda_i + (AR^2 + BR + C)$.

A major advantage of this analytical result is that the approximations for the elliptic integrals are independent of the input spin current and Gilbert damping. Additionally, the results obtained in this paper are valid for a broad range of R, as opposed to those in prior studies [12,30], which are valid only for $R < R_c$ (= 5.09).
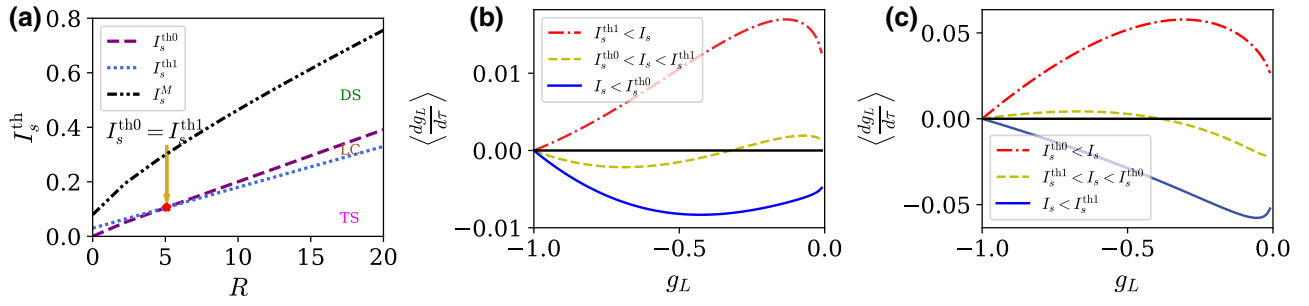
FIG. 3. (a) Threshold spin current versus R. The arrow marks $R = R_c = 5.094$, where $I_s^{th0} = I_s^{th1}$. For $I_s^{thm} < I_s$, deterministic switching (DS) occurs. The thermal-switching region is marked TS, while the regime of limit cycles is marked LC. (b) For $R = 3.0$ and $I_s^{th1} < I_s$, the rate of change of macrospin energy is positive for $g_L \in (-1, 0)$. For $I_s^{th0} < I_s < I_s^{th1}$, thermal assistance is required to cause magnetization reversal. For $I_s < I_s^{th0}$, switching occurs due to thermal noise. (c) For $R = 15.0$ and $I_s^{th} < I_s$, reversal is deterministic, as the rate of energy change is always positive. For $I_s < I_s^{th1}$, thermal switching dominates, while in the intermediate region, thermal assistance is required to cause switching. These results correspond to $\alpha = 0.03$.

In the case where $R \to \infty$, the elliptic integrals in Eq. (7) are approximated using Eqs. (8b) and (8c). This simplifies the energy-flow equation, and the switching time is given by (see Appendix C for details)

$$
\tau_s = \frac{1}{2\alpha \left[\tilde{I}_s - 0.5R \right]} \left\{ \log \left[\frac{1 + g_{L_f}}{1 + g_{L_i}} \right] \right.
$$
$$
+ \frac{b(a-4)(a-8)}{32(a-b)} \log \left[\frac{1 + g_{L_f} - a}{1 + g_{L_i} - a} \right]
$$
$$
\left. - \frac{a(b-4)(b-8)}{32(a-b)} \log \left[\frac{1 + g_{L_f} - b}{1 + g_{L_i} - b} \right] \right\}, \quad (11)
$$

where a and b are the roots of the quadratic equation $x^2 - \left[(160\tilde{I}_s - 60R)/(16\tilde{I}_s - 7R) \right] x + \left[(256\tilde{I}_s - 128R)/(16\tilde{I}_s - 7R) \right] = 0$.

In the limit of $R \to 0$, the free-energy density in Eq. (3) simplifies to $g_L = -(\mathbf{m} \cdot \hat{\mathbf{n}}_e)^2$, which represents a uniaxial-anisotropy ferromagnet. In this case, the switching time is given by [12–14,24,31] (see Appendix D for details)

$$
\tau_s = \frac{1}{2\alpha(\tilde{I}_s^2 - 1)} \left\{ \tilde{I}_s \left[\log \left(\frac{1 + \sqrt{-g_{L_i}}}{1 + \sqrt{-g_{L_f}}} \right) \right. \right.
$$
$$
\left. - \log \left(\frac{1 - \sqrt{-g_{L_i}}}{1 - \sqrt{-g_{L_f}}} \right) \right] - \log \left(\frac{1 + g_{L_i}}{1 + g_{L_f}} \right)
$$
$$
\left. + 2 \log \left(\frac{\tilde{I}_s - \sqrt{-g_{L_i}}}{\tilde{I}_s - \sqrt{-g_{L_f}}} \right) \right\}. \quad (12)
$$

A. Equilibrium Distribution and Average Switching Time

In the absence of an input spin current, the magnetization is considered to be in thermal equilibrium in its stable energy well. An average switching time, $\langle \tau_s \rangle$, is obtained by averaging τ_s over the equilibrium energy distribution, which in the case of a large energy barrier is the Boltzmann distribution, given by

$$
w_{eq}(\mathbf{m}) = \frac{1}{Z(\Delta_0, R)} \exp \left[-\Delta_0 g_L(\mathbf{m}, 0) \right], \quad (13)
$$

where $Z(\Delta_0, R)$ is the partition function. We evaluate the Boltzmann distribution function in terms of the free-energy random variable (see Appendix E for details of the transformation). Accordingly, the probability density function (PDF) is

$$
\rho(g_L) = \frac{1}{Z(\Delta_0, R)} \frac{\exp(-\Delta_0 g_L)}{\sqrt{-g_L}}, \quad (14)
$$

where the partition function is

$$
Z(\Delta_0, R) = \int_{-1}^{0} \rho(x) \, dx = \frac{2 \exp(\Delta_0) F(\sqrt{\Delta_0})}{\sqrt{\Delta_0}}. \quad (15)
$$

The cumulative distribution function (CDF) is

$$
P(g_L) = \int_{-1}^{g_L} \rho(x) \, dx
$$
$$
= 1 - \exp\left[-\Delta_0 (1 + g_L) \right] \frac{F\left(\sqrt{-\Delta_0 g_L} \right)}{F\left(\sqrt{\Delta_0} \right)}, \quad (16)
$$

where $F(x) = \exp(-x^2) \int_0^x \exp(y^2) \, dy$ is the Dawson integral.

The magnetization is considered to have switched successfully when it has crossed the separatrix ($g_{L_f} = 0$) and consequently moved into the adjoining energy well. Once the magnetization has moved into the target energy well, the spin current can be switched off. The magnetization

eventually settles into its stable well due to its intrinsic damping. Therefore, $\langle \tau_s \rangle$ is given by

$$\langle \tau_s \rangle = \int_{-1}^{0} \rho\left(g_{L_i}\right) \tau_s \left(g_{L_i}, g_{L_f} = 0\right) dg_{L_i}. \quad (17)$$

B. Distribution Functions and Write Error Rate

If we define a random variable T_s for the switching time, the fraction of macrospins in an ensemble that have switched from an antiparallel to a parallel state at time τ_s is

$$P_{T_s}[\tau_s] = 1 - \int_{-1}^{g_{L_i}(\tau_s)} \frac{1}{Z(\Delta_0, R)} \frac{\exp(-\Delta_0 x)}{\sqrt{-x}} dx$$

$$= \exp\left\{-\Delta_0 \left[1 + g_{L_i}(\tau_s)\right]\right\} \frac{F\left[\sqrt{-\Delta_0 g_{L_i}(\tau_s)}\right]}{F\left(\sqrt{\Delta_0}\right)}, \quad (18)$$

where $g_{L_i}(\tau_s)$ corresponds to the initial energy for a switching time τ_s. An exact analytical expression for $g_{L_i}(\tau_s)$ is not feasible; however, the monotonically decreasing nature of Eqs. (10)–(12) makes it possible to numerically invert them [32]. $P_{T_s}[\tau_s]$ includes all macrospins in an ensemble with $g_L \geq g_{L_i}(\tau_s)$, as we consider only deterministic switching in this paper. $P_{T_s}[\tau_s]$ also refers to the probability of switching of a macrospin with $\tau_s \geq T_s$; therefore, $P_{T_s}[\tau_s] = \Pr[T_s \leq \tau_s]$. Finally, the probability density function of the switching time is

$$\rho_{T_s}(\tau_s) = \frac{dP_{T_s}[\tau_s]}{d\tau_s}$$

$$= \frac{\sqrt{\Delta_0}}{2F(\sqrt{\Delta_0})} \frac{\exp\left\{-\Delta_0[1 + g_{L_i}(\tau_s)]\right\}}{\sqrt{-g_{L_i}(\tau_s)}} \left|\frac{dg_{L_i}(\tau_s)}{d\tau_s}\right|. \quad (19)$$

The write-error rate (WER) quantifies the probability of unsuccessful spin-torque switching of the magnet. Using Eq. (18), the WER is given by

$$1 - P_{T_s}[\tau]$$

$$= 1 - \exp\left\{-\Delta_0 \left[1 + g_{L_i}(\tau_s)\right]\right\} \frac{F\left[\sqrt{-\Delta_0 g_{L_i}(\tau_s)}\right]}{F\left(\sqrt{\Delta_0}\right)}. \quad (20)$$

C. Model Validation

To benchmark our analytical results, we solve Eq. (1) numerically using the Heun integration scheme, implemented in CUDA and run in parallel on GPUs. The numerical simulations are calibrated against published results to ensure their accuracy [13,24,33]. For all numerical simulations presented in this paper, the time step of the integration
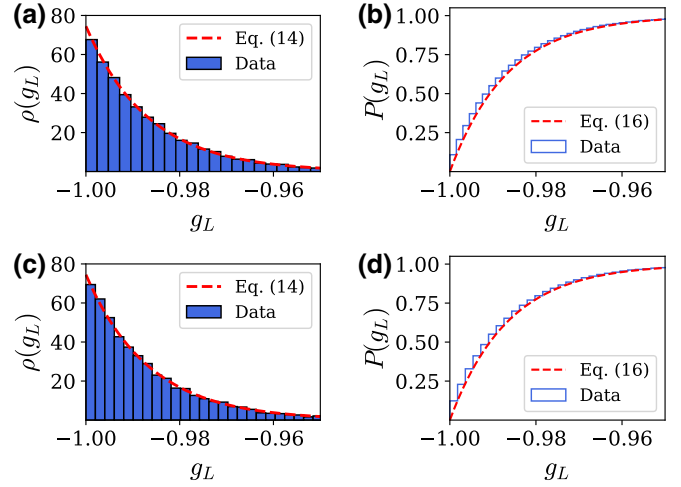

FIG. 4. Histogram distribution of g_L using $\alpha = 0.03$ and $\Delta_0 = 75$. The results in the top panel correspond to $R = 3.0$, while those in the bottom panel are for $R = 30.0$. (a),(c) PDF; (b),(d) CDF.

is set to 0.3 ps. The Gilbert damping constant of the free layer is taken as $\alpha = 0.03$, while its cross-section area is assumed to be $Ar = \pi/4 \times 30 \times 15$ nm^2. Different values of R, each signifying a different ferromagnetic material with its respective M_s and K_u values, are considered. The thickness of the free layer is adjusted to achieve an energy barrier $\Delta_0 = 75$ [34]. Simulation results corresponding to values of α other than 0.03 are not reported for brevity, as the key features and trends of the switching dynamics remain the same. Finally, a wide range of R values is selected to show the applicability and robustness of the model. Figure 4 shows that Eqs. (14) and (16) describe the equilibrium numerical distribution very well.

IV. RESULTS AND DISCUSSION

Figure 5 shows the average switching time $\langle t_s \rangle = \left[(1 + \alpha^2)/\gamma\mu_0\right] (\langle \tau_s \rangle/H_k)$ of an ensemble of 10^4 independent macrospins as a function of the input spin current density $J_s = \Delta_0(4ek_BT/\hbar)(I_s/Ar)$, where e is the electron charge and $\hbar$ is the reduced Planck constant. For $R = 0.001$ (uniaxial anisotropy), there is excellent agreement between the numerical data and the closed-form solutions given in Eqs. (10) and (12) for moderate to large current levels. However, for current levels approaching the threshold switching current, the numerical results predict a lower average switching time, as the presence of thermal noise aids the switching process.

For $R \gg 1$ (biaxial anisotropy), $\langle t_s \rangle$ obtained using Eq. (11) predicts a larger average switching time than the numerical results for current levels comparable to the threshold value. This is expected, since the analytical
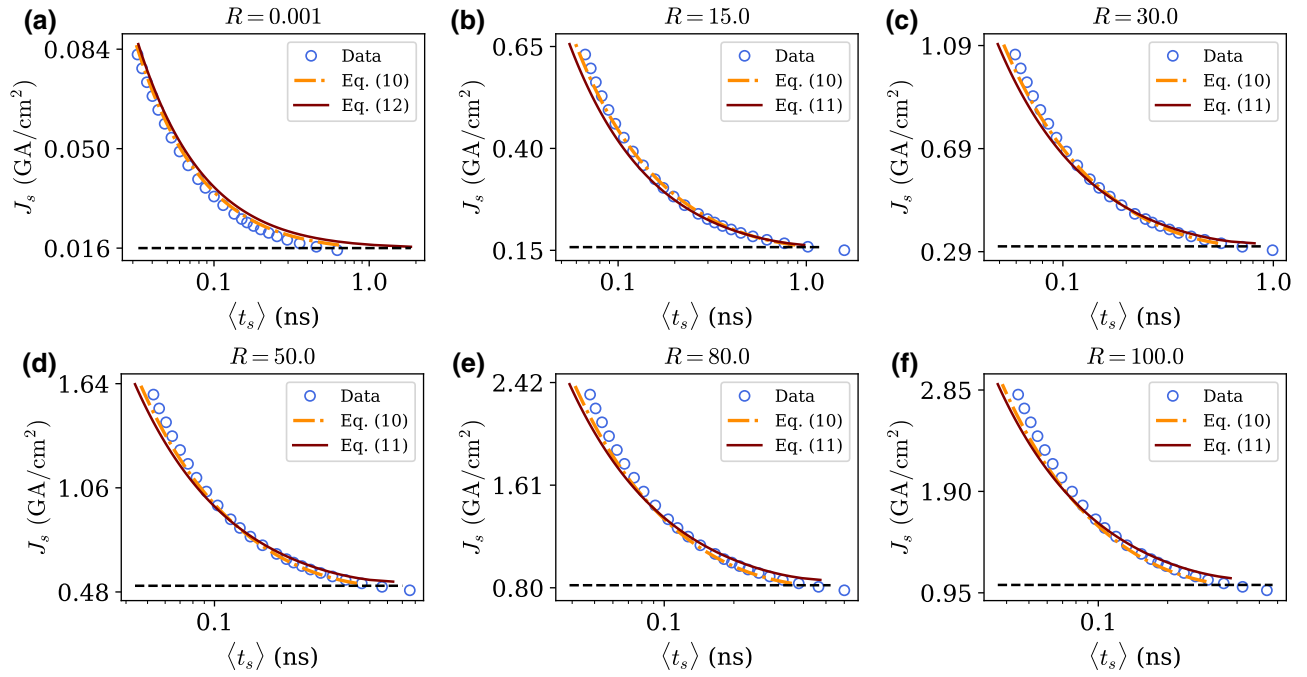
FIG. 5.   Average switching time $\langle t_s \rangle = [(1 + \alpha^2)/\gamma \mu_0](\langle \tau_s \rangle / H_k)$ as a function of injected spin current density J_s for different values of R. The horizontal dashed black line in each figure is the threshold current demarcating the region of deterministic switching from that of thermal activation. The numerical data is obtained for an ensemble of 10^4 macrospins.

solutions neglect the effect of thermal noise during the switching process. As the input spin current density (J_s) increases beyond the threshold value J_s^{th0} [$= \Delta_0(4ek_BT/\hbar)(I_s^{\text{th0}}/Ar)$], the agreement between the analytical and numerical results improves. Near the threshold current level, the average switching time obtained from Eq. (10) is slightly lower than that obtained from the numerical results. This slight deviation is due to the quadratic approximation for the elliptic integrals in Eq. (7). Note that for $J_s > J_s^{\text{thM}}$ [$= \Delta_0(4ek_BT/\hbar)(I_s^{\text{thM}}/Ar)$], the validity of CEOA is not fully justified, as the average rate of energy change becomes large [24]. However, the general trend of the switching time stays the same, as the error due to CEOA does not increase abruptly but rather increases slowly. We observe that the error between the switching times obtained from solution of Eqs. (1) and (10) is between 5 and 12% for different values of R, for currents as large as $2J_s^{\text{thM}}$.

The PDF and WER obtained analytically from Eqs. (19) and (20) are compared with numerical solutions in Figs. 6–9 for an ensemble of 10^6 macrospins. It is observed that in the uniaxial case, the accuracy of the PDF and WER improves as the applied spin current increases [33]. For $R \geq 15$, the accuracy of the analytical solutions also increases as the current increases from J_s^{th0} toward J_s^{thM}. Although for spin currents larger than J_s^{thM} the accuracy of the analytical results reduces as the validity of CEOA becomes questionable, the model is fairly robust and predicts the numerical results well. To arrive at the analytical

results reported in Figs. 7–9, we numerically invert Eq. (11), due to its simplicity.

Figure 10 shows the average switching time for a few ferromagnetic thin films. Both the numerical and the analytical average switching times [Eqs. (1) and (10)] are for the material parameters listed in Table III. The larger value
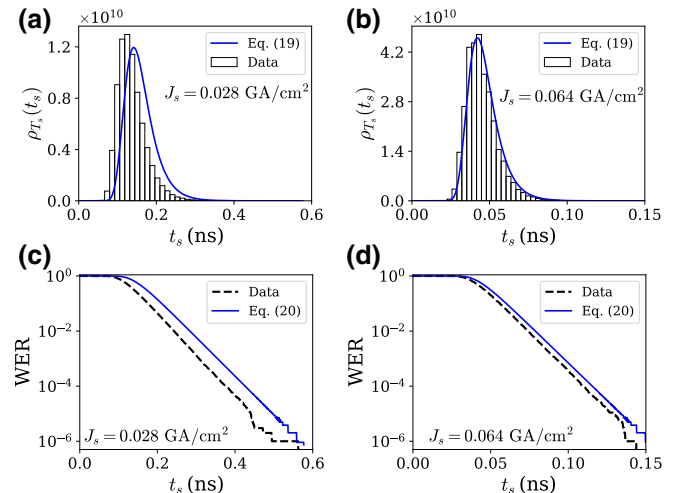

FIG. 6.   Results for $R = 0.001$ (uniaxial anisotropy) and $J_s^{\text{th1}} = 1.6 \times 10^7$ A/cm^2. The top panel shows the PDF, while the bottom panel shows the WER. The accuracy of the analytical results improves as the input spin-current density increases with respect to J_s^{th1}.
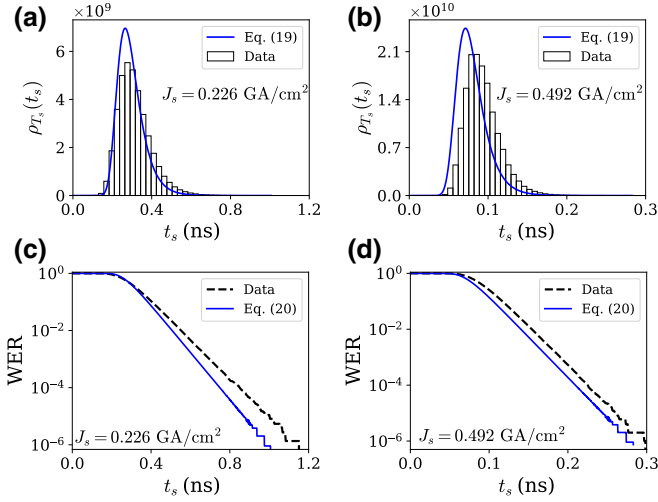
FIG. 7. Results for $R = 15$ and $J_s^{\text{th0}} = 1.58 \times 10^8$ A/cm². The top panel shows the PDF, while the bottom panel shows the WER. The accuracy of the analytical solutions improves for spin-current densities larger than the threshold current density. Although CEOA is strictly valid for $J_s < J_s^{\text{thM}} = 2.1 J_s^{\text{th0}}$, the analytical model agrees well with the numerical solution for currents larger than J_s^{thM}.

of the spin current is chosen to demonstrate the applicability of the model beyond the strict validity of CEOA. As expected, the average switching time decreases for the larger value of the spin current. Also, for two materials with similar values of R, the switching time is lower for the material with a higher damping constant.
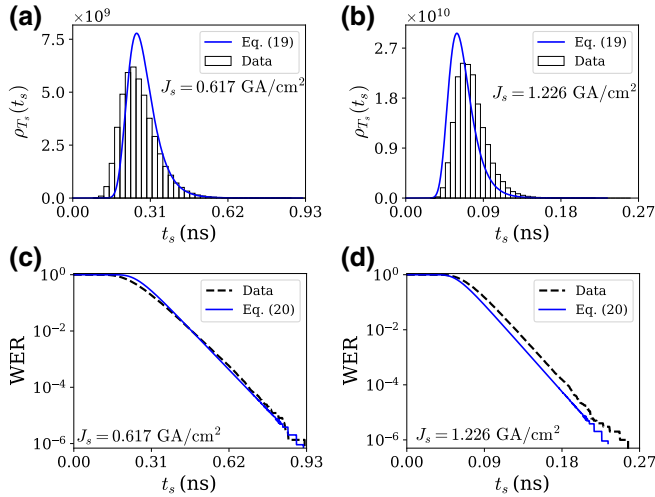

FIG. 8. Results for $R = 50$ and $J_s^{\text{th0}} = 5.14 \times 10^8$ A/cm². The top panel shows the PDF, while the bottom panel shows the WER. The accuracy of the analytical results improves for spin-current densities larger than the threshold current density. The analytical model is fairly accurate even for $J_s > J_s^{\text{thM}} = 1.6 J_s^{\text{th0}}$, showing the robustness of our model.
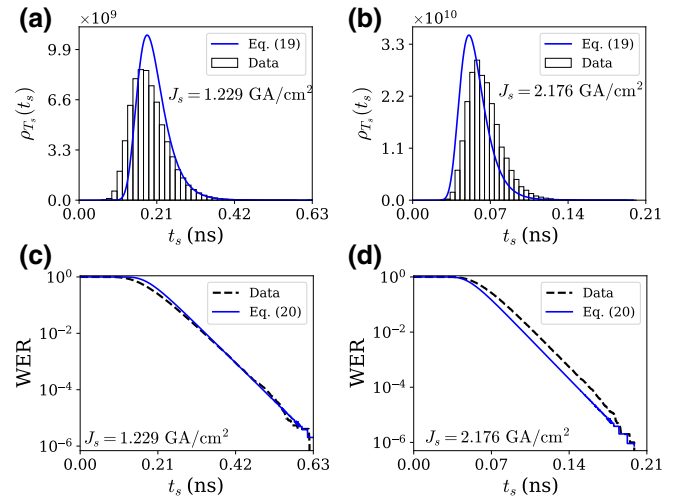

FIG. 9. Results for $R = 100$ and $J_s^{\text{th0}} = 1.02 \times 10^9$ A/cm². The top panel shows the PDF, while the bottom panel shows the WER. The accuracy of the analytical results improves as J_s increases above J_s^{th0}. For $J_s > J_s^{\text{thM}} = 1.42 J_s^{\text{th0}}$, the accuracy of the analytical model reduces but continues to predict the numerical results with good accuracy.

A. Applications of spin-torque memory in the deterministic and thermally activated regimes

Conventional artificial-intelligence hardware accelerators based on von Neumann architectures suffer from significant power dissipation and communication overheads [40–42]. The introduction of nonvolatile magnetic memory into hybrid magnetic-silicon hardware could lead to lower power dissipation even for computationally expensive machine-learning tasks that routinely process large-scale datasets [43]. Neuromorphic architectures that colocate memory and computing elements are well suited to leveraging the nonvolatile spin-torque memory discussed in this paper. Such memory can be used as cache memory in general-purpose processors and for storing the weights pertinent to machine-learning tasks in a neuromorphic architecture [44].

Beyond serving as nonvolatile memory elements, the device architectures illustrated in Fig. 1 can also function
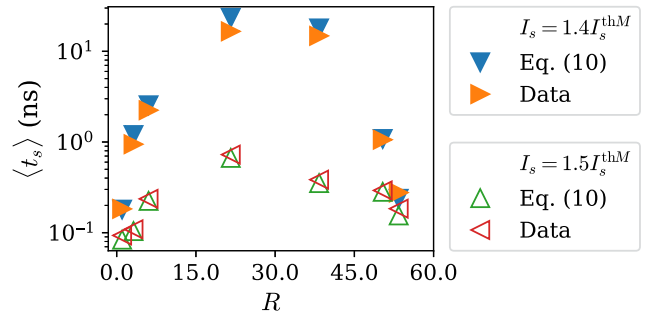

FIG. 10. Average switching time for materials listed in Table III for two different currents, viz., $I_s = 1.4 I_s^{\text{thm}}$ and $I_s = 1.5 I_s^{\text{thM}}$.

TABLE III. List of common ferromagnetic materials, and their saturation magnetization M_s, uniaxial-anisotropy energy density K_u, and damping constant α.

Material	M_s (T)	K_u (MJ/m^3)	R	α	Ref.
Terfenol-D	1.0	0.39	1.04	0.1	[28]
Co	1.81	0.41	3.2	0.02	[35]
$Co_{0.6}Fe_{0.2}B_{0.2}$	1.2	0.095	6.0	0.015	[36,37]
NiMnSb	0.84	0.013	21.6	0.002	[35,38]
Fe	2.15	0.048	38.3	0.001	[35,39]
EuO	2.36	0.044	50.4	0.015	[35]
Fe-Ga-B	1.63	0.0198	53.4	0.1	[28]

as stochastic oscillators when operated in the thermally activated regime, for which $I_s < I_s^{\text{th}m}$. In this regime, the input spin current controls the rate at which the ferromagnetic thin film fluctuates between its two stable states [45]. Several concepts of neuromorphic computing have been proposed in the literature based on fluctuating magnets [44,46–48]. Moreover, if the constraint on the energy barrier is relaxed, the thresholding behavior of monodomain ferromagnets can mimic neural dynamics in an energy-efficient manner. In the case of multidomain effects, it has been shown that the structures shown in Fig. 1 can implement synaptic functionality [46,49], which is necessary for weight storage and update in neuromorphic hardware. Energy-efficient probabilistic computing for image processing and machine learning [50], and spin logic [51,52] are a few other possible applications in the thermally activated regime.

The thermally activated regime is also important from the point of view of studying the fundamental scaling laws of the spin-torque-driven dynamics of ferromagnetic materials. The scaling of the energy-barrier height of ferromagnets with the spin current is typically modeled as $\Delta = \Delta_0 \left(1 - I_s/I_s^{\text{th}1}\right)^{\xi}$. The value of ξ in the limit of $I_s \ll I_s^{\text{th}1}$ has been a subject of debate for over two decades [21,53–56]. Recently, it was theoretically postulated [54,55] that $\xi = 2$ in the case of uniaxial anisotropy ($R \to 0$). For the biaxial case ($R \gg 1$), however, the value of ξ was estimated to be 2.2 by Taniguchi and coworkers based on the Fokker-Planck (FP) equation [56]. Pinna and coworkers obtained the current dependence of ξ via applying the principle of action [24,55] to Eq. (7). In our opinion, it might be possible to construct a FP representation from Eq. (7) and solve it with the rational approximations in Eqs. (8b)–(8c) to arrive at an analytical expression for ξ. However, a full analysis in the thermally activated regime is outside the scope of this paper.

V. CONCLUSION

The analytical models for the average switching time, the probability distribution function of the switching times,

and the write-error rate developed in this paper for thin-film magnets with biaxial anisotropy show good agreement with numerical results for moderate to large spin-current densities. In the vicinity of the threshold spin-current density, the error between the analytical and numerical data is significant due to thermal noise. For very large spin-current densities, the constant-energy-orbit-averaging approach adopted in this paper becomes inadequate, even though the error between the numerical and analytical results is well under the tolerance limit. The models in this paper should complement experimental results and aid the analysis, design, and development of nonvolatile memory driven by both spin-transfer and spin-orbit torques.

APPENDIX A: SIMPLIFYING Eq. (6)

Without any loss of generality, we assume the easy and hard axes to coincide with $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$, respectively. Next, we assume the spin polarization of the fixed layer to be in the plane of the magnet at an angle ϕ from the easy axis, and therefore $\hat{\mathbf{n}}_p = \cos\phi\,\hat{\mathbf{x}} + \sin\phi\,\hat{\mathbf{y}}$. Therefore, for zero external magnetic field ($\mathbf{H}_a = 0$), the effective magnetic field is $\mathbf{h}_{\text{eff}} = m_x\hat{\mathbf{x}} - Rm_z\hat{\mathbf{z}} + \mathbf{h}_T$, where $\mathbf{h}_T$ is the thermal field. Therefore, the three components of Eq. (1) are

$$\frac{\partial m_x}{\partial \tau} = Rm_y m_z + \alpha m_x \left(1 - m_x^2 + Rm_z^2\right)$$
$$+ I_s \left[\cos\phi - m_x \left(\cos\phi\, m_x + \sin\phi\, m_y\right)\right]$$
$$- \alpha I_s m_z \sin\phi$$
$$+ n_{S,x}, \tag{A1a}$$

$$\frac{\partial m_y}{\partial \tau} = -(R+1)\, m_x m_z - \alpha m_y \left(m_x^2 - Rm_z^2\right)$$
$$+ I_s \left[\sin\phi - m_y \left(\cos\phi\, m_x + \sin\phi\, m_y\right)\right]$$
$$+ \alpha I_s m_z \cos\phi$$
$$+ n_{S,y}, \tag{A1b}$$

and

$$\frac{\partial m_z}{\partial \tau} = m_x m_y - \alpha m_z \left(R + m_x^2 - Rm_z^2\right)$$
$$- I_s m_z \left(\cos\phi\, m_x + \sin\phi\, m_y\right)$$
$$+ \alpha I_s \left(\sin\phi\, m_x - \cos\phi\, m_y\right)$$
$$+ n_{S,z}, \tag{A1c}$$

where each of $n_{S,p}$ denotes a thermal-noise component in the Stratonovich sense. From Eqs. (4) and (5), we have

$$\begin{pmatrix} n_{S,x} \\ n_{S,y} \\ n_{S,z} \end{pmatrix} = \mathcal{D} \circ \begin{pmatrix} h_{T,x} \\ h_{T,y} \\ h_{T,z} \end{pmatrix} = \sqrt{D}\mathcal{D} \circ \begin{pmatrix} \dot{W}_x \\ \dot{W}_y \\ \dot{W}_z \end{pmatrix}, \qquad \text{(A2)}$$

where $\dot{\mathbf{W}}$ represents a 3D stochastic Wiener process, each of whose components is a Gaussian random variable with zero mean and unit standard deviation. $\mathcal{D}$ is referred to as the diffusion matrix and is given by

$$\begin{pmatrix} \alpha\left(1 - m_x^2\right) & m_z - \alpha m_x m_y & -m_y - \alpha m_x m_z \\ -m_z - \alpha m_x m_y & \alpha\left(1 - m_y^2\right) & m_x - \alpha m_y m_z \\ m_y - \alpha m_x m_z & -m_x - \alpha m_y m_z & \alpha\left(1 - m_z^2\right) \end{pmatrix}.$$

Substituting Eqs. (A1a) and (A1c) into Eq. (6) leads to

$$\begin{aligned} \frac{\partial g_L}{\partial \tau} &= 2Rm_z \left[m_x m_y - \alpha m_z \left(R - g_L\right) \right. \\ &\quad - I_s m_z \left(\cos\phi\, m_x + \sin\phi\, m_y\right) \\ &\quad + \alpha I_s \left(\sin\phi\, m_x - \cos\phi\, m_y\right) \right] \\ &\quad - 2m_x \left\{ Rm_y m_z + \alpha m_x \left(1 + g_L\right) \right. \\ &\quad + I_s \left[\cos\phi - m_x \left(\cos\phi\, m_x + \sin\phi\, m_y\right)\right] \\ &\quad - \alpha I_s \sin\phi\, m_y \right\} \\ &\quad + 2\sqrt{D} \begin{pmatrix} d_x & d_y & d_z \end{pmatrix} \circ \begin{pmatrix} \dot{W}_x \\ \dot{W}_y \\ \dot{W}_z \end{pmatrix} \qquad \text{(A3a)} \end{aligned}$$

$$\text{and} \quad \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} = \begin{pmatrix} Rm_z\mathcal{D}_{31} - m_x\mathcal{D}_{11} \\ Rm_z\mathcal{D}_{32} - m_x\mathcal{D}_{12} \\ Rm_z\mathcal{D}_{33} - m_x\mathcal{D}_{13} \end{pmatrix}. \qquad \text{(A3b)}$$

It can be observed from Eq. (A3a) that each of the damping, spin current, and thermal noise contributes to the rate of energy change. In addition, the last term in Eq. (A3a) can be simplified to $\sqrt{d_x^2 + d_y^2 + d_z^2}\sqrt{\dot{W}_x^2 + \dot{W}_y^2 + \dot{W}_Z^2}$, which leads to

$$\begin{aligned} \frac{\partial g_L}{\partial \tau} &= -2\alpha \left\{ (1 + R)\, m_x^2 + g_L \left(R - g_L\right) \right. \\ &\quad + \frac{I_s}{\alpha} \left[(1 + g_L)\cos\phi\, m_x + g_L \sin\phi\, m_y\right] \\ &\quad + I_s \left[\sin\phi\, m_x m_y - R\left(\sin\phi\, m_x m_z - \cos\phi\, m_y m_z\right)\right] \right\} \\ &\quad + 2\sqrt{\frac{\alpha}{\Delta_0}}\sqrt{(1 + R)\, m_x^2 + g_L \left(R - g_L\right)} \circ \dot{W}_{g_L}, \end{aligned}$$
$$\text{(A4)}$$

where $\dot{W}_{g_L}$ is 1D white Gaussian noise, and it acts away from the constant-energy orbit along a normal [26].

We now average $\partial g_L/\partial \tau$ over a constant-energy orbit, as the nonconservative effects act over long timescales, whereas we are more interested in studying their effects on the rapid periodic motion. In the limit of zero damping, with zero spin current at absolute zero temperature, the constant-energy solutions to Eqs. (A1a)–(A1c) in the antiparallel well are given by [23,24,26]

$$m_x^c(t) = -\sqrt{\frac{R - g_L}{R + 1}}\, \mathrm{dn}\left[\sqrt{R - g_L}\, t, k^2\right], \qquad \text{(A5a)}$$

$$m_y^c(t) = \sqrt{1 + g_L}\, \mathrm{sn}\left[\sqrt{R - g_L}\, t, k^2\right], \qquad \text{(A5b)}$$

$$m_z^c(t) = \sqrt{\frac{1 + g_L}{R + 1}}\, \mathrm{cn}\left[\sqrt{R - g_L}\, t, k^2\right], \qquad \text{(A5c)}$$

where $k^2 \equiv R[(1 + g_L)/(R - g_L)]$, $0 < k^2 < 1$, and sn[·], cn[·], and dn[·] are Jacobi elliptic functions. The magnetization oscillates around the easy axis with a time period $T(g_L) = (4/\sqrt{R - g_L})K\left(k^2\right)$, where $K\left(k^2\right)$ is the complete elliptic integral of the first kind. In Eq. (A4), only the time averages of m_x and m_x^2 are nonzero, whereas those of m_y, $m_x m_y$, $m_x m_z$, and $m_y m_z$ are zero due to their periodic nature. To evaluate these averages, we geometrically parameterize [24,26] the constant-energy orbit in the antiparallel well as a function of the parameter ω as

$$\cosh^2(\omega) - \sinh^2(\omega) = 1, \qquad \text{(A6a)}$$

$$\frac{1}{-g_L}m_x^2 - \frac{R}{-g_L}m_z^2 = 1, \qquad \text{(A6b)}$$

which leads to the form

$$m_x^c = -\sqrt{-g_L}\cosh(\omega), \qquad \text{(A7a)}$$

$$m_z^c = \pm\sqrt{\frac{-g_L}{R}}\sinh(\omega), \qquad \text{(A7b)}$$

$$m_y^c = \pm\sqrt{\frac{R - g_L}{R}}\sqrt{1 - \zeta^2 \cosh^2(\omega)}, \qquad \text{(A7c)}$$

$$\zeta^2 = -\frac{g_L(R + 1)}{R - g_L} = 1 - k^2, \qquad \text{(A7d)}$$

where $|\omega| < \cosh^{-1}(1/\zeta)$. We define the time average of a function $p(t)$ as

$$\langle p(t)\rangle = \frac{4}{T(g_L)}\int_0^{t=m_y:\sqrt{1+g_L}\to 0} p(t)\, dt, \qquad \text{(A8)}$$

and the time averages of $\langle m_x \rangle$ and $\langle m_x^2 \rangle$ are evaluated as

$$
\begin{aligned}
\langle m_x \rangle &= \frac{4}{T(g_L)} \int_0^{\cosh^{-1}(1/\zeta)} m_x(\omega) \left| \frac{\sqrt{R} m_z}{R m_y m_z} \right| d\omega \\
&= -\frac{4}{T(g_L)} \frac{\sqrt{-g_L}}{\sqrt{R - g_L}} \int_0^{\cosh^{-1}(1/\zeta)} \frac{\cosh(\omega)\, d\omega}{\sqrt{1 - \zeta^2 \cosh^2(\omega)}} \\
&= -\frac{4}{T(g_L)} \frac{\sqrt{-g_L}}{\zeta \sqrt{R - g_L}} \int_0^{\cosh^{-1}(1/\zeta)} \\
&\quad \times \frac{\cosh(\omega)\, d\omega}{\sqrt{(1 - \zeta^2)/\zeta^2 - \sinh^2(\omega)}} \\
&= -\frac{\pi}{2} \frac{\sqrt{R - g_L}}{\sqrt{1 + R}} \frac{1}{K(\zeta)} = -\frac{\pi}{2} \sqrt{\frac{R - g_L}{R + 1}} \frac{1}{K(R, g_L)},
\end{aligned}
\tag{A9}
$$

and

$$
\begin{aligned}
\langle m_x^2 \rangle &= \frac{4}{T(g_L)} \int_0^{\cosh^{-1}(1/\zeta)} m_x^2(\omega) \left| \frac{m_x \sqrt{R} m_z}{m_x R m_y m_z} \right| d\omega \\
&= \frac{-4}{T(g_L)} \frac{g_L}{\sqrt{R - g_L}} \int_0^{\cosh^{-1}(1/\zeta)} \frac{\cosh^2(\omega)\, dw}{\sqrt{1 - \zeta^2 \cosh^2(\omega)}} \\
&= \frac{-1}{K(\zeta)} \frac{g_L}{\sqrt{1 - \zeta^2}} \int_0^{\cosh^{-1}(1/\zeta)} \\
&\quad \times \frac{\cosh^2(\omega)\, d\omega}{\sqrt{1 - [\zeta^2/(1 - \zeta^2)] \sinh^2(\omega)}} \\
&= \frac{-1}{K(\zeta)} \frac{g_L}{\zeta} \int_0^1 \frac{\sqrt{1 - (1 - 1/\zeta^2) u^2}\, du}{\sqrt{1 - u^2}} \\
&= \frac{-1}{K(\zeta)} \frac{g_L}{\zeta^2} E(1 - \zeta^2) = \frac{R - g_L}{R + 1} \frac{E(R, g_L)}{K(R, g_L)},
\end{aligned}
\tag{A10}
$$

where we use $dt = \{[dp(\omega)/d\omega]/[dp(t)/dt]\}d\omega$. Substituting $\langle m_x \rangle$ and $\langle m_x^2 \rangle$ in Eq. (A4) results in the average rate of energy flow,

$$
\begin{aligned}
\left\langle \frac{\partial g_L}{\partial \tau} \right\rangle &= -\frac{\pi \alpha}{K(R, g_L)} \sqrt{\frac{R - g_L}{1 + R}} \\
&\quad \times \left\{ -\frac{I_s}{\alpha} \cos\phi\, (1 + g_L) + \frac{2}{\pi} \sqrt{1 + R} \sqrt{R - g_L} \right. \\
&\quad \left. \times [E(R, g_L) + g_L K(R, g_L)] \right\} - 2\sqrt{\frac{\alpha}{\Delta_0}} \\
&\quad \times \sqrt{\frac{(R - g_L)}{K(R, g_L)}} \sqrt{E(R, g_L) + g_L K(R, g_L)} \circ \dot{W}_{g_L}.
\end{aligned}
\tag{A11}
$$

In this paper, we assume the spin polarization $\hat{\mathbf{n}}_p$ to be collinear with the easy axis $\hat{\mathbf{n}}_e$, and so $\phi = 0$, which leads to Eq. (7).

APPENDIX B: SOLVING THE INTEGRAL IN Eq. (9)

In Eq. (9), we substitute $R - g_L = u^2$, so that we have

$$
\begin{aligned}
\tau_s &= \frac{1}{2A\alpha} \frac{R + 4}{R + 2} \int_{\sqrt{R - g_{L_i}}}^{\sqrt{R - g_{L_f}}} \frac{u^2 - [R(1 + R)/(R + 4)]}{u^2 - [R(1 + R)/(R + 2)]} \\
&\quad \times \left[\frac{du}{P(u)} \right],
\end{aligned}
\tag{B1}
$$

where $P(u) = u^5 - (B/A + 2R)u^3 + (\tilde{I}_s/A\sqrt{1 + R})u^2 + [R^2 + (B/A)R + C/A]u - \tilde{I}_s\sqrt{1 + R}/A$ and the λ_i are the roots of the polynomial $P(u)$, which are evaluated numerically. Using partial fractions to resolve the denominator of the integral, we finally have

$$
\begin{aligned}
\tau_s &= \frac{1}{2\alpha(R + 2)} \left\{ \sum_{i=1}^{5} \frac{N}{D} \log\left[\frac{(R - g_{L_f}) - \lambda_i}{(R - g_{L_i}) - \lambda_i} \right] \right. \\
&\quad + \sqrt{\frac{R(1 + R)}{R + 2}} \left(\frac{\log\left[\frac{\sqrt{(R - g_{L_f})(R + 2)} - \sqrt{R(1 + R)}}{\sqrt{(R - g_{L_i})(R + 2)} - \sqrt{R(1 + R)}} \right]}{A \prod_{n=1}^{5} \left[\frac{R(1 + R)}{R + 2} - \lambda_i \right]} \right. \\
&\quad \left. \left. + \frac{\log\left[\frac{\sqrt{(R - g_{L_f})(R + 2)} + \sqrt{R(1 + R)}}{\sqrt{(R - g_{L_i})(R + 2)} + \sqrt{R(1 + R)}} \right]}{A \prod_{n=1}^{5} \left[\frac{R(1 + R)}{R + 2} + \lambda_i \right]} \right) \right\}.
\end{aligned}
\tag{B2}
$$

APPENDIX C: $R \to \infty$

For large values of R and for Eq. (7) in the deterministic regime, along with rational approximations for the elliptic integrals, Eqs. (8b) and (8c) can be simplified to

$$
\begin{aligned}
\frac{\partial g_L}{\partial \tau} &= 4\alpha \left[\frac{1 - g_L}{3 - g_L} \right] \left\{ \tilde{I}_s(1 + g_L) - R\left[g_L \frac{3 - g_L}{2 - 2g_L} \right. \right. \\
&\quad \left. \left. + 1 - \frac{(1 + g_L)}{4} \frac{(1 + g_L)^2 - 28(1 + g_L) + 64}{4(1 + g_L)^2 - 40(1 + g_L) + 64} \right] \right\}.
\end{aligned}
\tag{C1}
$$

Substituting $1 + g_L = x$ in the above equation leads us to

$$
\frac{dx}{d\tau} = 4\alpha \left[\frac{x-2}{x-4}\right] \left\{ \tilde{I}_s x - R \left[(x-1)\frac{x-4}{2x-4}\right.\right.
$$
$$
\left.\left. +1 - \frac{x}{4}\frac{x^2 - 28x + 64}{4x^2 - 40x + 64}\right]\right\}, \tag{C2}
$$

which can then be simplified and rearranged into an integral of the form

$$
\int_{1+g_{L_i}}^{1+g_{L_f}} \frac{(x-4)(x-8)\,dx}{x\left[x^2 - Ex + F\right]} = \frac{\alpha}{4}\left(16\tilde{I}_s - 7R\right)\tau_s, \tag{C3}
$$

where $E = (160\tilde{I}_s - 60R)/(16\tilde{I}_s - 7R)$ and $F = (256\tilde{I}_s - 128R)/(16\tilde{I}_s - 7R)$. If the roots of the quadratic equation $x^2 - Ex + F = 0$ are a and b, then the switching time τ_s can be evaluated as

$$
\tau_s = \frac{1}{2\alpha\left[\tilde{I}_s - 0.5R\right]}\left\{ \log\left[\frac{1+g_{L_f}}{1+g_{L_i}}\right]\right.
$$
$$
+ \frac{b(a-4)(a-8)}{32(a-b)}\log\left[\frac{1+g_{L_f} - a}{1+g_{L_i} - a}\right]
$$
$$
\left. - \frac{a(b-4)(b-8)}{32(a-b)}\log\left[\frac{1+g_{L_f} - b}{1+g_{L_i} - b}\right]\right\}. \tag{C4}
$$

APPENDIX D: UNIAXIAL LIMIT

In the case of the uniaxial limit, Eq. (7) in the deterministic regime reduces to

$$
\left\langle\frac{\partial g_L}{\partial\tau}\right\rangle = 2\alpha\sqrt{-g_L}\left(1 + g_L\right)\left(\tilde{I}_s - \sqrt{-g_L}\right), \tag{D1}
$$

which can then be integrated as

$$
\int_{g_{L_i}}^{g_{L_f}} \frac{dg_L}{\sqrt{-g_L}\left(1 + g_L\right)\left(\tilde{I}_s - \sqrt{-g_L}\right)} = 2\alpha\tau_s. \tag{D2}
$$

Substituting $g_L = -u^2$, using partial fractions to separate the terms in the denominator, and integrating with proper limits leads us to

$$
\tau_s = \frac{1}{2\alpha\left(\tilde{I}_s^2 - 1\right)}\left(\tilde{I}_s\left\{\log\left[\frac{1+\sqrt{-g_{L_i}}}{1+\sqrt{-g_{L_f}}}\right]\right.\right.
$$
$$
\left.- \log\left[\frac{1-\sqrt{-g_{L_i}}}{1-\sqrt{-g_{L_f}}}\right]\right\} - \log\left[\frac{1+g_{L_i}}{1+g_{L_f}}\right]
$$
$$
\left.+ 2\log\left[\frac{\tilde{I}_s/\alpha - \sqrt{-g_{L_i}}}{\tilde{I}_s/\alpha - \sqrt{-g_{L_f}}}\right]\right). \tag{D3}
$$

APPENDIX E: Eqs. (14) and (16)

If a random variable X has a certain probability $P_X[x]$, then the probability for a random variable $Y = -X^2$ can be obtained as

$$
P_Y[y] = P\left[Y \leq y\right] = P\left[-X^2 \leq y\right]
$$
$$
= 1 - P\left[|X| \leq \sqrt{-y}\right]
$$
$$
= 1 - P\left[-\sqrt{-y} \leq X \leq \sqrt{-y}\right]. \tag{E1}
$$

Now, if we assume that X represents the distribution of the magnetization m_x while Y represents that of the energy g_L, then one can obtain the following probability for $R \to 0$ using Eq. (13):

$$
P[g_L] = 1 - \frac{2}{Z(\Delta_0, 0)}\int_0^{\sqrt{-g_L}} dx\,\exp\left(\Delta_0 m_x^2\right)
$$
$$
= 1 - \frac{2}{Z(\Delta_0, 0)\sqrt{\Delta_0}}\int_0^{\sqrt{-g_L}} dm_x\,\sqrt{\Delta_0}\exp\left(\Delta_0 m_x^2\right)
$$
$$
= 1 - \frac{2\exp(\Delta_0 g_L)}{Z(\Delta_0, 0)\sqrt{\Delta_0}\exp(\Delta_0 g_L)}\int_0^{\sqrt{-\Delta_0 g_L}} du\,\exp\left(u^2\right)
$$
$$
= 1 - \frac{2\exp(-\Delta_0 g_L)}{Z(\Delta_0, 0)\sqrt{\Delta_0}}F\left(\sqrt{-\Delta_0 g_L}\right). \tag{E2}
$$

Here we substitute $\sqrt{\Delta_0}m_x = u$ and use $F(x) = \exp(-x^2)\int_0^x dy\,\exp(y^2)$, the Dawson integral. Since $P[-1] = 0$, therefore,

$$
Z(\Delta_0, 0) = \frac{2\exp(\Delta_0)}{\sqrt{\Delta_0}}F\left(\sqrt{\Delta_0}\right). \tag{E3}
$$

Equations (E2) and (E3) together lead to Eq. (16).

In order to arrive at the PDF, we now differentiate Eq. (E2) with respect to g_L to get

$$
\rho_Y[g_L] = \frac{1}{Z(\Delta_0, 0)}\frac{\exp\left(-\Delta_0 g_L\right)}{\sqrt{-g_L}}
$$
$$
= \frac{\sqrt{\Delta_0}}{2F\left(\sqrt{\Delta_0}\right)}\frac{\exp\left[-\Delta_0\left(1 + g_L\right)\right]}{\sqrt{-g_L}}, \tag{E4}
$$

where we use $Z(\Delta_0, R) = Z(\Delta_0, 0)$.

---

[1] A. D. Kent and D. C. Worledge, A new spin on magnetic memories, Nat. Nanotechnol. **10**, 187 (2015).

[2] A. V. Sadovnikov, S. A. Odintsov, E. N. Beginin, S. E. Sheshukova, Y. P. Sharaevskii, and S. A. Nikitov, Toward nonlinear magnonics: Intensity-dependent spin-wave switching in insulating side-coupled magnetic stripes, Phys. Rev. B **96**, 144428 (2017).

[3] T. Chen, R. K. Dumas, A. Eklund, P. K. Muduli, A. Houshang, A. A. Awad, P. Dürrenfeld, B. G. Malm, A. Rusu, and J. Åkerman, Spin-torque and spin-Hall nano-oscillators, Proc. IEEE **104,** 1919 (2016).

[4] J. C. Slonczewski, Current-driven excitation of magnetic multilayers, J. Magn. Magn. Mater. **159,** L1 (1996).

[5] L. Berger, Emission of spin wave s by a magnetic multilayer traversed by a current, Phys. Rev. B **54,** 9353 (1996).

[6] J. Sinova, S. O. Valenzuela, J. Wunderlich, C. H. Back, and T. Jungwirth, Spin Hall effects, Rev. Mod. Phys. **87,** 1213 (2015).

[7] L. Liu, C.-F. Pai, Y. Li, H. W. Tseng, D. C. Ralph, and R. A. Buhrman, Spin-torque switching with the giant spin Hall effect of tantalum, Science **336,** 555 (2012).

[8] D. Sander, The magnetic anisotropy and spin reorientation of nanostructures and nanoscale films, J. Phys.: Condens. Matter **16,** R603 (2004).

[9] L. Liu, O. J. Lee, T. J. Gudmundsen, D. C. Ralph, and R. A. Buhrman, Current-Induced Switching of Perpendicularly Magnetized Magnetic Layers Using Spin Torque from the Spin Hall Effect, Phys. Rev. Lett. **109,** 096602 (2012).

[10] Y.-W. Oh, S.-H. C. Baek, Y. Kim, H. Y. Lee, K.-D. Lee, C.-G. Yang, E.-S. Park, K.-S. Lee, K.-W. Kim, G. Go, J.-R. Jeong, B.-C. Min, H.-W. Lee, K.-J. Lee, and B.-G. Park, Field-free switching of perpendicular magnetization through spin–orbit torque in antiferromagnet/ferromagnet/oxide structures, Nat. Nanotechnol. **11,** 878 (2016).

[11] S. Fukami, T. Anekawa, C. Zhang, and H. Ohno, A spin–orbit torque switching scheme with collinear magnetic easy axis and current configuration, Nat. Nanotechnol. **11,** 621 (2016).

[12] J. Z. Sun, Spin-current interaction with a monodomain magnetic body: A model study, Phys. Rev. B **62,** 570 (2000).

[13] D. Pinna, A. D. Kent, and D. L. Stein, Spin-transfer torque magnetization reversal in uniaxial nanomagnets with thermal noise, J. Appl. Phys. **114,** 033901 (2013).

[14] M. d'Aquino, V. Scalera, and C. Serpico, Analysis of switching times statistical distributions for perpendicular magnetic memories, J. Magn. Magn. Mater. **475,** 652 (2019).

[15] W. F. Brown, Jr., Thermal fluctuations of a single-domain particle, Phys. Rev. **130,** 1677 (1963).

[16] J. Park, G. Rowlands, O. Lee, D. Ralph, and R. Buhrman, Macrospin modeling of sub-ns pulse switching of perpendicularly magnetized free layer via spin-orbit torques for cryogenic memory applications, Appl. Phys. Lett. **105,** 102404 (2014).

[17] R. Moreno, R. F. L. Evans, S. Khmelevskyi, M. C. Muñoz, R. W. Chantrell, and O. Chubykalo-Fesenko, Temperature-dependent exchange stiffness and domain wall width in Co, Phys. Rev. B **94,** 104433 (2016).

[18] D. Bedau, H. Liu, J. Sun, J. Katine, E. Fullerton, S. Mangin, and A. Kent, Spin-transfer pulse switching: From the dynamic to the thermally activated regime, Appl. Phys. Lett. **97,** 262502 (2010).

[19] M. Baumgartner, K. Garello, J. Mendil, C. O. Avci, E. Grimaldi, C. Murer, J. Feng, M. Gabureac, C. Stamm, Y. Acremann, S. Finizio, S. Wintz, J. Raabe, and P. Gambardella, Spatially and time-resolved magnetization dynamics driven by spin–orbit torques, Nat. Nanotechnol. **12,** 980 (2017).

[20] E. Grimaldi, V. Krizakova, G. Sala, F. Yasin, S. Couet, G. S. Kar, K. Garello, and P. Gambardella, Single-shot dynamics of spin–orbit torque and spin transfer torque switching in three-terminal magnetic tunnel junctions, Nat. Nanotechnol. **15,** 111 (2020).

[21] D. M. Apalkov and P. B. Visscher, Spin-torque switching: Fokker-Planck rate calculation, Phys. Rev. B **72,** 180405(R) (2005).

[22] T. Taniguchi, Y. Utsumi, M. Marthaler, D. S. Golubev, and H. Imamura, Spin torque switching of an in-plane magnetized system in a thermally activated region, Phys. Rev. B **87,** 054406 (2013).

[23] I. D. Mayergoyz, G. Bertotti, and C. Serpico, *Nonlinear Magnetization Dynamics in Nanosystems* (Elsevier, 2009).

[24] D. Pinna, A. D. Kent, and D. L. Stein, Thermally assisted spin-transfer torque dynamics in energy space, Phys. Rev. B **88,** 104405 (2013).

[25] K. A. Newhall and E. Vanden-Eijnden, Averaged equation for energy diffusion on a graph reveals bifurcation diagram and thermally assisted reversal times in spin-torque driven nanomagnets, J. Appl. Phys. **113,** 184105 (2013).

[26] D. Pinna, D. L. Stein, and A. D. Kent, Spin-torque oscillators with thermal noise: A constant energy orbit approach, Phys. Rev. B **90,** 174405 (2014).

[27] Y. Yu, C. B. Muratov, and R. O. Moore, Importance sampling for thermally induced switching and non-switching probabilities in spin-torque magnetic nanodevices, IEEE Trans. Magn. **55,** 1 (2019).

[28] N. Kani, Ph.D. thesis, Georgia Institute of Technology, (2017).

[29] S. Hollos, Rational approximations for complete elliptic integrals, http://www.exstrom.com/math/elliptic/ellipint.html (2012), Accessed: 2019-07-31.

[30] 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 **62,** 164 (2014).

[31] H. Liu, D. Bedau, J. Sun, S. Mangin, E. Fullerton, J. Katine, and A. Kent, Dynamics of spin-torque switching in all-perpendicular spin valve nanopillars, J. Magn. Magn. Mater. **358,** 233 (2014).

[32] A file to numerically evaluate the inverse of a function in Python, https://github.com/alvarosg/pynverse/blob/master/pynverse/inverse.py (2016).

[33] N. Kani, S. Rakheja, and A. Naeemi, A probability-density function approach to capture the stochastic dynamics of the nanomagnet and impact on circuit performance, IEEE Trans. Electron Devices **63,** 4119 (2016).

[34] K. C. Chun, H. Zhao, J. D. Harms, T.-H. Kim, J.-P. Wang, and C. H. Kim, A scaling roadmap and performance evaluation of in-plane and perpendicular MTJ based STT-MRAMs for high-density cache memory, IEEE J. Solid-State Circuits **48,** 598 (2012).

[35] J. M. Coey, *Magnetism and Magnetic Materials* (Cambridge University Press, 2010).

[36] G. Rowlands, T. Rahman, J. Katine, J. Langer, A. Lyle, H. Zhao, J. Alzate, A. Kovalev, Y. Tserkovnyak, Z. Zeng,

H. W. Jiang, K. Galatsis, Y. M. Huai, P. K. Amiri, K. L. Wang, I. N. Krivorotov, and J.-P. Wang, Deep subnanosecond spin torque switching in magnetic tunnel junctions with combined in-plane and perpendicular polarizers, Appl. Phys. Lett. **98**, 102509 (2011).

[37] S. Yakata, H. Kubota, Y. Suzuki, K. Yakushiji, A. Fukushima, S. Yuasa, and K. Ando, Influence of perpendicular magnetic anisotropy on spin-transfer switching current in CoFeB/MgO/CoFeB magnetic tunnel junctions, J. Appl. Phys. **105**, 07D131 (2009).

[38] P. Dürrenfeld, F. Gerhard, M. Ranjbar, C. Gould, L. Molenkamp, and J. Åkerman, Spin Hall effect-controlled magnetization dynamics in NiMnSb, J. Appl. Phys. **117**, 17E103 (2015).

[39] H. Yoda *et al.*, High efficient spin transfer torque writing on perpendicular magnetic tunnel junctions for high density MRAMs, Curr. Appl. Phys. **10**, e87 (2010).

[40] H. Nagasaka, N. Maruyama, A. Nukada, T. Endo, and S. Matsuoka, in *International Conference on Green Computing* (IEEE, Chicago, Illinois, 2010), p. 115.

[41] M. Abadi *et al.*, in *12th* {USENIX} *Symposium on Operating Systems Design and Implementation (*{OSDI} *16)* (PeerJ, Savannah, GA, 2016), p. 265.

[42] G. Indiveri and S.-C. Liu, Memory and information processing in neuromorphic systems, Proc. IEEE **103**, 1379 (2015).

[43] N. K. Upadhyay, H. Jiang, Z. Wang, S. Asapu, Q. Xia, and J. Joshua Yang, Emerging memory devices for neuromorphic computing, Adv. Mater. Technol. **4**, 1800589 (2019).

[44] J. Grollier, D. Querlioz, and M. D. Stiles, Spintronic nanodevices for bioinspired computing, Proc. IEEE **104**, 2024 (2016).

[45] K. Y. Camsari, S. Salahuddin, and S. Datta, Implementing p-bits with embedded MTJ, IEEE Electron Device Lett. **38**, 1767 (2017).

[46] A. F. Vincent, J. Larroque, N. Locatelli, N. B. Romdhane, O. Bichler, C. Gamrat, W. S. Zhao, J.-O. Klein, S. Galdin-Retailleau, and D. Querlioz, Spin-transfer torque magnetic memory as a stochastic memristive synapse for neuromorphic systems, IEEE. Trans. Biomed. Circuits Syst. **9**, 166 (2015).

[47] J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubuta, S. Yuasa, M. D. Stiles, and J. Grollier, Neuromorphic computing with nanoscale spintronic oscillators, Nature **547**, 428 (2017).

[48] A. Sengupta and K. Roy, Encoding neural and synaptic functionalities in electron spin: A pathway to efficient neuromorphic computing, Appl. Phys. Rev. **4**, 041105 (2017).

[49] K. Yue, Y. Liu, R. K. Lake, and A. C. Parker, A brain-plausible neuromorphic on-the-fly learning system implemented with magnetic domain wall analog memristors, Sci. Adv. **5**, eaau8170 (2019).

[50] N. Rangarajan, A. Parthasarathy, N. Kani, and S. Rakheja, Energy-efficient computing with probabilistic magnetic bits–performance modeling and comparison against probabilistic CMOS logic, IEEE Trans. Magn. **53**, 1 (2017).

[51] R. Venkatesan, S. Venkataramani, X. Fong, K. Roy, and A. Raghunathan, in *2015 Design, Automation & Test in Europe Conference & Exhibition (DATE)* (IEEE, Grenoble, France, 2015), p. 1575.

[52] K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, Stochastic P-Bits for Invertible Logic, Phys. Rev. X **7**, 031014 (2017).

[53] Z. Li and S. Zhang, Thermally assisted magnetization reversal in the presence of a spin-transfer torque, Phys. Rev. B **69**, 134416 (2004).

[54] T. Taniguchi and H. Imamura, Thermally assisted spin transfer torque switching in synthetic free layers, Phys. Rev. B **83**, 054432 (2011).

[55] D. Pinna, A. Mitra, D. Stein, and A. Kent, Thermally assisted spin-transfer torque magnetization reversal in uniaxial nanomagnets, Appl. Phys. Lett. **101**, 262401 (2012).

[56] T. Taniguchi, Y. Utsumi, and H. Imamura, Thermally activated switching rate of a nanomagnet in the presence of spin torque, Phys. Rev. B **88**, 214414 (2013).