Switching time of spin-torque-driven magnetization in biaxial ferromagnets

Ankit Shukla[*]

Holonyak Micro and Nanotechnology Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Arun Parthasarathy[†]

Department of Electrical and Computer Engineering, New York University, Brooklyn 11201, USA

Shaloo Rakheja[‡]

Holonyak Micro and Nanotechnology Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801
(Dated: June 24, 2020)

We analytically model the magnetization switching time of a biaxial ferromagnet driven by an antidamping-like spin torque. The macrospin magnetization dynamics is mapped to an energy-flow equation, wherein a rational-function approximation of the elliptic integrals for moderate spin current and small damping results in a closed-form expression of the switching time. Randomness in the initial angle of magnetization gives the distribution function of the switching time. The analytic model conforms to the results obtained from Monte Carlo simulation for a broad range of material parameters. Our results can ameliorate design and benchmarking of in-plane spin torque magnetic memory by obviating 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 non-magnetic (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 fabrication complexity [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 the 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. 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 expres-

———

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

sions of the switching time for the uniaxial anisotropy, but not for the more general biaxial anisotropy presented in this work.

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 ~ 50 nm for Co, Fe, Ni [16, 17]. The macrospin model has been applied successfully to explain measured probability distribution of switching dynamics in ferromagnetic thin films with dimensions on the order of 50 nm [18]. Inhomogeneous switching dynamics, such as due to domain wall nucleation and propagation, have been experimentally measured in large ferromagnetic samples [19, 20]. However, consideration of multi-domain switching of the ferromagnet is out of scope of this work.

Within the macrospin approximation there are three equivalent approaches to analyze the magnetization dynamics: perturbative approach [12], constructing a Fokker-Planck representation [21, 22], or using a constant-energy stochastic equation [23–26]. In this work, we adopt the 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 into a tractable one-dimensional (1D) problem at low temperature and for low-to-moderate applied spin current [24–26].
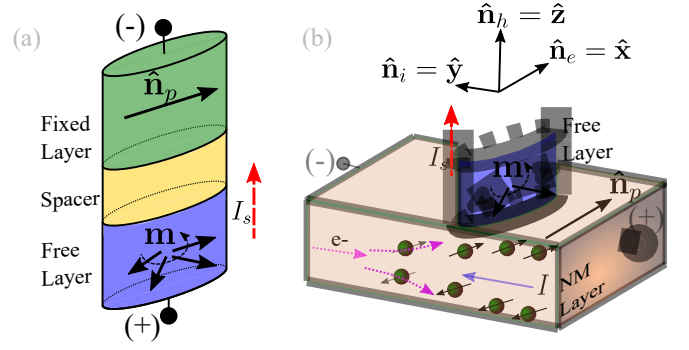
The switching dynamics is modeled as a slow perturbation of the rapid constant-energy gyration around the easy axis (Sec. II). In the deterministic limit, a closed-form expression of the switching time as a function of input spin current, initial magnetization energy, and material parameters is obtained. Average switching time and the probability distribution of the switching time follow for an initial Boltzmann distributed ensemble of spins (Sec. III). Analytic results are shown to conform to those 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 magnetization subject to an effective magnetic field, intrinsic damping, and spin torque is described by the Landau-Lifshitz-Gilbert (LLG) equation. Using dimensionless form of physical parameters listed in Table I, the LLG equation is [24, 27]

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

where $\mathbf{m}$ is the normalized free-layer magnetization, α is the Gilbert damping constant, I_s is the input spin cur-

coincide with $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ axes, respectively. Using Eq. (3) in Eq. (2), we obtain $\mathbf{h} = m_x\hat{\mathbf{x}} - Rm_z\hat{\mathbf{z}} + \mathbf{H}_a/H_k$ where $R = D_h = M_s/H_k$.

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 the thermal bath and neglecting Joule heating, the diffusion coefficient

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

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 thermal energy $k_B T$.

A. Constant-Energy Orbit Averaging (CEOA)

The energy of a macrospin is conserved when damping, thermal field (noise), external magnetic field, and spin torque are absent. In this case, $\mathbf{m}$ 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, $\mathbf{m}$ loses energy, eventually relaxing to a stable equilibrium state ($m_x = \pm 1$). A non-zero 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 time-scale corresponding to perpendicular diffusion of magnetization from one constant-energy orbit to another as a result of damping, spin torque, and themal field. Figure 2(b) shows one such trajectory where the magnetization switches from an anti-parallel well to a parallel well under the simultaneous effects of damping, input spin current, and thermal field

The rate of change of macrospin energy due to the non-conservative torques is given as

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

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 [24, 26] (see
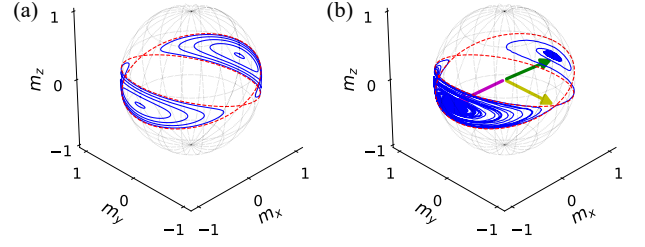

FIG. 2. (a) Constant-energy curves for $g_L < 0$ in both anti-parallel and parallel wells for $R = 15$. (b) Switching trajectory of a magnetization initially in the anti-parallel well for $R = 15$ and a large input spin current. External field is assumed to be absent. The arrows shown in the figure mark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

Appendix A)

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

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 $\frac{\partial g_L}{\partial \tau}$ 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 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, $I_s^{\text{th0}} = \alpha\left[\frac{2}{\pi}\sqrt{R(1 + R)}\right]$ 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 energy for deterministic dynamics. It is positive for the complete range of macrospin energy only for $I_s^{\text{thm}} = \max\left(I_s^{\text{th0}}, I_s^{\text{th1}}\right) < I_s$ which is consistent with Fig. 3(a). For other values of current, thermal assistance is required for switching. The CEOA is valid when the variation in macrospin energy over one precessional cycle is small, i.e. $|T\langle\frac{\partial g_L}{\partial \tau}\rangle| \ll \max[|g_L|] = 1$. To satisfy this constraint on the variation of macrospin energy, the maximum spin current is given as $I_s^{\text{thM}} = I_s^{\text{th0}}\left[1 + 1/\left(8\alpha\sqrt{R}\right)\right]$.
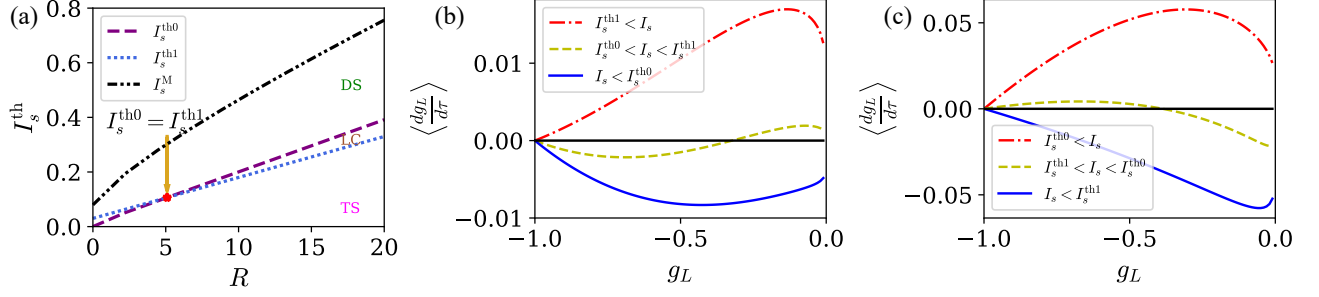
FIG. 3. (a) Threshold spin current versus R. The arrow marks $R = R_c = 5.094$ where $I_s^{\text{th0}} = I_s^{\text{th1}}$. For $I_s^{\text{thm}} < I_s$, deterministic switching (DS) occurs. Thermal switching region is denoted as TS, while the regime of limit cycles is marked as LC. (b) For $R = 3.0$ and $I_s^{\text{th1}} < I_s$, the rate of change of macrospin energy is positive for $g_L \in (-1, 0)$. For $I_s^{\text{th0}} < I_s < I_s^{\text{th1}}$, thermal assistance is necessary to cause magnetization reversal. For $I_s < I_s^{\text{th0}}$, switching occurs due to thermal noise. (c) For $R = 15.0$ and $I_s^{\text{th0}} < I_s$, reversal is deterministic as the rate of energy change is always positive. For $I_s < I_s^{\text{th1}}$, thermal switching dominates, while in the intermediate region, thermal assistance is required to cause switching. These results correspond to $\alpha = 0.03$.

III. ANALYTIC 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}. Analytic solutions of the switching time presented in this section rely on one or more of the following rational approximations [29] for elliptic integrals of Eq. (7)—

$$\frac{2}{\pi}\left[E\left(R, g_L\right) + g_L K\left(R, g_L\right)\right] = A(R)g_L^2 + B(R)g_L + C(R),$$
(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 $x = R\frac{(1+g_L)}{(R-g_L)}$. Defining $\tilde{I}_s = I_s/\alpha$ and using Eqs. (8a) and (8b) together in Eq. (7), we obtain the magnetization switching time (see Appendix B for derivation)

$$\tau_s = \frac{1}{4\alpha}\int_{g_{L_i}}^{g_{L_f}}\left[\frac{3R - g_L(R+4)}{R - g_L(R+2)}\right]\sqrt{\frac{1+R}{R-g_L}}\times$$
$$\frac{\partial g_L}{\left[\tilde{I}_s\left(1+g_L\right) - \sqrt{1+R}\sqrt{R-g_L}\left(Ag_L^2 + Bg_L + C\right)\right]}.$$
(9)

In the above equation, the parameters A, B, and C are functions of R given as $k_1 + k_2 R^{k_3}$, 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) is simplified and integrated using partial fractions to arrive at the closed-

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

A major advantage of this analytic result is that the approximations of elliptic integrals are independent of the input spin current and Gilbert damping. Additionally, the results obtained in this work are valid for a broad range of R as opposed to prior works [12, 30] that are valid only for $R < R_c \ (= 5.09)$.

For the case of $R \to \infty$, 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

TABLE II. 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
	A	0.35661	-0.51244	-0.38689
$[1, 3]$	B	1.05148	-0.55504	-0.28598
	C	0.61670	0.03018	-1.00153
	A	0.20223	-0.38439	-0.68424
$[3, 50]$	B	0.81746	-0.34729	-0.63939
	C	0.61765	0.02994	-1.08243
	A	0.17370	-0.51992	-0.97986
$[50, 100]$	B	0.78501	-0.48295	-0.97726
	C	0.61755	0.02625	-1.01442

as (see Appendix C for details)

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

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

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 as [12–14, 24, 31] (see Appendix D for details)

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

A. Equilibrium Distribution and Average Switching Time

In the absence of 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 as

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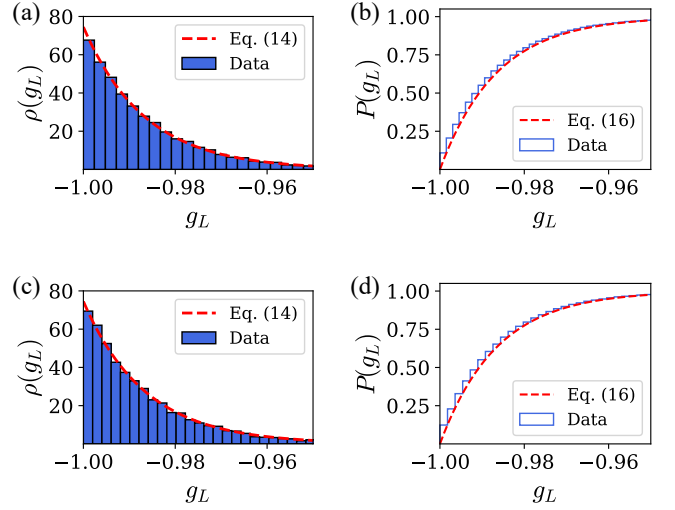w_{eq}(\mathbf{m}) = \frac{1}{Z(\Delta_0, R)} \exp\left(-\Delta_0 g_L(\mathbf{m}, 0)\right), \qquad (13)
$$

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

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

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

where the partition function

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

The cumulative distribution function (CDF) is

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

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

The magnetization is considered to have switched successfully when it crosses the separatrix $\left(g_{L_f} = 0\right)$ and consequently moves into the adjoining energy well. Once the magnetization moves into the target energy well, the spin current could be switched off. The magnetization would eventually settle into its stable well due to its intrinsic damping. Therefore, $\langle\tau_s\rangle$ is given as

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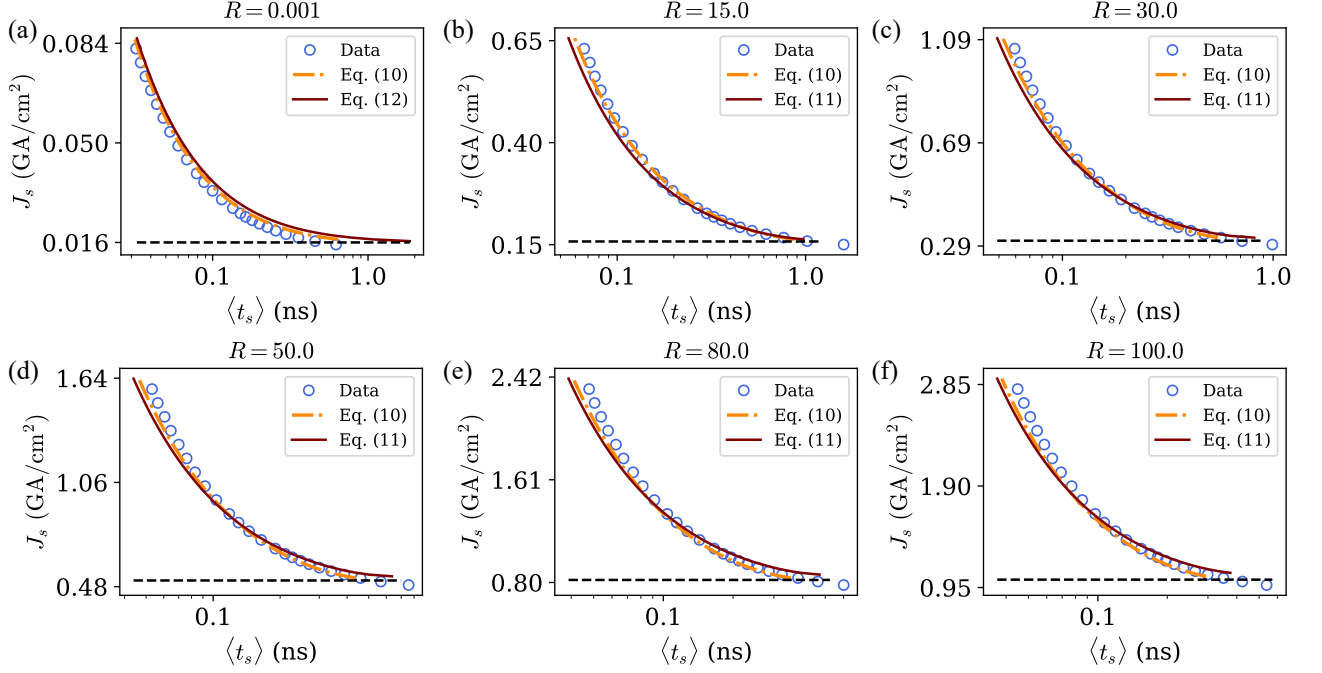$$

FIG. 5. Average switching time $\langle t_s \rangle = \frac{(1+\alpha^2)}{\gamma \mu_0} \frac{\langle \tau_s \rangle}{H_k}$ 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 region of deterministic switching from that of thermal activation. The numerical data is obtained for an ensemble of 10^4 macrospins.

B. Distribution Functions and Write Error Rate

Defining a random variable switching time T_s, the fraction of macrospins in an ensemble that have switched from an anti-parallel to a parallel state at time τ_s is

$$
\begin{aligned}
P_{T_s}\left[\tau_s\right] &= 1 - \int_{-1}^{g_{L_i}(\tau_s)} \frac{1}{Z(\Delta_0, R)} \frac{\exp\left(-\Delta_0 x\right)}{\sqrt{-x}} dx \\
&= \exp\left(-\Delta_0\left(1 + g_{L_i}(\tau_s)\right)\right) \frac{F\left(\sqrt{-\Delta_0 g_{L_i}(\tau_s)}\right)}{F\left(\sqrt{\Delta_0}\right)},
\end{aligned}
\tag{18}
$$

where $g_{L_i}(\tau_s)$ corresponds to the initial energy for switching time τ_s. An exact analytic 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 [32] them. $P_{T_s}[\tau_s]$ includes all macrospins in an ensemble with $g_L \geq g_{L_i}(\tau_s)$ as we consider only deterministic switching in this work. $P_{T_s}[\tau_s]$ 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

$$
\begin{aligned}
\rho_{T_s}(\tau_s) &= \frac{dP_{T_s}\left[\tau_s\right]}{d\tau_s} \\
&= \frac{\sqrt{\Delta_0}}{2F(\sqrt{\Delta_0})} \frac{\exp\left(-\Delta_0(1 + g_{L_i}(\tau_s))\right)}{\sqrt{-g_{L_i}(\tau_s)}} \left|\frac{dg_{L_i}(\tau_s)}{d\tau_s}\right|.
\end{aligned}
\tag{19}
$$

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

$$
\begin{aligned}
\text{WER} &= 1 - P_{T_s}[\tau] \\
&= 1 - \exp\left(-\Delta_0\left(1 + g_{L_i}(\tau_s)\right)\right) \frac{F\left(\sqrt{-\Delta_0 g_{L_i}(\tau_s)}\right)}{F\left(\sqrt{\Delta_0}\right)}.
\end{aligned}
\tag{20}
$$

C. Model Validation

To benchmark our analytic results, we solve Eq. (1) numerically using the Heun integration scheme implemented in CUDA and run in parallel on GPUs. Numerical simulations were calibrated against published results to ensure their accuracy [13, 24, 33]. For all numerical simulations presented in this work, the time step of integration was set as 0.3 ps. The Gilbert damping constant of the free-layer is taken as $\alpha = 0.03$, while its cross-sectional 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 switching dynamics remain the same. Finally, a wide
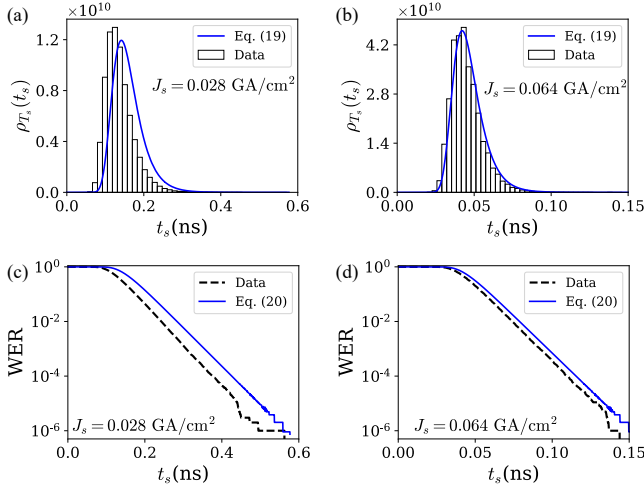
FIG. 6. All results reported for $R = 0.001$ (uniaxial anisotropy) and $J_s^{\mathrm{th1}} = 1.6 \times 10^7$ A/cm^2. The top panel shows the PDF, while the bottom panel shows the WER. The accuracy of analytical results improves as the input spin current density increases with respect to J_s^{th1}.
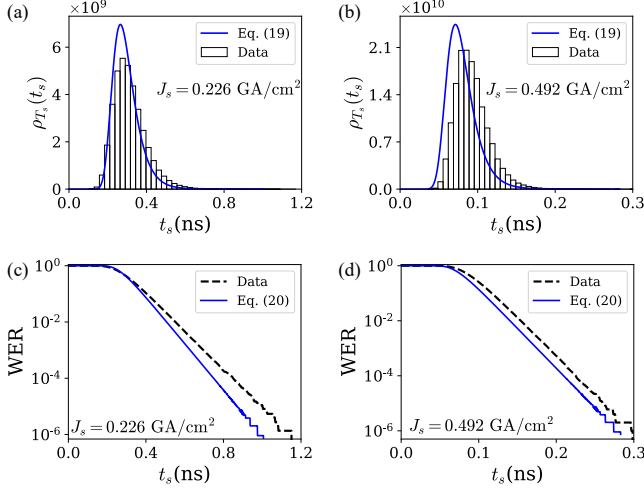

FIG. 8. All results are reported for $R = 50$ and $J_s^{\mathrm{th0}} = 5.14 \times 10^8$ A/cm^2. The top panel shows the PDF, while bottom panel shows the WER. The accuracy of analytic results improves for spin current density larger than the threshold current density. The analytic model is fairly accurate even for $J_s > J_s^{\mathrm{thM}} = 1.6 J_s^{\mathrm{th0}}$ showing the robustness of our model.

FIG. 7. All results are reported for $R = 15$ and $J_s^{\mathrm{th0}} = 1.58 \times 10^8$ A/cm^2. The top panel shows the PDF, while the bottom panel shows the WER. The accuracy of analytical solutions improves for spin current densities larger than the threshold current density. Though CEOA is strictly valid for $J_s < J_s^{\mathrm{thM}} = 2.1 J_s^{\mathrm{th0}}$, the analytic model agrees well with the numerical solution for current larger than J_s^{thM}.

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.
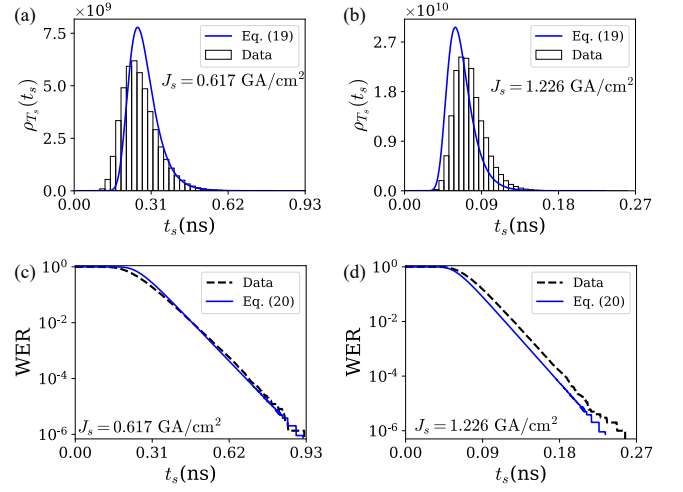
IV. RESULTS AND DISCUSSION

Figure 5 shows the average switching time of an ensemble of 10^4 independent macrospins, $\langle t_s \rangle = \frac{(1+\alpha^2)}{\gamma \mu_0} \frac{\langle \tau_s \rangle}{H_k}$ as a function of the input spin current density, $J_s = \Delta_0 (4 e k_B T / \hbar)(I_s / Ar)$, where e is the electron charge, and $\hbar$ is the reduced Planck's constant. For $R = 0.001$ (uniaxial anisotropy), there is an excellent agreement between numerical data and closed-form solutions given in Eqs. (10) and (12) for moderate to large current levels. However, for current levels approaching the threshold switching currents, 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 compared to the numerical results for current levels comparable to the threshold value. This is expected since the analytic 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} ($= \Delta_0 (4 e k_B T / \hbar)(I_s^{\mathrm{th0}} / Ar)$), the agreement between analytic and numerical results improves. Near the threshold current level, the average switching time obtained from Eq. (10) is slightly lower than that obtained from numerical results. This slight deviation is due to the quadratic approximation of elliptic integrals in Eq. (7). Note that for $J_s > J_s^{\mathrm{thM}}(= \Delta_0 (4 e k_B T / \hbar)(I_s^{\mathrm{thM}} / Ar))$, the validity of CEOA is not fully justified as the average rate of energy change becomes large [24]. However, the general trend of 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

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

Materials	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]
FeGaB	1.63	0.0198	53.4	0.1	[28]

time from solution of Eqs. (1) and (10) is between $5-12\%$ for different values of R for current as large as $2J_s^{\mathrm{thM}}$.

The PDF and WER obtained using analytic results of Eqs. (19) and (20) are compared against numerical solution in Figs. 6-9 for an ensemble of 10^6 macrospins. It is observed that for 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 analytic solutions also increases as current increases from J_s^{th0} toward J_s^{thM}. Though for spin currents larger than J_s^{thM}, the accuracy of analytic results reduces as the validity of CEOA becomes questionable, the model is fairly robust and predicts numerical results well. To arrive at analytic results reported in Figs. 7-9, we numerically invert Eq. (11) due to its simplicity.
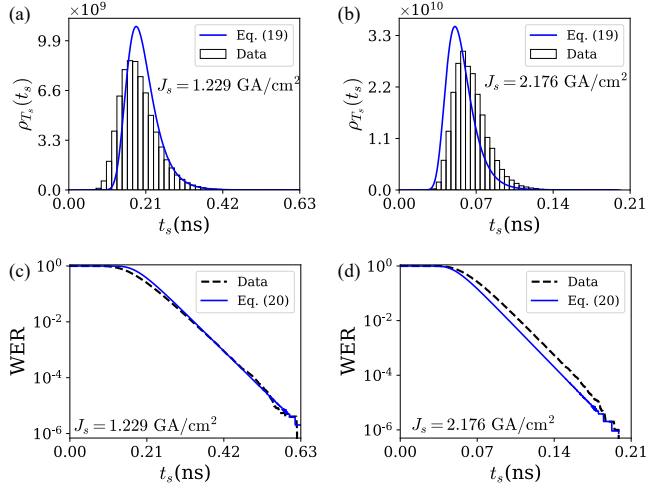

FIG. 9. All results are reported for $R = 100$ and $J_s^{\mathrm{th0}} = 1.02 \times 10^9$ A/cm^2. The top panel shows the PDF, while the bottom panel shows the WER. The accuracy of analytical results improves as J_s increases above J_s^{th0}. For $J_s > J_s^{\mathrm{thM}} = 1.42 J_s^{\mathrm{th0}}$, the accuracy of the analytic model reduces but continues to predict the numerical results with good accuracy.

Figure 10 shows the average switching time for a few ferromagnetic thin films. Both numerical and analytic average switching times (Eqs. (1) and (10)) are for the

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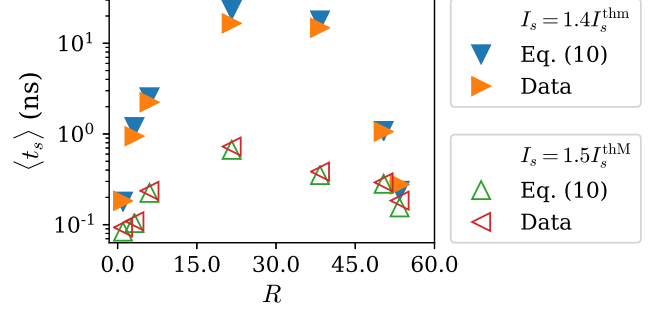

FIG. 10. Average switching time for materials listed in Table (III) for two different currents, viz. $I_s = 1.4 I_s^{\mathrm{thm}}$ and $I_s = 1.5 I_s^{\mathrm{thM}}$.

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

Conventional artificial intelligence (AI) hardware accelerators based on von Neumann architectures suffer from significant power dissipation and communication overheads [40–42]. The introduction of non-volatile magnetic memory in a 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 co-locate memory and compute elements are well suited to leverage the non-volatile spin-torque memory discussed in this work. Such memory can be used as a cache memory in general-purpose processors and for storing the weights pertinent to machine learning tasks in a neuromorphic architecture [44].

Beyond serving as non-volatile memory elements, the device architectures illustrated in Fig. 1 can also function as stochastic oscillators when operated in the thermally activated regime for which $I_s < I_s^{\mathrm{thm}}$. 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, by relaxing the constraint on the energy barrier, the thresholding behavior of mono-domain ferromagnets can mimic neural dynamics in an energy efficient manner. In the case of multi-domain effects, it has been shown that the structures of Fig. 1 can implement synaptic functionality [46, 49], which is necessary for weight storage and update in neu-

romorphic hardware. Energy efficient probabilistic computing for image processing and machine learning [50], and spin logic [51, 52] are 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 spin torque driven dynamics of ferromagnetic materials. The scaling of the energy barrier height of ferromagnets with spin current is typically modeled as $\Delta = \Delta_0 \left(1 - I_s/I_s^{\text{th1}}\right)^\xi$. The value of ξ in the limit of $I_s \ll I_s^{\text{th1}}$ has been a subject of debate for over two decades [21, 53–56]. Recently, it was theoretically postulated [54, 55] that $\xi = 2$ for 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-Plank (FP) equation [56]. Pinna and coworkers note the current dependence of ξ via the the principle of action [24, 55] on Eq. (7). In our opinion, it might be possible to construct a FP representation from Eq. (7) and solve it with the rational approximations Eqs. (8b)-(8c) to arrive at an analytical expression for ξ. However, full analysis in the thermally activated regime is out of scope of this work.

V. CONCLUSION

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

Appendix A: Simplifying Eq. (6)

Without any loss of generality, we consider the easy and hard axes to coincide with the $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ axes, respectively. Next, we consider spin polarization of the fixed layer to be in the plane of the magnet at an angle ϕ with the easy axis, 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 $\mathbf{h}_{\text{eff}} = m_x\hat{\mathbf{x}} - Rm_z\hat{\mathbf{z}} + \mathbf{h}_T$, where $\mathbf{h}_T$ is the thermal field. Therefore, the three components of Eq. (1) are

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

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

and

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

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

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

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

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

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

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

and

$$
\begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} = \begin{pmatrix} Rm_z\mathcal{D}_{31} - m_x\mathcal{D}_{11} \\ Rm_z\mathcal{D}_{32} - m_x\mathcal{D}_{12} \\ Rm_z\mathcal{D}_{33} - m_x\mathcal{D}_{13} \end{pmatrix}.
\tag{A3b}
$$

It can be observed from Eq. (A3a) that each of damping, spin current and thermal noise contribute to the rate of energy change. In addition, the last term in Eq. (A3a) is 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

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

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

Now average $\frac{\partial g_L}{\partial \tau}$ over a constant-energy orbit as the non-conservative effects act over long time-scales whereas we are more interested in studying their effects on rapid periodic motion. In the limit of zero damping, zero spin current at absolute zero temperature, the constant energy solutions to Eqs. (A1a)-(A1c) in the anti-parallel well are given as [23, 24, 26]

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

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

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

where $k^2 \equiv R\frac{1 + g_L}{R - g_L}$, $0 < k^2 < 1$ and $sn[\cdot]$, $cn[\cdot]$, $dn[\cdot]$ are Jacobi elliptic functions. Magnetization oscillates around the easy axis with a time period, $T(g_L) = \frac{4}{\sqrt{R - g_L}}K(k^2)$, where $K(k^2)$ is the complete elliptic integral of first kind. In Eq. (A4) only the time averages of m_x and m_x^2 are non-zero while 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 parametrize [24, 26] the constant-energy orbit in the anti-parallel well as a function of parameter ω as

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

$$
\frac{1}{-g_L}m_x^2 - \frac{R}{-g_L}m_z^2 = 1,
\tag{A6b}
$$

which leads to the form

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

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

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

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

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

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

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

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

and

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

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

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

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

Appendix B: Solving the integral in Eq. (9)

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

$$\tau_s = \frac{1}{2A\alpha} \frac{R+4}{R+2} \int\limits_{\sqrt{R-g_{L_i}}}^{\sqrt{R-g_{L_f}}} \frac{u^2 - \frac{R(1+R)}{R+4}}{u^2 - \frac{R(1+R)}{R+2}} \left[\frac{du}{P(u)} \right], \quad \text{(B1)}$$

where $P(u) = u^5 - (B/A + 2R)u^3 + \frac{\tilde{I}_s}{A\sqrt{1+R}}u^2 + (R^2 + (B/A)R + C/A)u - \tilde{I}_s\sqrt{1+R}/A$ and 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

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

Appendix C: $R \to \infty$

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

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

Now substituting $1 + g_L = x$ in the previous equation leads us to

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

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

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

where $E = \frac{160\tilde{I}_s - 60R}{16\tilde{I}_s - 7R}$ and $F = \frac{256\tilde{I}_s - 128R}{16\tilde{I}_s - 7R}$. 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

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

Appendix D: Uniaxial Limit

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

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

which can then be integrated as

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

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

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

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

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

$$P_Y[y] = \mathbb{P}[Y \leq y] = \mathbb{P}[-X^2 \leq y]$$
$$= 1 - \mathbb{P}[|X| \leq \sqrt{-y}] \quad \text{(E1)}$$
$$= 1 - \mathbb{P}[-\sqrt{-y} \leq X \leq \sqrt{-y}]$$

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

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

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

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

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

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

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

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

[1] A. D. Kent and D. C. Worledge, A new spin on magnetic memories, Nature Nanotechnology **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, Physical Review 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, Proceedings of the IEEE **104**, 1919 (2016).

[4] J. C. Slonczewski, Current-driven excitation of magnetic multilayers, Journal of Magnetism and Magnetic Materials **159**, L1 (1996).

[5] L. Berger, Emission of spin waves by a magnetic multilayer traversed by a current, Physical Review B **54**, 9353 (1996).

[6] J. Sinova, S. O. Valenzuela, J. Wunderlich, C. H. Back, and T. Jungwirth, Spin Hall effects, Reviews of Modern Physics **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, Journal of Physics: Condensed 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, Physical Review Letters **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, Nature Nanotechnol-

ogy **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, Nature Nanotechnology **11**, 621 (2016).

[12] J. Z. Sun, Spin-current interaction with a monodomain magnetic body: A model study, Physical Review 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, Journal of Applied Physics **114**, 033901 (2013).

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

[15] W. F. Brown Jr, Thermal fluctuations of a single-domain particle, Physical Review **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, Applied Physics Letters **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, Physical Review 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, Applied Physics Letters **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, Nature Nanotechnology **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, Nature Nanotechnology , 1 (2020).

[21] D. M. Apalkov and P. B. Visscher, Spin-torque switching: Fokker-Planck rate calculation, Physical Review 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, Physical Review 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, Physical Review 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, Journal of Applied Physics **113**, 184105 (2013).

[26] D. Pinna, D. L. Stein, and A. D. Kent, Spin-torque oscillators with thermal noise: A constant energy orbit approach, Physical Review 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 Transactions on Magnetics (2019).

[28] N. Kani, *"Modeling of magnetization dynamics and applications to spin-based logic and memory devices"*, 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 Transactions on 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, Journal of Magnetism and Magnetic Materials **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 Transactions on 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 Journal of 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, Applied Physics Letters **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, Journal of Applied Physics **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, Journal of Applied Physics **117**, 17E103 (2015).

[39] H. Yoda, T. Kishi, T. Nagase, M. Yoshikawa, K. Nishiyama, E. Kitagawa, T. Daibou, M. Amano, N. Shimomura, S. Takahashi, T. Kai, M. Nakayama, H. Aikawa, S. Ikegawa, M. Nagamine, J. Ozeki, S. Mizukami, M. Oogane, Y. Ando, S. Yuasa, K. Yakushiji, H. Kubota, Y. Suzuki, Y. Nakatani, T. Miyazaki, and K. Ando, High efficient spin transfer torque writing on perpendicular magnetic tunnel junctions for high density MRAMs, Current Applied Physics **10**, e87 (2010).

[40] H. Nagasaka, N. Maruyama, A. Nukada, T. Endo, and S. Matsuoka, Statistical power modeling of GPU kernels using performance counters, in *International conference on green computing* (IEEE, 2010) pp. 115–122.

[41] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, S. Monga, Rajat snd Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, Tensorflow: A system for large-scale machine learning, in *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)* (2016) pp. 265–283.

[42] G. Indiveri and S.-C. Liu, Memory and information processing in neuromorphic systems, Proceedings of the 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, Advanced Materials Technologies **4**, 1800589 (2019).

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

[45] K. Y. Camsari, S. Salahuddin, and S. Datta, Implementing p-bits with embedded MTJ, IEEE Electron Device Letters **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 Transactions on Biomedical Circuits and Systems **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, Applied Physics Reviews **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,

Science advances **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 Transactions on Magnetics **53**, 1 (2017).

[51] R. Venkatesan, S. Venkataramani, X. Fong, K. Roy, and A. Raghunathan, Spintastic: Spin-based stochastic logic for energy-efficient computing, in *2015 Design, Automation & Test in Europe Conference & Exhibition (DATE)* (IEEE, 2015) pp. 1575–1578.

[52] K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, Stochastic p-bits for invertible logic, Physical Review X **7**, 031014 (2017).

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

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

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

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