

AVERAGED EQUATION FOR ENERGY DIFFUSION ON A GRAPH REVEALS BIFURCATION DIAGRAM AND THERMALLY ASSISTED REVERSAL TIMES IN SPIN-TORQUE DRIVEN NANOMAGNETS

KATHERINE A. NEWHALL AND ERIC VANDEN-EIJNDEN

COURANT INSTITUTE OF MATHEMATICAL SCIENCE, NEW YORK UNIVERSITY, NEW YORK,
NEW YORK 10012

ABSTRACT. Driving nanomagnets by spin-polarized currents offers exciting prospects in magnetoelectronics, but the response of the magnets to such currents remains poorly understood. We show that an averaged equation describing the diffusion of energy on a graph captures the low-damping dynamics of these systems. From this equation we obtain the bifurcation diagram of the magnets, including the critical currents to induce stable precessional states and magnetization switching, as well as the mean times of thermally assisted magnetization reversal in situations where the standard reaction rate theory of Kramers is no longer valid. These results agree with experimental observations and give a theoretical basis for a Néel-Brown-type formula with an effective energy barrier for the reversal times.

Manipulating thin-film magnetic elements with spin-polarized currents besides external magnetic fields [12] has generated a lot of recent interest in applications to magnetoelectronic devices that offer low power memory storage without the use of moving parts [2]. Understanding the response of the magnet to such currents is nontrivial, however, because they apply a nonconservative force, called spin-transfer torque (STT), on the system. Like other nongradient systems with no Lyapunov function, the phase portrait of nanomagnets in the presence of STT can be quite complex, and include limit cycles or chaotic trajectories besides fixed points. When the applied fields and/or currents are nonstationary, or in the presence of thermal noise, the situation is even worse. In particular, Kramers' reaction rate theory [14, 10] is no longer applicable and the Néel-Brown formula [5] for the mean magnetization reversal time is not valid since there is no well-defined energy associated with STT.

Nanomagnets typically operate in a regime where the nonconservative parts of the dynamics, including the effects of damping, STT, and thermal noise, act on time-scales that are much longer than that of the energy-conserving Hamiltonian part. Trajectories remain close to periodic Hamiltonian orbits for a long time, and slowly drift from one orbit to another due to damping, STT, and thermal noise. This separation of time scales can be exploited, using averaging techniques developed by Freidlin and Wentzell [9, 8] (see also [20, 4]), to reduce the dynamics to that of an energy diffusing on a graph. We show here that this reduced description permits to explain the features of nanomagnets subject to STT that are observed experimentally. Specifically, we obtain the full bifurcation diagram of the system at zero temperature and determine the critical spin-polarized currents needed to induce stable precessional states [15, 3] and magnetization switching [15, 19]. At finite temperature, we calculate the mean times of thermally assisted magnetization reversals [15, 18], and give expressions for the effective energy barriers conjectured to exist [15, 18, 1, 16, 17]. These

results are complementary to those obtained in [6], using the geometrical Minimum Action Method (gMAM) [11], for situations with small thermal noise and stronger damping.

We will focus on magnetic systems in which the magnetization has constant strength M_s in the direction of a unit vector $\mathbf{m}(t) = (m_x(t), m_y(t), m_z(t))$ whose evolution is governed by

$$(1) \quad \dot{\mathbf{m}} = -\mathbf{m} \times \mathbf{h}_{\text{eff}} + \mathbf{m} \times (\mathbf{m} \times (-\alpha \mathbf{h}_{\text{eff}} + a_J \mathbf{m}_p)).$$

This is the standard stochastic Landau-Lifshitz-Gilbert (LLG) equation written in non-dimensional form (details of the nondimensionalization can be found in the Appendix A), with an additional STT term, $\mathbf{m} \times (\mathbf{m} \times a_J \mathbf{m}_p)$, modeling the transfer of angular momentum to the magnetization from the electron spin in a polarized current of strength a_J directed along unit vector \mathbf{m}_p [21]. For simplicity, here we take $\mathbf{m}_p = (1, 0, 0)$ and a constant strength a_J , but these could straightforwardly be generalized to any direction and a time-varying strength. The other terms in (1) are standard: α is the damping parameter and $\mathbf{h}_{\text{eff}} = -\nabla_{\mathbf{m}} E + \sqrt{\frac{2\alpha\epsilon}{1+\alpha^2}} \boldsymbol{\eta}(t)$ is the effective field, which in turn is the sum of the negative gradient of the energy per volume, $E(\mathbf{m})$, and a term accounting for thermal effects with $\boldsymbol{\eta}(t)$ being three-dimensional white-noise and $\epsilon = k_B T / \mu_0 M_s^2 \nu$ the non-dimensional temperature (ν is the magnet volume and μ_0 the permeability of free-space). Here we take $E(\mathbf{m}) = \beta_y m_y^2 + \beta_z m_z^2 - h_x m_x$ with $\beta_y < \beta_z = 1/2$, corresponding to biaxial anisotropy with a planar applied field of amplitude h_x in the x -direction.

The presence of the STT term in (1), which is nongradient and nonconservative, complicates the analysis of this equation even in the absence of thermal noise ($\epsilon = 0$). In particular, the magnetic energy $E(\mathbf{m})$ is not a Lyapunov function for the system, and it is unclear if the STT term provides additional damping, driving, or something entirely different. To understand the effect of this term, we take advantage of the separation of time scales that arises when both the damping and the strength of the polarized current are weak, $\alpha \ll 1$ and $a_J \ll 1$. In this regime, \mathbf{m} moves rapidly along the energy conserving Hamiltonian orbits in Fig. 1(a) and drifts slowly in the direction perpendicular to these orbits. This slow motion can be captured by tracking the evolution of the energy $E(\mathbf{m})$ along with an index to distinguish between disconnected orbits with the same energy. This information is encoded in the graph shown in Fig. 1(b), whose topology is directly related to the energy function, $E = \beta_y m_y^2 + \beta_z m_z^2 - h_x m_x$, and changes based on its form and values of parameters. For example, when $|h_x| < 2\beta_y$ the graph has four branches, as shown in Fig. 1(b), which meet at the saddle point of the energy that corresponds to the homoclinic orbits connecting the two green points on the surface of the sphere in Fig. 1(a). We will use the indexes 1 and 2 (3 and 4) for the lower (higher) energy branches in Fig. 1(b), which correspond to orbits on the front-right and back-left (top and bottom) of the sphere in Fig. 1(a), respectively.

To deduce the effective dynamics on the graph when α and a_J are small, we follow Freidlin and Wentzell [9] to remove the direct dependence on $\mathbf{m}(t)$ from $\dot{E} = \nabla_{\mathbf{m}} E \cdot \dot{\mathbf{m}}$ by time-averaging the coefficients in this equation over the Hamiltonian orbits of constant energy depicted in Fig. 1(a). The resulting averaged equation for the energy on branch j is (see Appx. B)

$$(2) \quad \dot{E} = -\alpha A_j(E) + a_J B_j(E) + 2\alpha\epsilon C_j(E) + \sqrt{2\alpha\epsilon A_j(E)} \xi(t),$$

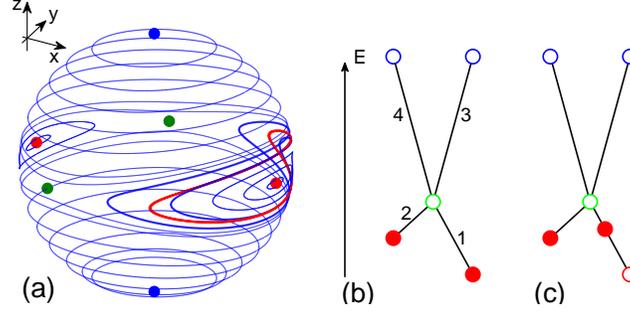


FIGURE 1. (color online) (a): Hamiltonian orbits (blue lines) of the unit magnetization vector solution of $\dot{\mathbf{m}} = -\mathbf{m} \times \nabla_{\mathbf{m}} E$ along with fixed points of this equation that are also critical points of the energy: minima (red dots), saddle points with energy E_b (green dots), and maxima (blue dots). (b) and (c): Graphs (not drawn to scale) in which each energy point along the edges corresponds to an orbit of \mathbf{m} shown in (a) with this energy. The numbers indicate the label for each energy branch. Branch 1: $m_x > -h_x/2\beta_y$ and $E < E_b$, branch 2: $m_x < -h_x/2\beta_y$ and $E < E_b$, branch 3: $m_z > 0$ and $E > E_b$ and branch 4: $m_z < 0$ and $E > E_b$. The ends of each branch correspond to the fixed points in (a) and the circles indicate the location and stability of the fixed points of $\dot{E} = -\alpha A_j(E) + a_j B_j(E)$: filled circles are stable and open circles are unstable. The graph in (b) corresponds to a situation where the energy minima are stable fixed points. The graph in (c) is a representative case when a solution of (4) exists, leading to a new stable fixed point at $E = E_0$; this new fixed point corresponds to a stable limit cycle like the one shown in red in (a).

written in Ito's form, where $\xi(t)$ is a 1D white-noise and

$$\begin{aligned}
 A_j(E) &= 4(\beta_y^2 \langle m_y^2 \rangle_j + \beta_z^2 \langle m_z^2 \rangle_j - E^2) - 4Eh_x \langle m_x \rangle_j + h_x^2(1 - \langle m_x^2 \rangle_j) \\
 B_j(E) &= 2E \langle m_x \rangle_j + h_x(1 + \langle m_x^2 \rangle_j) \\
 C_j(E) &= \beta_y + \beta_z - 3E - 2h_x \langle m_x \rangle_j .
 \end{aligned}
 \tag{3}$$

Here $\langle \cdot \rangle_j$ denotes the time-average over one period along the orbit with constant energy E corresponding to branch j of the graph in Fig. 1(b). As we show in the Appx. C, the averages in (3) can be evaluated asymptotically near the critical points, and this information turns out to be sufficient to calculate the bifurcation diagram and the mean times of magnetization reversal that we obtain below. Away from the critical points, the averages (3) must be evaluated numerically, which we do by using a symplectic implicit mid-point integrator to evolve \mathbf{m} via $\dot{\mathbf{m}} = \mathbf{m} \times \nabla_{\mathbf{m}} E$ along an orbit with prescribed energy to compute the time averages. Note also that (2) requires a matching condition where the branches on the energy graph meet [9]; these conditions are discussed in the Appx. D.

Next we use the reduced equation (2) to obtain the bifurcation diagram of the system at zero temperature, $\epsilon = 0$, and determine the fixed points of $\dot{E} = -\alpha A_j(E) + a_j B_j(E)$ and their stability. The coefficients $A_j(E)$ and $B_j(E)$ encode the separate effects of the damping and the STT on the energy, respectively, and it can be checked that they are both zero at

the critical points of the Hamiltonian. The energies at these points are

$$E_{a,1} = -h_x \quad \text{and} \quad E_{a,2} = h_x$$

corresponding to the two energy minima on the lower branches 1 and 2 where $\mathbf{m} = (\pm 1, 0, 0)$, respectively;

$$E_b = \beta_y + h_x^2/4\beta_y$$

corresponding to the saddle point in energy where all four branches meet where $\mathbf{m} = (-h_x/2\beta_y, \pm\sqrt{1 - h_x^2/4\beta_y^2}, 0)$; and

$$E_c = \beta_z + h_x^2/4\beta_z$$

corresponding to the two energy maxima on the upper branches 3 and 4 where $\mathbf{m} = (-h_x/2\beta_z, 0, \pm\sqrt{1 - h_x^2/4\beta_z^2})$. These critical points can merge and disappear when the applied field crosses the critical values $h_x = \pm 2\beta_y$ and $h_x = \pm 2\beta_z$. In addition, only the two energy minima can ever be stable, and one of them loses stability when another nontrivial fixed point in energy, E_0 , appears on one of the energy branches, where the energy lost by damping, $-\alpha A_j(E_0)$, is exactly compensated by the energy gained by STT, $a_j B_j(E_0)$:

$$(4) \quad -\alpha A_j(E_0) + a_j B_j(E_0) = 0.$$

The stable fixed point at E_0 does not correspond to a stable fixed point of the original dynamics, but rather to a stable limit cycle (precessional state), see Fig. 1 for a schematic illustration. The location and stability of the fixed points identified above are shown in Fig. 2(a) as a function of a_j for a fixed value of h_x , highlighting that magnetization reversal can be achieved by varying the strength of the spin-polarized current, a_j .

We can also calculate the full bifurcation diagram shown in Fig. 2(b), which is remarkably similar to the experimental one (see Fig. 2a in [15]). One of the energy minima loses its stability and the precessional state appears when $E_{a,1}$ or $E_{a,2}$ solves (4), i.e. when a_j is given by ($j = 1, 2$)

$$(5) \quad a_j = \alpha \lim_{x \downarrow E_{a,j}} \frac{A_j(x)}{B_j(x)} = \alpha \sigma_j (\beta_y + \beta_z + \sigma_j h_x),$$

where $\sigma_1 = 1$ and $\sigma_2 = -1$. The corresponding boundaries on the bifurcation diagram are shown as dashed lines in Fig. 2(b). The limit in (5) was obtained using asymptotic expansions of the coefficients; details can be found in the Appx. C.1. The precessional state exists in the region between the dashed and the solid lines in Fig. 2(b). Beyond these solid lines only one stable state remains. This occurs when E_b solves (4), meaning that the strength of the current required to induce switching is ($j = 1, 2$)

$$(6) \quad \begin{aligned} a_j &= \alpha \lim_{x \uparrow E_b} \frac{A_j(x)}{B_j(x)} \equiv \lambda_j \\ &= \alpha \sigma_j \frac{4d_j(\beta_z^2 - \beta_y^2) - b_j(4\beta_y^2 + h_x^2)}{\sigma_j b_j h_x + \pi \sqrt{b_j} \beta_y (1 - h_x^2/4\beta_y^2)} \end{aligned}$$

where $b_j = (\sqrt{S/\beta_z} + \sigma_j(\beta_z - \beta_y)h_x/2\beta_y\beta_x)^2$, $d_j = 1 - (\sqrt{S/\beta_z} - \sigma_j h_x/2\beta_z)^2$, and $S = (\beta_z - \beta_y)(\beta_z - h_x^2/4\beta_y)$; details can be found in the Appx. C.2.

Next we study thermally induced magnetization reversal. To this end we use the reduced system in (2) with $\epsilon > 0$ to calculate the mean transition times (i.e. dwell times) between the

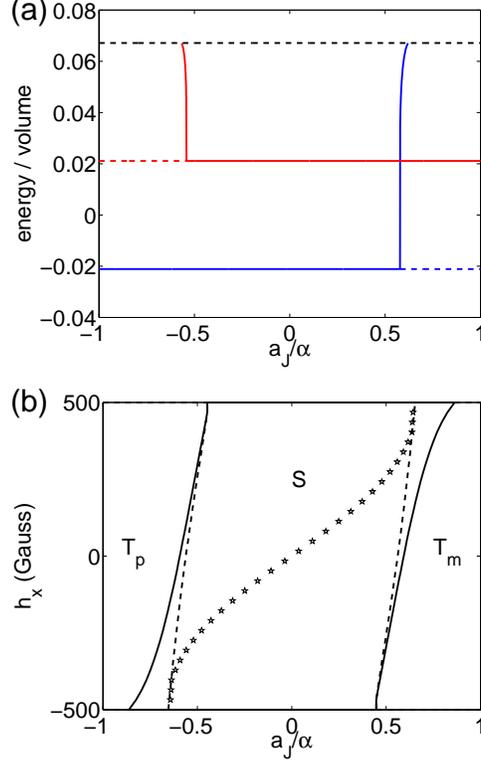


FIGURE 2. (color online) (a): Bifurcation diagram as a function of a_J at the fixed value of $h_x = 80.64$ G: the stable (solid line) and unstable (dashed line) fixed points of $\dot{E} = -\alpha A_j(E) + a_J B_j(E)$ are shown for both energy branch 1 (blue) and 2 (red). (The remaining two unstable fixed point with higher energy are not shown.) (b): Bifurcation diagram as a function of a_J and h_x (converted to units of Gauss by $\mu_0 M_s 10^4$, $M_s = 954930$ A/m, to easily compare with experimental data shown in Fig. 2a in [15]). The dashed lines on either side of region S correspond to the current required to first initiate a stable precessional state (a_J given by (5)) while the solid lines correspond to the current required to induce switching (a_J given by (6)). Beyond the solid lines, only one stable fixed point remains: in region T_p (T_m) it is $m_x = +1$ ($m_x = -1$) and for h_x beyond the region shown, only one lower energy branch remains. The stars indicate where the mean thermally induced switching times from branch 1 to 2 and branch 2 to 1, computed via (7), are equal. For both figures, the values $\alpha = 0.003$, $\beta_y = 0.0654$, and $\beta_z = 0.5$ are used.

stable fixed points of the deterministic dynamics identified before. The mean time $\tau_j(x)$ to transition from energy x on branch $j = 1, 2$ to the fixed point on the other branch satisfies

$$(7) \quad \begin{aligned} & [-\alpha A_j(x) + a_J B_j(x) + 2\alpha \epsilon C_j(x)] \tau_j'(x) \\ & + \alpha \epsilon A_j(x) \tau_j''(x) = -1 \end{aligned}$$

with a matching condition (see Appx. D) to prescribe transitions through the center node of the graphs shown in Figs. 1(b), (c) as well as an absorbing condition at the target state. Equation (7) is valid at any temperature and its solution can be expressed in terms of

integrals involving the coefficients $A_j(x)$, etc. Evaluating these integrals numerically leads to the results shown in Fig. 5. We can also evaluate these integrals asymptotically in the limit when the temperature is small ($\epsilon \ll 1$), in which case they are dominated by the known behavior of the coefficients near the critical points. These calculations are tedious but straightforward and reported in the Appx. E. In situations where the system transits from the stable minimum $E_{a,1}$ or $E_{a,2}$ on branch 1 or 2 to the stable point (minimum $E_{a,2}$ or $E_{a,1}$ or precessional state E_0) on the other branch we obtain ($j = 1, 2$)

$$(8) \quad \tau_j \sim \frac{\gamma_j}{\gamma_1 + \gamma_2} \frac{\epsilon}{\alpha} \frac{e^{(1-\frac{\alpha I}{\lambda_j})(E_b - E_{a,j})/\epsilon}}{2(\beta_y + \beta_z \pm h_x)(1 - \frac{\alpha I}{\lambda_j})^2}$$

whereas in situations when switching occurs from the stable precessional state E_0 we obtain ($j = 1, 2$ depending on whether E_0 is on branch 1 or 2)

$$(9) \quad \tau_j \sim \frac{\gamma_j}{\gamma_1 + \gamma_2} \frac{\epsilon}{\alpha} \frac{e^{(1-\frac{\alpha I}{\lambda_j})(E_b - E_0)/\epsilon}}{A_j(E_0)(1 - \frac{\alpha I}{\lambda_j})^2}.$$

Here λ_j , defined in (6), is the critical current to induce switching at zero temperature, and

$$\begin{aligned} \gamma_1 &= 4d_2(\beta_z^2 - \beta_y^2) - b_2(4\beta_y^2 + h_x^2) \\ \gamma_2 &= 4d_1(\beta_z^2 - \beta_y^2) - b_1(4\beta_y^2 + h_x^2) \end{aligned}$$

with b_j and d_j defined after (6). The results in (8) and (9) agree with the experimental observations [15, 18] and the theoretical predictions [1, 17] that the effect of STT on the dwell times can be captured via a Néel-Brown-type formula with an effective energy scaling linearly with the current strength. We stress, however, that these previous theoretical works had to assume the existence of such a formula, whereas (8) and (9) fall out naturally from the asymptotic analysis, and give explicit expressions not only for the effective energy but also the prefactors and their dependency on the strength of the current producing STT.

In summary, we have shown how the dynamical behavior of nanomagnets driven by spin-polarized currents can be understood in the low-damping regime by mapping their evolution to the diffusion of an energy on a graph. We thereby obtained the full bifurcation diagram of the magnet at zero-temperature as well as the mean times of thermally assisted magnetization reversal. These results agree with experimental observations and give explicit expressions for the dwell times in terms of a Néel-Brown-type formula with an effective energy, thereby settling the issue of the existence of such a formula. We carried the analysis for micromagnets that are of the specific type considered by Li and Zhang [16], but the method presented in this paper is general and can be applied to other situations with different geometry, applied fields that are time-dependent or not, etc. It can also be applied to systems in which the magnetization varies spatially in the sample. In these situations, the graph of the energy will be more complicated, but the general procedure to reduce the dynamics to a diffusion on this graph remains the same. Such a study will be the object of a future publication.

The research of E. V.-E. was supported in part by NSF grant DMS07-08140 and ONR grant N00014-11-1-0345.

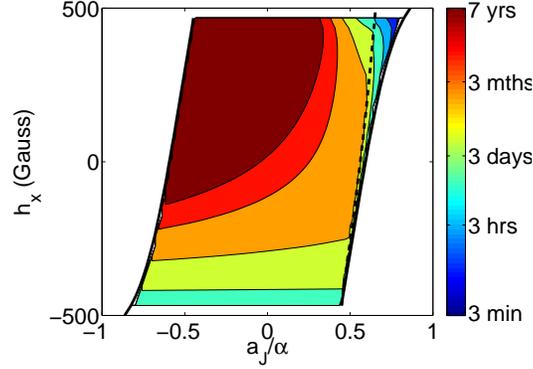


FIGURE 3. (color online) Contour plot of the mean first passage time (converted to seconds by $\gamma\mu_0 M_s/(1 + \alpha^2)$, $\gamma = 2.21 \times 10^5$ m/A·s, $M_s = 954930$ A/m) from the fixed point on energy branch $j = 1$, to the fixed point on branch 2. (The times from branch 2 to 1 would be the same figure, rotated 180° .) These times are computed using (7) and plotted as a function of h_x (converted to units of Gauss by $\mu_0 M_s 10^4$) and a_J . The shape of the plotted region is identical to that shown in Fig. 2(b). Between the dashed line (a_J in (5)) and solid black line (a_J in (6)), the starting point is E_0 (Eq. (4)); everywhere else it is the minimum energy, $E_{a,1}$. The values $\alpha = 0.003$, $\epsilon = 0.01$, $\beta_y = 0.654$, and $\beta_z = 0.5$ are used. Switching times greater than 25 years are all colored in dark red.

APPENDIX A. NONDIMENSIONALIZATION OF THE GOVERNING EQUATION

We focus on magnetic systems of the type considered in Li and Zhang [16]. They consist of a thicker fixed-magnetization layer, separated from a thinner layer whose magnetization we wish to model. This free-layer's magnetization vector $\mathbf{M} = (M_x, M_y, M_z)$, assumed to have constant strength $|\mathbf{M}| = M_s$ (M_s is the saturation magnetization), is governed by the (dimensional) stochastic Landau-Lifshitz-Gilbert (LLG) equation

$$(10) \quad \frac{d\mathbf{M}}{dt} = -\gamma^* \mathbf{M} \times \mathbf{H}_{\text{eff}} - \frac{\gamma^* \alpha}{M_s} \mathbf{M} \times (\mathbf{M} \times \mathbf{H}_{\text{eff}})$$

where $\gamma^* = \gamma/(1 + \alpha^2)$, γ is the gyromagnetic ratio (units rad/s·T) and α is the non-dimensional Gilbert damping constant. The effective magnetic field (units T),

$$(11) \quad \mathbf{H}_{\text{eff}} = -\nabla_{\mathbf{M}} \tilde{E} + \sqrt{\frac{2\alpha k_B T}{\gamma M_s \nu}} \boldsymbol{\eta}(\tilde{t}),$$

is the sum of the 3-component white noise, $\boldsymbol{\eta}(\tilde{t})$, and the gradient of \tilde{E} , the energy per unit volume (units N·m/m³). The noise amplitude, derived from the fluctuation-dissipation theorem, will be discussed later, ν is the volume of the magnetic element, its temperature is T , and the Boltzmann constant is $k_B = 1.38 \times 10^{-23}$ N·m/K. The energy per unit volume varies across different materials. Here, we use

$$(12) \quad \tilde{E} = \mu_0 \mathbf{H}_{\text{ext}} \cdot \mathbf{M} + \frac{\mu_0 H_k}{2M_s} M_y^2 + \frac{\mu_0}{2} M_z^2,$$

the sum of the contributions from the external field, \mathbf{H}_{ext} , and the anisotropy in the magnet along the y direction, given by H_k (hard axis z , easy axis x). The permeability of free space is $\mu_0 = 4\pi \times 10^{-7}$ T·m/A.

In order to non-dimensionalize Eq. (10), we first write $\mathbf{M} = M_s \mathbf{m}$ where $\mathbf{m} = (m_x, m_y, m_z)$ is a unit vector, $\tilde{E} = \mu_0 M_s^2 E$ where E is the non-dimensional energy per volume, and $\mathbf{H}_{eff} = \mu_0 M_s \mathbf{h}_{eff}$. With these substitutions we obtain

$$(13) \quad \begin{aligned} \frac{d\mathbf{m}}{d\tilde{t}} &= -\gamma^* \mu_0 M_s \mathbf{m} \times \mathbf{h}_{eff} - \gamma^* \mu_0 M_s \alpha \mathbf{m} \times (\mathbf{m} \times \mathbf{h}_{eff}) \\ \mathbf{h}_{eff} &= -\nabla_{\mathbf{m}} E + \sqrt{\frac{2\alpha k_B T}{\gamma \mu_0 M_s \mu_0 M_s^2 \nu}} \boldsymbol{\eta}(\tilde{t}) \\ E &= \frac{\mathbf{H}_{ext}}{M_s} \cdot \mathbf{m} + \frac{1}{2} \frac{H_k}{M_s} m_y^2 + \frac{1}{2} m_z^2. \end{aligned}$$

After non-dimensionalizing time to $t = \tilde{t}/\gamma^* \mu_0 M_s$, the resulting non-dimensional equation is

$$(14) \quad \begin{aligned} \frac{d\mathbf{m}}{dt} &= -\mathbf{m} \times \mathbf{h}_{eff} - \alpha \mathbf{m} \times (\mathbf{m} \times \mathbf{h}_{eff}) \\ \mathbf{h}_{eff} &= -\nabla_{\mathbf{m}} E + \sqrt{\frac{2\alpha k_B T}{(1 + \alpha^2) \mu_0 M_s^2 \nu}} \boldsymbol{\eta}(t). \end{aligned}$$

Note that the appearance of $(1 + \alpha^2)$ was because γ , not γ^* appeared in the denominator of the white-noise amplitude factor.

To Eq. (14), we add the spin-transfer torque (STT) term, $a_J \mathbf{m} \times (\mathbf{m} \times \mathbf{m}_p)$, describing the response of the system to a spin-polarized current with (non-dimensional) strength a_J generated from the fixed magnetic layer, with magnetization unit vector \mathbf{m}_p . The non-dimensional strength $a_J = \gamma^* \mu_0 \eta(\theta) \mu_B I / e \nu$, contains the dependence on the current, I , and the structure of the fixed and free magnetic layers through $\eta(\theta) = q / (A + B \cos(\theta))$ where $\cos(\theta) = \mathbf{m} \cdot \mathbf{m}_p$ [21]. The current strength is divided by the volume, ν , of the free layer, so as to produce a force per volume, matching the energy per volume contribution already in Eq. (14). We take the value of a_J to be constant here, though our results could easily be extended to the time-varying case.

For simplicity, we consider both the spin-polarized current direction, \mathbf{m}_p , and the external field, \mathbf{H}_{ext} , to be directed along the x -axis: $\mathbf{m}_p = (1, 0, 0)$ and $\mathbf{H}_{ext} = (h_x M_s, 0, 0)$. The final non-dimensional LLG equation that appears as Eq. (1) in the text is

$$(15) \quad \frac{d\mathbf{m}}{dt} = -\mathbf{m} \times \mathbf{h}_{eff} - \mathbf{m} \times (\mathbf{m} \times (\alpha \mathbf{h}_{eff} - a_J \mathbf{m}_p))$$

where $\mathbf{h}_{eff} = -\nabla_{\mathbf{m}} E + \sqrt{\frac{2\alpha\epsilon}{1+\alpha^2}} \boldsymbol{\eta}(t)$ with $\epsilon = k_B T / \mu_0 M_s^2 \nu$ and

$$(16) \quad E = \beta_y m_y^2 + \beta_z m_z^2 - h_x m_x$$

and where $\beta_y = H_k / 2M_s < \beta_z = 1/2$. Recall that M_s is the constant amplitude of the dimensional magnetization vector and ν the volume of the magnetic element.

We return to consider the amplitude of the noise, $\sqrt{2\alpha\epsilon/(1+\alpha^2)}$ in Eq. (14), taken from the paper by Kohn, Reznikoff and Vanden-Eijnden [13]. The parameter ϵ is the non-dimensional temperature; it is the thermal energy, $k_B T$, non-dimensionalized by $\mu_0 M_s^2 \nu$,

where ν is the volume of the magnet and $\mu_0 M_s^2$ is the energy per volume scaling. The α appears in the noise term so that the equilibrium distribution is the Gibbs distribution,

$$\rho(m) = Z^{-1} e^{-E(m)/\epsilon},$$

where Z^{-1} is the normalization factor; it is the integral of $e^{-E(m)/\epsilon}$ over the surface of the m -sphere. This fact is shown in Appendix A of the paper by Kohn, Reznikoff and Vanden-Eijnden [13]. It was done by switching to the (θ, z) representation of the magnetic vector to represent this two degrees of freedom system by two variables. From the non-dimensional form of the noise term, we worked backwards through the non-dimensionalization procedure to obtain the dimensional noise amplitude in Eq. (10).

A.1. Converting to Ito form. We convert the Strotonovich SDE (15) into an Ito SDE, which will be needed in the next section to determine the averaged evolution of the energy. First, we write the Strotonovich SDE in the form

$$(17) \quad \frac{d\mathbf{m}}{dt} = \mathbf{a}_c - \alpha \mathbf{a}_d + a_J \mathbf{a}_p + \sqrt{\frac{2\alpha\epsilon}{1+\alpha^2}} \mathcal{B}\boldsymbol{\eta}(t)$$

where the conservative term is

$$\mathbf{a}_c = \begin{pmatrix} 2(\beta_z - \beta_y)m_y m_z \\ -m_z(2\beta_z m_x + h_x) \\ m_y(2\beta_y m_x + h_x) \end{pmatrix}$$

the damping term is

$$\mathbf{a}_d = \begin{pmatrix} -h_x(1 - m_x^2) - 2m_x(\beta_y m_y^2 + \beta_z m_z^2) \\ m_y(h_x m_x + 2\beta_y(1 - m_y^2) - 2\beta_z m_z^2) \\ m_z(h_x m_x - 2\beta_y m_y^2 + 2\beta_z(1 - m_z^2)) \end{pmatrix}$$

the spin-torque transfer term is

$$\mathbf{a}_p = \begin{pmatrix} m_x^2 - 1 \\ m_x m_y \\ m_x m_z \end{pmatrix}$$

and the diffusion matrix is

$$\mathcal{B} = \begin{pmatrix} \alpha(1 - m_x^2) & m_z - \alpha m_x m_y & -m_y - \alpha m_x m_z \\ -m_z - \alpha m_x m_y & \alpha(1 - m_y^2) & m_x - \alpha m_y m_z \\ m_y - \alpha m_x m_z & -m_x - \alpha m_y m_z & \alpha(1 - m_z^2) \end{pmatrix}.$$

In order to convert to the Ito form, the drift term obtains the correction

$$\mathbf{a}_I = \frac{2\alpha\epsilon}{1+\alpha^2} \frac{1}{2} \sum_{j,k} \mathcal{B}_{kj} \partial_k \mathcal{B}_{ij} = -2\alpha\epsilon \mathbf{m},$$

making the Ito SDE for the magnetization direction

$$(18) \quad \frac{d\mathbf{m}}{dt} = \mathbf{a}_c - \alpha \mathbf{a}_d + a_J \mathbf{a}_p - 2\alpha\epsilon \mathbf{m} + \sqrt{\frac{2\alpha\epsilon}{1+\alpha^2}} \mathcal{B}\boldsymbol{\eta}(t).$$

APPENDIX B. ENERGY EQUATION

Here, we derive the energy equation (2) in the text. Using the rules of Ito calculus, we compute $\dot{E} = \nabla_{\mathbf{m}} E \cdot \dot{\mathbf{m}}$ to obtain the Ito SDE for the evolution of the energy in the form

$$(19) \quad \dot{E} = -\alpha A(\mathbf{m}) + a_j B(\mathbf{m}) + 2\alpha\epsilon C(\mathbf{m}) + \sqrt{2\alpha\epsilon} \sqrt{A(\mathbf{m})} \xi(t),$$

where $\xi(t)$ is 1D white noise, the strength of which is computed from the combination of the three strengths of the three independent components of the white noise, $\boldsymbol{\eta}(t)$, in Eq. (15). Notice since the term \mathbf{a}_c in Eq. (18) conserves energy, it has no corresponding term in Eq. (19). The remaining terms in Eq. (18) have corresponding terms in Eq. (19): the dissipative term, \mathbf{a}_d , leads to

$$A(\mathbf{m}) = 4(\beta_y^2 m_y^2 + \beta_z^2 m_z^2 - E^2) - 4E h_x m_x + h_x^2 (1 - m_x^2),$$

the spin-torque transfer terms, \mathbf{a}_p , leads to

$$B(\mathbf{m}) = 2E m_x + h_x (1 + m_x^2),$$

and the correction term for Ito calculus (∂_j indicates partial derivative with respect to the j th element of \mathbf{m}),

$$\frac{2\alpha\epsilon}{1 + \alpha^2} \frac{1}{2} \sum_{i,j} [\mathcal{B}\mathcal{B}^T]_{ij} \partial_i \partial_j E(\mathbf{m}) = 2\alpha\epsilon \left(\beta_y (1 - m_y^2) + \beta_z (1 - m_z^2) \right),$$

together with the contribution from \mathbf{a}_I give

$$C(\mathbf{m}) = \beta_y + \beta_z - 3E - 2h_x m_x.$$

The three separate contributions to the noise term,

$$\begin{aligned} b_x(\mathbf{m}) &= 2(\beta_z - \beta_y) m_y m_z - \alpha [h_x (1 - m_x^2) + 2m_x (\beta_y m_y^2 + \beta_z m_z^2)] \\ b_y(\mathbf{m}) &= -m_z (h_x + 2\beta_z m_x) + \alpha [h_x m_x m_y + 2m_y (\beta_y (1 - m_y^2) - \beta_z m_z^2)] \\ b_z(\mathbf{m}) &= h_x m_y + 2\beta_y m_x m_y + \alpha [h_x m_x m_z + 2m_z (\beta_z (1 - m_z^2) - \beta_y m_y^2)], \end{aligned}$$

simplify so that $\sqrt{b_x^2 + b_y^2 + b_z^2} = \sqrt{(1 + \alpha^2) A(\mathbf{m})}$. To remove the explicit dependence on the magnetization vector \mathbf{m} from Eq. (19), following Freidlin and Wentzell [9], we average the coefficients appearing in the backwards Kolmogorov equation for SDE (19) over one period, $T_j(E)$, at constant energy,

$$\langle f(\mathbf{m}) \rangle_j = \frac{1}{T_j(E)} \int_0^{T_j(E)} f(\mathbf{m}(t)) dt.$$

The subscript $j = 1, 2, 3, 4$ indicates that the average corresponds to one connected orbit of \mathbf{m} with constant energy E on branch j of the energy graph (Fig. 1(b) in the text). The resulting averaged coefficient backwards Kolmogorov equation corresponds to the averaged coefficient SDE (2) in the text.

APPENDIX C. ASYMPTOTIC EXPANSIONS OF THE AVERAGED ENERGY COEFFICIENTS

While explicit expressions for the averaged coefficients are unknown, we work out asymptotic expansions near the two minima and the saddle point of the energy. These expressions show how the coefficients $A_j(E)$ and $B_j(E)$ go to zero near these points, and are used to determine the accessibility of these critical points in energy. The expressions are also used to determine the critical spin polarized current a_j for the emergence of a stable precessional

state (Eq. (5) in the text), and for the critical current to induce switching of the magnetization vector (Eq. (6) in the text). In addition, they are used in Sec. D in determining the switching probabilities between the lower energy branches and in Sec. E in asymptotically approximating the mean switching time.

C.1. Approximation Near the Minima. In order to determine the scaling of the averaged coefficients $A_1(E)$, $B_1(E)$ and $C_1(E)$, near the energy minimum on branch 1, $\mathbf{m}_0 = (1, 0, 0)$, we create a series expansion about this point for the solution to the Hamiltonian system,

$$(20) \quad \dot{\mathbf{m}} = -\mathbf{m} \times \nabla E = \begin{pmatrix} 2(\beta_z - \beta_y)m_y m_z \\ -m_z(2\beta_z m_x + h_x) \\ m_y(2\beta_y m_x + h_x) \end{pmatrix},$$

then compute the averages exactly as a function of the energy. We consider the expansion for \mathbf{m} ,

$$\mathbf{m} = (1, 0, 0) + \delta \mathbf{m}_1 + \delta^2 \mathbf{m}_2 + O(\delta^3),$$

(where the subscripts denote terms in the expansion, and not energy branches) and determine the averaged coefficients up to $O(\delta^2)$, requiring the first non-zero term above the zeroth term for each component of \mathbf{m} . This expansion must also satisfy $|\mathbf{m}|^2 = 1$ up to order δ^2 ,

$$\begin{aligned} |\mathbf{m}^2| &= (1 + \delta m_{1,x} + \delta^2 m_{2,x})^2 + \delta^2 m_{1,y}^2 + \delta^2 m_{1,z}^2 + O(\delta^3) \\ &= 1 + 2\delta m_{1,x} + \delta^2 (m_{1,x}^2 + 2m_{2,x} + m_{1,y}^2 + m_{1,z}^2) + O(\delta^3). \end{aligned}$$

To satisfy this constraint at $O(\delta)$, we must have that $m_{1,x} = 0$, and then at $O(\delta^2)$,

$$(21) \quad m_{2,x} = -\frac{1}{2}(m_{1,y}^2 + m_{1,z}^2).$$

The $O(\delta)$ expansion of the system (20) is

$$(22) \quad \dot{\mathbf{m}}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2\beta_z - h_x \\ 0 & 2\beta_y + h_x & 0 \end{pmatrix} \mathbf{m}_1.$$

This is consistent with $m_{1,x} = 0$ obtained from the length constraint on \mathbf{m} above. The remaining two components are

$$\begin{aligned} m_{1,y} &= c_1 \cos \omega t \\ m_{1,z} &= c_1 \frac{\omega}{2\beta_z + h_x} \sin \omega t \end{aligned}$$

where $\omega = \sqrt{(2\beta_y + h_x)(2\beta_z + h_x)}$ and the $O(\delta)$ initial condition is assumed to be of the form $\mathbf{m}_1(0) = (0, c_1, 0)$. No higher order terms are needed for the y and z components, but we must consider the next order term for the x component.

For the component $m_{2,x}$ at $O(\delta^2)$, the expansion of system (20) gives

$$\dot{m}_{2,x} = 2(\beta_z - \beta_y)m_{1,y}m_{1,z}.$$

The solution to this is consistent with the definition of $m_{2,x}$ in (21) need to satisfy the length constraint. By taking the derivative of $m_{2,x}$ given by (21),

$$\dot{m}_{2,x} = -m_{1,y}\dot{m}_{1,y} - m_{1,z}\dot{m}_{1,z},$$

and then substituting the $O(\delta)$ DEQ for $\dot{m}_{1,y}$ and $\dot{m}_{1,z}$, we see that

$$\dot{m}_{2,x} = 2(\beta_z - \beta_y)m_{1,y}m_{1,z}.$$

Combining these results we have the expansions

$$\begin{aligned} m_x &\sim 1 - \delta^2 \frac{c_1^2}{2} \left[\cos^2 \omega t + \frac{2\beta_y + h_x}{2\beta_z + h_x} \sin^2 \omega t \right] \\ m_y &\sim \delta c_1 \cos \omega t \\ m_z &\sim \delta c_1 \sqrt{\frac{2\beta_y + h_x}{2\beta_z + h_x}} \sin \omega t \end{aligned}$$

where $\omega = \sqrt{(2\beta_y + h_x)(2\beta_z + h_x)}$. These solutions correspond to a trajectory with initial condition $(1 - \delta^2 c_1^2/2, \delta c_1, 0)$ and constant energy

$$(23) \quad E = -h_x + \delta^2 c_1^2 (\beta_y + h_x/2).$$

To determine the averaged coefficients, we average the functions of \mathbf{m} over one period, $T = 2\pi/\omega$ (note that T does not depend on the energy in this expansion),

$$\langle f(\mathbf{m}) \rangle = \frac{1}{T} \int_0^T f(\mathbf{m}(t)) dt.$$

Taking $c_1 = 1$ (without loss of generality), we have

$$\begin{aligned} \langle m_x \rangle &\sim 1 - \frac{\delta^2}{4} \left(1 + \frac{2\beta_y + h_x}{2\beta_z + h_x} \right) \\ \langle m_x^2 \rangle &\sim 1 - \frac{\delta^2}{2} \left(1 + \frac{2\beta_y + h_x}{2\beta_z + h_x} \right) \\ \langle m_y^2 \rangle &\sim \frac{\delta^2}{2} \\ \langle m_z^2 \rangle &\sim \frac{\delta^2}{2} \left(\frac{2\beta_y + h_x}{2\beta_z + h_x} \right). \end{aligned}$$

To write the averages in term of the energy, we solve for δ^2 as a function of E from Eq. (23),

$$\frac{\delta^2}{2} = \frac{E + h_x}{2\beta_y + h_x},$$

and obtain

$$(24) \quad \begin{aligned} A_1(E) &\sim 2(\beta_y + \beta_z + h_x)(E + h_x) \\ B_1(E) &\sim 2(E + h_x) \\ C_1(E) &\sim \beta_y + \beta_z + h_x - \left(1 + \frac{2\beta_y}{2\beta_y + h_x} + \frac{2\beta_z}{2\beta_z + h_x} \right) (E + h_x). \end{aligned}$$

These approximations are in good agreement with the numerically obtained averaged coefficients, shown in Fig. 4. The limit in Eq. (5) in the text is obtained from the ratio of $A_1(E)$ to $B_1(E)$.

We repeat the above procedure for energy branch 2, where the minimum is $\mathbf{m} = (-1, 0, 0)$, using the series expansion

$$\mathbf{m} = (-1, 0, 0) + \delta \mathbf{m}_1 + \delta^2 \mathbf{m}_2 + O(\delta^3).$$

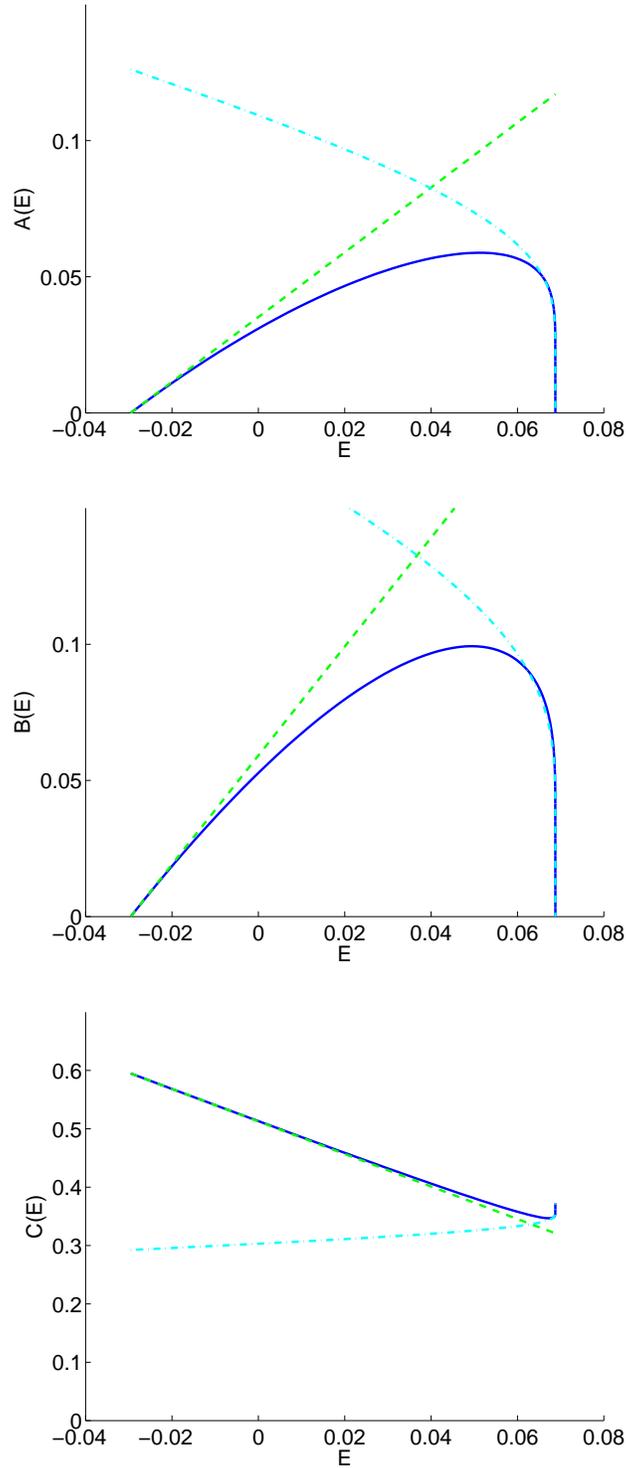


FIGURE 4. Comparison of the averaged energy equation coefficients on energy branch 1 computed by averaging the numerically integrated trajectories given by $\dot{\mathbf{m}} = -\mathbf{m} \times \nabla E$ (blue solid line) to the asymptotic approximations in Eq. (24) near the minimum energy (green dashed line) and the asymptotic approximations in Eq. (36) near the saddle point in energy (cyan dash-dot line) for non-zero h_x . The values $\beta_y = 0.06$, $\beta_z = 0.5$ and $h_x = 0.03$ were used.

From the constraint on the length of \mathbf{m} we similarly find that $m_{1,x} = 0$ and

$$m_{2,x} = \frac{1}{2}(m_{1,y}^2 + m_{1,z}^2).$$

The $O(\delta)$ expansion of the Hamiltonian system has the solution

$$\begin{aligned} m_{1,y} &= \cos \omega t \\ m_{1,z} &= -\sqrt{\frac{2\beta_y - h_x}{2\beta_z - h_x}} \sin \omega t, \end{aligned}$$

where $\omega = \sqrt{(2\beta_y - h_x)(2\beta_z - h_x)}$ and we have taken an $O(\delta)$ initial condition $(0,1,0)$. Combining and averaging as above, using

$$\frac{\delta^2}{2} = \frac{E - h_x}{2\beta_y - h_x},$$

we obtain

$$\begin{aligned} (25) \quad A_2(E) &\sim 2(\beta_y + \beta_z - h_x)(E - h_x) \\ B_2(E) &\sim -2(E - h_x) \\ C_2(E) &\sim \beta_y + \beta_z - h_x - \left(1 + \frac{2\beta_y}{2\beta_y - h_x} + \frac{2\beta_z}{2\beta_z - h_x}\right)(E - h_x) \end{aligned}$$

for the averaged energy coefficients on energy branch 2. The limit in Eq. (5) in the text is obtained from the ratio of $A_2(E)$ to $B_2(E)$.

C.2. Approximation Near the Saddle Point. We repeat a similar process to the one in the above section, and determine the scaling of the period of the orbit, which goes to infinity as the energy approaches its saddle point value, by again using a series expansion of the solution. To determine the approximation of the averaged coefficients, we must also consider their value along the entire orbit in order to compute the time-average integral. The integral is dominated by the values the components of \mathbf{m} take along the homoclinic orbit connecting the two fixed points. We consider this integral last.

Orbits with either energy slightly less than or slightly more than the saddle point value, E_b , spend the majority of their time near the two fixed points. To approximate the period, we expand the solution of the Hamiltonian system, Eq. (20), about one fixed point and determine the time it takes the trajectory to leave an $O(1)$ size box around this fixed point. We must handle the two cases, $h_x = 0$ and $h_x \neq 0$, separately, but in the end we obtain the same scaling of the period with the energy in both cases.

First, we consider the simple case when $h_x = 0$, and consider the expansion of the solution about one of the fixed points, $\mathbf{m}_0 = (0, 1, 0)$, in the form

$$\mathbf{m} = (0, 1, 0) + \delta \mathbf{m}_1 + \delta^2 \mathbf{m}_2 + O(\delta^3).$$

From the constraint on the length of the vector,

$$1 = |\mathbf{m}|^2 = 1 + 2\delta m_{1,y} + \delta^2(2m_{2,y} + m_{1,x}^2 + m_{1,y}^2 + m_{1,z}^2) + O(\delta^3),$$

we see that at $O(\delta)$ it must be that $m_{1,y} = 0$, and at $O(\delta^2)$,

$$(26) \quad m_{2,y} = -\frac{1}{2}(m_{1,x}^2 + m_{1,z}^2).$$

The constant energy of this expansion can be found in terms of the initial conditions, and is

$$E = \beta_y + 2\beta_y \delta^2 m_{2,y}(0) + \beta_z \delta^2 m_{1,z}^2(0) + O(\delta^3).$$

Due to the constraint, we only need to find the $O(\delta)$ solutions, then construct $m_{2,y}$ from Eq. (26).

The expansion of the Hamiltonian system (20) leads to the $O(\delta)$ set of equations,

$$\begin{aligned}\dot{m}_{1,x} &= 2(\beta_z - \beta_y)m_{1,z} \\ \dot{m}_{1,y} &= 0 \\ \dot{m}_{1,z} &= 2\beta_y m_{1,x},\end{aligned}$$

which is solved by

$$(27) \quad \begin{aligned}m_{1,x} &= \cosh \omega t \\ m_{1,z} &= \frac{\omega}{2(\beta_z - \beta_y)} \sinh \omega t\end{aligned}$$

where $\omega^2 = 4\beta_y(\beta_z - \beta_y)$. We have used the initial conditions $(\delta, 1 - \frac{1}{2}\delta^2, 0)$ and find the energy scales with δ as,

$$E = \beta_y - \beta_y\delta^2 + O(\delta^3).$$

The approximate solutions (27) do not lead to complete orbits as in the previous section, rather, we can estimate the period by computing the time for the trajectory to leave a box around the saddle point. Specifically, we approximate the period as quadruple the time it takes the m_x component to increase to some $O(1)$ value c away from its initial value. (We would obtain the same scaling if we chose to use the m_z component instead.) Note that we are exploiting the symmetry of the system: there are two saddle points which the trajectory spends equal time near, and the trajectory spends an equal amount of time entering the box and traveling to the initial condition as it does leaving the box. Therefore, we wish to solve

$$\delta \cosh\left(\omega \frac{T}{4}\right) = c$$

for the period T , of the orbit, which we expect to be large. We approximate $\cosh \omega T/4 = (e^{\omega T/4} + e^{-\omega T/4})/2$ as $e^{\omega T/4}/2$ and obtain

$$(28) \quad T = \frac{2}{\omega} \log \frac{1}{\delta^2} + O(1) = \frac{2}{\omega} \log \frac{1}{E_b - E} + O(1),$$

where $\omega = 2\sqrt{\beta_y(\beta_z - \beta_y)}$ and the saddle point energy, $E_b = \beta_y$.

We repeat the above procedure for the second case, $h_x \neq 0$. Again, we expand the solution about the fixed point, $(-a, \sqrt{1 - a^2}, 0)$ where $a = h_x/2\beta_y$, in the form

$$\mathbf{m} = (-a, \sqrt{1 - a^2}, 0) + \delta \mathbf{m}_1 + \delta^2 \mathbf{m}_2 + O(\delta^3).$$

This time, the constraint on the length leads to

$$\begin{aligned}1 = |\mathbf{m}|^2 &= 1 + \delta(-2am_{1,x} + 2\sqrt{1 - a^2}m_{1,y}) \\ &+ \delta^2(-2am_{2,x} + 2\sqrt{1 - a^2}m_{2,y} + m_{1,x}^2 + m_{1,y}^2 + m_{1,z}^2) + O(\delta^3)\end{aligned}$$

providing the requirements

$$(29) \quad \begin{aligned}m_{1,y} &= \frac{a}{\sqrt{1 - a^2}} m_{1,x} \\ \sqrt{1 - a^2} m_{2,y} &= am_{2,x} - \frac{1}{2}(m_{1,x}^2 + m_{1,y}^2 + m_{1,z}^2).\end{aligned}$$

Although the $O(\delta)$ solution of m_y is no longer zero, these $O(\delta)$ terms do not contribute to the value of the energy, and the solution to $m_{2,x}$ is required. The constant energy is again obtained from the initial conditions of the solution, and is

$$E = \beta_y \left(\sqrt{1 - a^2} + \delta m_{1,y}(0) + \delta^2 m_{2,y}(0) \right)^2 + \beta_z \delta^2 m_{1,z}^2 - h_x (-a + \delta m_{1,x}(0) + \delta^2 m_{2,x}(0)) + O(\delta^3).$$

We see that the $m_{2,z}$ solution is not needed; it does not contribute to the $O(\delta^2)$ expansion of the energy.

The expansion of the Hamiltonian system (20) leads to the $O(\delta)$ set of equations,

$$\begin{aligned} \dot{m}_{1,x} &= 2(\beta_z - \beta_y) \sqrt{1 - a^2} m_{1,z} \\ \dot{m}_{1,y} &= -(-2a\beta_z + h_x) m_{1,z} \\ \dot{m}_{1,z} &= 2\beta_y \sqrt{1 - a^2} m_{1,x}, \end{aligned}$$

which are satisfied by

$$\begin{aligned} m_{1,x} &= \cosh \omega t \\ m_{1,z} &= \frac{\omega}{2(\beta_z - \beta_y) \sqrt{1 - a^2}} \sinh \omega t \end{aligned}$$

where $\omega^2 = 4\beta_y(\beta_z - \beta_y)(1 - a^2)$. Note these solutions reduce to those in (27) when $h_x = 0$ since $a = 0$ in this case. The $O(\delta^2)$ expansion of the Hamiltonian system is,

$$\begin{aligned} \dot{m}_{2,x} &= 2(\beta_z - \beta_y) \sqrt{1 - a^2} m_{2,z} + 2(\beta_z - \beta_y) m_{1,y} m_{1,z} \\ \dot{m}_{2,y} &= -(2a\beta_z + h_x) m_{2,z} \\ \dot{m}_{2,z} &= 2\beta_y \sqrt{1 - a^2} m_{2,x} + 2\beta_y m_{1,x} m_{1,y} \end{aligned}$$

which have the same general solution as the $O(\delta)$ equations, but an additional particular solution to handle the non-homogeneous part. We only require the solution $m_{2,x}$, since we can then construct the needed $m_{2,y}$ solution from the length constraint (29). Taking another derivative of the $\dot{m}_{2,x}$ equation, and inserting the equation for $\dot{m}_{2,z}$, we use the method of undetermined coefficients, and obtain the solution

$$m_{2,x} = \cosh \omega t + \frac{h_x}{4\omega^2} [(\beta_z - \beta_y) \cosh 2\omega t + \frac{1}{2\omega} \sinh 2\omega t] + \frac{h_x}{2} (\beta_z - \beta_y) t^2.$$

The initial conditions at $O(\delta^2)$ are

$$\begin{aligned} m_{2,x}(0) &= 1 + \frac{h_x}{4\omega^2} (\beta_z - \beta_y) \\ \sqrt{1 - a^2} m_{2,y}(0) &= a m_{2,x}(0) - \frac{1}{2(1 - a^2)}. \end{aligned}$$

Using these, we can construct the expansion of the energy, which after much simplification reduces to

$$(30) \quad E = E_b - \delta^2 \beta_y + O(\delta^3)$$

where the saddle point energy, $E_b = \beta_y + \frac{h_x^2}{4\beta_y}$.

The leading order term of the m_x component is again used to determine the scaling of the period,

$$(31) \quad T = \frac{2}{\omega} \log \frac{1}{\delta^2} + O(1) = \frac{2}{\omega} \log \frac{1}{E_b - E} + O(1)$$

where $\omega^2 = 4\beta_y(\beta_z - \beta_y)(1 - h_x^2/4\beta_y^2)$ and $E_b = \beta_y + \frac{h_x^2}{4\beta_y}$. This reduces to the period in (28) when $h_x = 0$. For all values of h_x we can use (30) and (31) to describe the energy and period near the saddle point in energy, E_b .

To finish the approximation of the averaged energy coefficients, we derive the pre-factors for their $1/T(E)$ scalings. Recall that averages are defined as

$$\langle f(\mathbf{m}(t)) \rangle = \frac{1}{T(E)} \int_0^{T(E)} f(\mathbf{m}(t)) dt$$

for trajectories confined to energy E . (Note we will continue to omit the subscripts denoting energy branches as we only concern ourselves with branch 1 here.) Having already determined the scaling of T as $E \rightarrow E_b$, we now turn to the integral $\int_0^{T(E)} f(\mathbf{m}(t)) dt$, which is dominated by what happens on the homoclinic orbit with energy E_b . For the trajectories of the components of $\mathbf{m}(t)$, we use an approximate trajectory that starts at $t = 0$ at the point on the orbit midway between the two fixed points, where m_z is positive and $m_y = 0$. The trajectories are infinitely long and asymptotically approach the fixed points, therefore we approximate

$$\int_0^{T(E)} f(\mathbf{m}(t)) dt \sim 4 \int_0^\infty f(\mathbf{m}(t)) dt$$

and we take the averages to be approximated by

$$(32) \quad \langle f(\mathbf{m}(t)) \rangle \sim \frac{4}{T(E)} \int_0^\infty f(\mathbf{m}(t)) dt.$$

For the simple case when $h_x = 0$, the exact solution to the Hamiltonian system is

$$(33) \quad \begin{aligned} m_x(t) &= \sqrt{\frac{\beta_z - \beta_y}{\beta_z}} \operatorname{sech} \nu_0 t \\ m_y(t) &= \tanh \nu_0 t \\ m_z(t) &= \sqrt{\frac{\beta_y}{\beta_z}} \operatorname{sech} \nu_0 t \end{aligned}$$

where $\nu_0 = 2\sqrt{\beta_y(\beta_z - \beta_y)}$. Note, this also satisfies $E = \beta_y = \beta_y m_y^2 + \beta_z m_z^2$ and $m_x^2 + m_y^2 + m_z^2 = 1$ due to the identity $\tanh^2 x + \operatorname{sech}^2 x = 1$. Using (32) to compute averages, we have

$$\begin{aligned} \langle m_x^2 \rangle &\sim \frac{\beta_z - \beta_y}{\beta_z} \frac{4}{\log 1/\delta} \\ \langle m_z^2 \rangle &\sim \frac{\beta_y}{\beta_z} \frac{4}{\log 1/\delta} \\ \langle m_y^2 \rangle &= 1 - \langle m_x^2 \rangle - \langle m_z^2 \rangle \sim 1 - \frac{4}{\log 1/\delta} \end{aligned}$$

and the average energy coefficients are

$$(34) \quad \begin{aligned} A_1(E) &\sim \beta_y(\beta_z - \beta_y) \frac{16}{\log 1/\delta} \\ B_1(E) &\sim \frac{\beta_y(\beta_z - \beta_y)}{\beta_z} \frac{4\pi}{\log 1/\delta} \\ C_1(E) &\sim \beta_z - 2\beta_y. \end{aligned}$$

For non-zero h_x , an exact solution is unknown, but the components m_x and m_z are well approximated by sech functions, with appropriate values to match the exact solution at $t = 0$ and as $t \rightarrow \pm\infty$. These are

$$(35) \quad \begin{aligned} m_x &\sim \frac{-h_x}{2\beta_y} + A_x \operatorname{sech}\nu t = \frac{-h_x}{2\beta_y} + \left[\sqrt{\frac{\beta_z - \beta_y}{\beta_z} \left(1 - \frac{h_x^2}{4\beta_y\beta_z}\right)} + \frac{\beta_z - \beta_y}{2\beta_y\beta_z} h_x \right] \operatorname{sech}\nu t \\ m_z &\sim A_z \operatorname{sech}\nu t = \sqrt{1 - \left(\sqrt{\frac{\beta_z - \beta_y}{\beta_z} \left(1 - \frac{h_x^2}{4\beta_y\beta_z}\right)} - \frac{h_x}{2\beta_z} \right)^2} \operatorname{sech}\nu t \end{aligned}$$

where

$$\nu = \sqrt{4(\beta_y - h_x^2/4\beta_z)(\beta_z - \beta_y) + h_x \frac{\beta_z - \beta_y}{\beta_z} \sqrt{4(\beta_z - \beta_y)(\beta_z - \frac{h_x^2}{4\beta_y})}}$$

was found by matching the second derivative at $t = 0$ to the second derivative found from the Hamiltonian system. The coefficient A_x was found by noting that $m_y(0) = 0$ and then solving $\beta_y + h_x^2/(4\beta_y) = \beta_z(1 - m_x^2(0)) - h_x m_x(0)$ for A_x . Then, the coefficient $A_z = \sqrt{1 - m_x^2(0)}$. These solutions are also consistent with $m_x(t) \rightarrow \frac{-h_x}{2\beta_y}$ and $m_z(t) \rightarrow 0$ as $t \rightarrow \infty$. Furthermore, the solutions in (35) reduce to the above exact solutions in (33) when $h_x = 0$.

For computing the averages, we first note that

$$\begin{aligned} \langle m_x \rangle &= \frac{-h_x}{2\beta_y} + A_x \langle \operatorname{sech}\nu t \rangle \\ \langle m_x^2 \rangle &= \frac{h_x^2}{4\beta_y^2} - \frac{h_x}{\beta_y} A_x \langle \operatorname{sech}\nu t \rangle + A_x^2 \langle \operatorname{sech}^2\nu t \rangle \end{aligned}$$

and

$$\langle m_y^2 \rangle = 1 - \langle m_x^2 \rangle - \langle m_z^2 \rangle$$

using the constraint that the magnetization vector has unit length. After computing the approximate average defined in (32), we obtain

$$\begin{aligned} \langle m_x \rangle &\sim \frac{-h_x}{2\beta_y} + \frac{2\pi A_x}{\nu T(E)} \\ \langle m_x^2 \rangle &\sim \frac{h_x^2}{4\beta_y^2} - \frac{h_x 2\pi A_x}{\beta_y \nu T(E)} + \frac{4A_x^2}{\nu T(E)} \\ \langle m_z^2 \rangle &\sim \frac{4A_z^2}{\nu T(E)} \\ \langle m_y^2 \rangle &\sim 1 - \frac{h_x^2}{4\beta_y^2} + \left(\frac{\pi h_x A_x}{2\beta_y} - A_x^2 - A_z^2 \right) \frac{4}{\nu T(E)} \end{aligned}$$

where $T(E) \sim -\frac{2}{\omega} \log(E_b - E)$ when $E_b = \beta_y + h_x^2/4\beta_y$ and $\omega = 2\sqrt{(\beta_y - h_x^2/4\beta_y)(\beta_z - \beta_y)}$. While the same procedure could be carried out again for energy branch 2, we point out that due to symmetry in the system, we can write the averaged energy coefficients on energy branch $j = 1, 2$ (denoted by subscripts) as

$$(36) \quad \begin{aligned} A_j(E) &\sim \frac{4}{\nu T(E)} [4d_j(\beta_z^2 - \beta_y^2) - b_j(4\beta_y^2 + h_x^2)] \\ B_j(E) &\sim \frac{4}{\nu T(E)} \left[\sigma_j b_j h_x + \pi \sqrt{b_j} \beta_y \left(1 - \frac{h_x^2}{4\beta_y^2} \right) \right] \\ C_j(E) &\sim \beta_z - 2\beta_y + \frac{h_x^2}{4\beta_y} - \frac{4}{\nu T(E)} h_x \sqrt{b_j}, \end{aligned}$$

where

$$\begin{aligned} b_j &= \left(\sqrt{\frac{S}{\beta_z}} + \sigma_j (\beta_z - \beta_y) \frac{h_x}{2\beta_y \beta_x} \right)^2 \\ d_j &= 1 - \left(\sqrt{\frac{S}{\beta_z}} - \sigma_j \frac{h_x}{2\beta_z} \right)^2 \\ \nu^2 &= 4 \frac{S\beta_y}{\beta_z} + h_x (\beta_z - \beta_y) \frac{\sqrt{S}}{\beta_z}, \end{aligned}$$

$\sigma_1 = 1$ and $\sigma_2 = -1$, and $S = (\beta_z - \beta_y)(\beta_z - h_x^2/4\beta_y)$. The approximation (36) are in excellent agreement with the averages found via numerical integration, as can be seen in Fig. 4. The limit in Eq. (6) in the text is obtained from the ratios of $A_j(E)$ to $B_j(E)$.

APPENDIX D. MATCHING CONDITIONS

In this section, we derive the matching conditions for the mean first passage time equation (7) in the text. Matching conditions are required only at the saddle point where the energy branches meet [9] because it is a regular boundary point (see [7] for boundary point classification); it is accessible from the interior of each energy branch and the interior of each energy branch is accessible from it. On the other hand, no additional boundary conditions are required at the other ends of the energy branches, specifically the minima, as these are entrance boundary points; the interior of the energy branches are accessible from these points, but the expected passage time from the interior to these points is infinite. This coincides with the diffusion of the magnetization vector on the surface of the unit sphere. The original SDE for the magnetization vector contains no extra conditions prescribed at the single points corresponding to the energy minima and maxima. The classification of the boundary points was determined using the scalings of the averaged energy coefficients derived in Sec. C, and testing the convergence of certain integrals [7].

From the matching conditions, we are able to construct the probabilities of the energy switching from one branch to another (the matching conditions required to supplement Eq. (2) in the text) as well as the pre-factors for the mean first passage times in Eqs. (8) and (9) in the text describing the probability the system switches to the other lower energy branch rather than return to the original one. The derivation is based on the conservation of probability flux of the magnetization vector across the homoclinic orbit on the sphere with

energy equal to the saddle point energy, E_b . For ease of notation, any function evaluated at energy E_b should be interpreted as a limit as $E \rightarrow E_b$ from the interior of the energy branch.

Consider $\rho_j(E, t)$ to be the probability density for the energy while on branch j , normalized so that

$$\int_{E_{a,1}}^{E_b} \rho_1(E, t) dE + \int_{E_{a,2}}^{E_b} \rho_2(E, t) dE + \int_{E_b}^{E_c} \rho_3(E, t) dE + \int_{E_b}^{E_c} \rho_4(E, t) dE = 1.$$

These density functions are continuous at the saddle point in energy:

$$\rho_1(E_b, t) = \rho_2(E_b, t) = \rho_3(E_b, t) = \rho_4(E_b, t).$$

The functions $A_j(\mathbf{m}(t))$ are also continuous across the homoclinic orbit, therefore, if this orbit is approached from the higher or the lower energy branches, we have that

$$\int_0^{T_1(E_b)} A_1(\mathbf{m}(t)) dt + \int_0^{T_2(E_b)} A_2(\mathbf{m}(t)) dt = \int_0^{T_3(E_b)} A_3(\mathbf{m}(t)) dt + \int_0^{T_4(E_b)} A_4(\mathbf{m}(t)) dt$$

or equivalently

$$(37) \quad T_1(E_b)A_1(E_b) + T_2(E_b)A_2(E_b) = T_3(E_b)A_3(E_b) + T_4(E_b)A_4(E_b).$$

This provides the understanding for why the flux of the total probability, $T_j(E)\rho_j(E, t)$, and not simply the averaged probability, $\rho_j(E, t)$, is conserved across the homoclinic orbit.

The forward Kolmogorov equation for the total probability density, written in terms of the flux, $J_j[\cdot]$, on each branch $j = 1, 2, 3, 4$, is

$$(38) \quad \frac{\partial}{\partial t} T_j(E)\rho_j(E, t) = -\frac{\partial}{\partial E} J_j[T_j(E)\rho_j(E, t)]$$

where

$$J_j[T_j(E)\rho_j(E, t)] = \left[-\alpha A_j(E) + a_j B_j(E) + 2\alpha\epsilon C_j(E) \right] T_j(E)\rho_j(E, t) - \alpha\epsilon \frac{\partial}{\partial E} \left(A_j(E) T_j(E)\rho_j(E, t) \right).$$

Analogous to Eq. (37), the conservation of probability flux across the homoclinic orbit provides the matching condition for Eq. (38):

$$(39) \quad J_1[T_1(E_b)\rho_1(E_b, t)] + J_2[T_2(E_b)\rho_2(E_b, t)] = J_3[T_3(E_b)\rho_3(E_b, t)] + J_4[T_4(E_b)\rho_4(E_b, t)].$$

The differential equation (7) in the main text for the mean exit time, $\tau_j(E)$, from energy E , comes from the backwards Kolmogorov equation; it uses the adjoint operator to the one in (38). Therefore, Eq. (7)'s matching condition is the adjoint condition to the conservation of probability flux, Eq. (39). After dividing by $\alpha\epsilon$, the matching condition for Eq. (7) in the text is

$$(40) \quad A_1(E_b)T_1(E_b)\tau_1'(E_b) + A_2(E_b)T_2(E_b)\tau_2'(E_b) = A_3(E_b)T_3(E_b)\tau_3'(E_b) + A_4(E_b)T_4(E_b)\tau_4'(E_b).$$

This condition is equivalent to the condition stated by Fredlein and Wetzell [9].

From the exit time matching condition, (40), we derive the probabilities for the energy to switch branches in order to complete the stochastic differential equation (2) in the text describing the evolutions of the energy, as well as determine the pre-factor for the mean switching times between meta-stable states appearing in Eqs. (8) and (9) in the text.

We define the notation $P(j \rightarrow k)$ to be the probability the energy switches from energy branch j to energy branch k at the saddle point, E_b . In general, this probability is derived

from the coefficients of the matching condition (40) by breaking the integral within the coefficients $A_j(E_b)$ into the parts which lead to each of the other energy branches; these fractional parts out of the whole integral yield the probabilities $P(j \rightarrow k)$. Further simplifications are made by taking advantage of the symmetry of this particular magnetic system.

In general, the probability, $P(j \rightarrow k)$, to switch from branch j to branch k is

$$P(j \rightarrow k) = \frac{\int_0^{T_j(E_b)} A_j(\mathbf{m}(t)) 1_k(\mathbf{m}(t)) dt}{\int_0^{T_j(E_b)} A_j(\mathbf{m}(t)) dt}$$

where $1_k(\mathbf{m}(t)) = 1$ if $\mathbf{m}(t)$ is closer to orbits in branch k than any of the other branches besides the one in which it resides, and 0 otherwise. Immediately from Fig. 1(a) in the text, we see that

$$(41a) \quad P(1 \rightarrow 2) = P(2 \rightarrow 1) = P(3 \rightarrow 4) = P(4 \rightarrow 3) = 0$$

since $1_k(\mathbf{m}(t)) = 1$ at only two individual points (at the green dots). Exploiting the symmetry about the m_x - m_y plane, we have that

$$(41b) \quad P(1 \rightarrow 3) = P(1 \rightarrow 4) = \frac{1}{2},$$

$$(41c) \quad P(2 \rightarrow 3) = P(2 \rightarrow 4) = \frac{1}{2},$$

$$(41d) \quad P(4 \rightarrow 1) = P(3 \rightarrow 1),$$

and

$$(41e) \quad P(4 \rightarrow 2) = P(3 \rightarrow 2).$$

Using the above, we can rewrite

$$(41f) \quad P(3 \rightarrow 1) = \frac{1/2 \int_0^{T_1(E_b)} A_1(\mathbf{m}(t)) dt}{1/2 \int_0^{T_1(E_b)} A_1(\mathbf{m}(t)) dt + 1/2 \int_0^{T_2(E_b)} A_2(\mathbf{m}(t)) dt} = \frac{g_1}{g_1 + g_2}$$

where we define the notation

$$g_j = \int_0^{T_j(E_b)} A_j(\mathbf{m}(t)) dt$$

for $j = 1, 2$. We can take

$$g_j \approx 4d_j(\beta_z^2 - \beta_y^2) - b_j(4\beta_y^2 + h_x^2),$$

coming from Sec. C.2 with out term $4/\nu$, since g_j only appears as fractions. Similarly to (41f), we have that

$$(41g) \quad P(3 \rightarrow 2) = \frac{g_2}{g_1 + g_2}.$$

All together, the conditions (41) provide the switching probabilities for the stochastic energy equation (2) in the text.

The mean first passage time calculation requires the probability the energy switches from branch 1 to 2 or 2 to 1, which we can see from Eq. (41a) never happens along a direct path. Rather, the energy must first switch to one of the two higher energy branches. Conditioning

on which intermediate branch the energy switches to, we have that the probability the energy switches from branch 1 to branch 2 is

$$P(\text{switch from 1}) = P(1 \rightarrow 3)P(3 \rightarrow 2) + P(1 \rightarrow 4)P(4 \rightarrow 2).$$

Using the simplified probabilities in (41), we have that

$$P(\text{switch from 1}) = \frac{P(3 \rightarrow 2)}{2} + \frac{P(3 \rightarrow 2)}{2} = \frac{g_2}{g_1 + g_2}.$$

Similarly, the switching from energy branch 2 back to 1 is

$$P(\text{switch from 2}) = \frac{g_1}{g_1 + g_2}.$$

To match the notation in the text, we define the switching probability from branch j to be

$$(42) \quad P(\text{switch from } j) = \frac{\gamma_j}{\gamma_1 + \gamma_2}$$

where $\gamma_1 = g_2$ and $\gamma_2 = g_1$. The probabilities in (42) are precisely the pre-factors in Eqs. (8) and (9) in the text for the mean switching times.

APPENDIX E. MEAN FIRST PASSAGE TIME

In this section, we derive the mean first passage time equations (8) and (9) in the text. Rather than solve Eq. (7) in the text, it is simpler to find the transition time from starting point x on energy branch $j = 1, 2$ to the saddle point E_b , then account for the probability to transition to the other branch, rather than return to the same well. We therefore find the solution, $\tau_j(x)$, of

$$(43) \quad [-\alpha A_j(x) + a_j B_j(x) + 2\alpha \epsilon C_j(x)] \tau_j'(x) + \alpha \epsilon A_j(x) \tau_j''(x) = -1$$

with absorbing boundary condition $\tau_j(E_b) = 0$, and multiply it by the switching probability in Eq. (42). First we consider the solution valid for any temperature, then consider the limit of vanishing temperature.

The exact solution to (43) requires a second boundary condition. As $x \rightarrow E_a$ we know from Sec. C.1 that $A_j(x) \rightarrow 0$ and $B_j(x) \rightarrow 0$, which leaves the condition

$$2\alpha \epsilon C(E_a) \tau'(E_a) = -1.$$

Using integrating factors, we integrate (43) twice and obtain

$$(44) \quad \tau_j(x) = \frac{\gamma_j}{\gamma_1 + \gamma_2} \int_x^{E_b} \left(\frac{1}{2\alpha \epsilon (\beta_y + \beta_z \pm h_x)} + \frac{1}{\alpha \epsilon} I_j(y) \right) e^{(y-E_a + \frac{\alpha L}{\alpha} F_j(y))/\epsilon - G_j(y)} dy$$

where

$$I_j(y) = \int_{E_a}^y \frac{1}{A(z)} e^{-(z-E_a + \frac{\alpha L}{\alpha} F_j(z))/\epsilon + G_j(z)} dz$$

and where

$$F_j(z) = \int_{E_a}^z \frac{B_j(t)}{A_j(t)} dt \quad \text{and} \quad G_j(z) = \int_{E_a}^z \frac{2C_j(t)}{A_j(t)} dt.$$

The expression for γ_j was described in the previous section; it is

$$\begin{aligned} g_1 &\approx 4d_2(\beta_z^2 - \beta_y^2) - b_2(4\beta_y^2 + h_x^2) \\ g_2 &\approx 4d_1(\beta_z^2 - \beta_y^2) - b_1(4\beta_y^2 + h_x^2) \end{aligned}$$

where b_j and d_j are defined after Eq. (36).

For vanishing temperature ($\epsilon \rightarrow 0$), rather than approximate (44) directly, we notice that the solution has a boundary layer where the coefficients $A_j(x)$ and $B_j(x)$ go to zero: both near $x = E_{a,j}$, the minimum ($E_{a,1} = -h_x$ and $E_{a,2} = h_x$), and $x = E_b$, the saddle point in energy. We match the solution coming out of the boundary layer near E_b to determine the leading order expression for the mean first passage time from the meta-stable point. This meta-stable point is either $E_0 > E_{a,j}$ for values of a_j when a stable limit cycle exits on branch j or the minimum value, $E_{a,j}$. For simplicity of notation, we will drop the subscript j and only consider $j = 1$. The solution for $j = 2$ is derived similarly.

First, we consider the boundary layer near E_b , and rescale the energy by ϵ defining $\xi = (x - E_b)/\epsilon$ so that $\xi \rightarrow -\infty$ leaves the boundary layer. The rescaled equation for $g(\xi) = \tau(x(\xi))$ is

$$\left[-\alpha A(x(\xi)) + a_J B(x(\xi)) + 2\alpha\epsilon C(x(\xi)) \right] \frac{1}{\epsilon} g'(\xi) + \frac{\alpha\epsilon}{\epsilon^2} A(x(\xi)) g''(\xi) = -1$$

which to leading order reduces to

$$\left[-1 + \frac{a_J B(x(\xi))}{\alpha A(x(\xi))} \right] g'(\xi) + g''(\xi) = 0$$

with boundary condition $g(0) = 0$. We then have that

$$g'(\xi) = c \exp \left[\xi - \frac{a_J}{\alpha} \int_0^\xi \frac{B(E_b + \epsilon z)}{A(E_b + \epsilon z)} dz \right]$$

and integrating again yields

$$g(\xi) = c \int_0^\xi \exp \left[y - \frac{a_J}{\alpha} \int_0^y \frac{B(E_b + \epsilon z)}{A(E_b + \epsilon z)} dz \right] dy$$

where we have used the boundary condition $g(0) = 0$. By first expanding the integral in the exponent in term of ϵ ,

$$\int_0^y \frac{B(E_b + \epsilon z)}{A(E_b + \epsilon z)} dz = \int_0^y \frac{\alpha}{\lambda} + O(\epsilon) dz \sim \frac{\alpha}{\lambda} y,$$

where

$$\lambda = \alpha \lim_{x \uparrow E_b} \frac{A(x)}{B(x)} = \alpha \frac{4d(\beta_z^2 - \beta_y^2) - b(4\beta_y^2 + h_x^2)}{bh_x + \pi\sqrt{b}\beta_y \left(1 - \frac{h_x^2}{4\beta_y^2} \right)}$$

is the critical spin-polarized current required to induce switching (obtained for the asymptotic expansions in Sec. C.2), with b and d given by b_1 and d_1 after Eq. (36), we have that

$$g(\xi) = c \int_0^\xi e^{(1-a_J/\lambda)y} dy = \frac{c}{1 - \frac{a_J}{\lambda}} (e^{(1-a_J/\lambda)\xi} - 1)$$

and therefore

$$(45) \quad \tau(x) \approx \frac{c}{1 - \frac{a_J}{\lambda}} e^{(1-a_J/\lambda)(E_b-x)/\epsilon}$$

to leading order. We are left to determine the constant c . As we leave the boundary layer,

$$g(-\infty) = c \int_0^{-\infty} e^{(1-a_J/\lambda)y} dy = \frac{-c}{1 - \frac{a_J}{\lambda}},$$

and we see the solution becomes constant. We turn to consider the full solution in the outer region away from the boundary layer to match this constant.

Returning to Eq. (43), and dividing by $\alpha\epsilon A(x)$, we have

$$\frac{-1}{\epsilon} \left(1 - \frac{a_J B(x)}{\alpha A(x)} - \epsilon \frac{2C(x)}{A(x)} \right) \tau'(x) + \tau''(x) = \frac{-1}{\alpha\epsilon A(x)}.$$

We rewrite this as

$$(46) \quad [e^{\Phi(x)/\epsilon} \tau'(x)]' = \frac{-1}{\alpha\epsilon A(x)} e^{\Phi(x)/\epsilon}$$

where

$$\begin{aligned} \Phi(x) &\equiv \Phi_0(x) + \epsilon\Phi_1(x) \\ &= -x + \frac{a_J}{\alpha} \int_*^x \frac{B(y)}{A(y)} dy + \epsilon \int_*^{E_b} \frac{2C(y)}{A(y)} dy \end{aligned}$$

for some arbitrary point $*$. After integrating (46) from x_0 to E_b we have

$$e^{\Phi(E_b)/\epsilon} \tau'(E_b) - e^{\Phi(x_0)/\epsilon} \tau'(x_0) = \frac{-1}{\alpha\epsilon} \int_{x_0}^{E_b} \frac{1}{A(y)} e^{\Phi(y)/\epsilon} dy.$$

The constant from (45), enters through $\tau'(E_b) = g'(0)/\epsilon = c/\epsilon$. Combining with the above equation we have

$$(47) \quad c = \epsilon\tau'(x_0) e^{-(\Phi(E_b) - \Phi(x_0))/\epsilon} - \frac{1}{\alpha} \int_{x_0}^{E_b} \frac{1}{A(y)} e^{-(\Phi(E_b) - \Phi(y))/\epsilon} dy.$$

The integral in (47) is dominated by what happens near x_0 , and we have two cases, the first when x_0 is the solution to $-\alpha A(x) + a_J B(x) = 0$ in which case $A(x_0) \neq 0$, and the point x_0 is away from either boundary layer. The second is when x_0 is the minimum, and $1/A(x)$ must be canceled by the term generated from the integral of $C(x)/A(x)$ in $\Phi(x)$. In either case, we will need the expansion of

$$-(\Phi_0(E_b) - \Phi_0(x)) = E_b - x - \frac{a_J}{\alpha} \int_x^{E_b} \frac{B(y)}{A(y)} dy$$

in terms of ϵ defined by $\xi = (x - E_b)/\epsilon$. We then have

$$-(\Phi_0(E_b) - \Phi_0(x)) = -\epsilon\xi - \frac{a_J}{\alpha} [0 + \epsilon\xi(-1)\frac{\alpha}{\lambda} + O(\epsilon^2)] \sim -(1 - \frac{a_J}{\lambda})\epsilon\xi.$$

When x_0 is the solution to $-\alpha A(x_0) + a_J B(x_0) = 0$, the mean passage time, τ , is approximately constant at x_0 and therefore $\tau'(x_0) \approx 0$. The expansion of $-(\Phi_1(E_b) - \Phi_1(x))$ only contributes higher order terms to the exponent, and

$$\frac{-1}{\alpha} \int_{x_0}^{E_b} \frac{1}{A(y)} e^{-(\Phi(E_b) - \Phi(y))/\epsilon} dy \approx \frac{-1}{\alpha} \int_{-\infty}^0 \frac{1}{A(x_0)} e^{-(1 - \frac{a_J}{\lambda})\xi} \epsilon d\xi = \frac{\epsilon}{\alpha A(x_0) (1 - \frac{a_J}{\lambda})}.$$

This, together with $\tau'(x_0) = 0$, gives the constant in Eq. (45). Combining with the switching probability factor, we have Eq. (9) in the text,

$$(48) \quad \tau(E_0) \sim \frac{\gamma_1}{\gamma_1 + \gamma_2} \frac{\epsilon}{\alpha A(E_0) (1 - \frac{a_J}{\lambda})^2} e^{(1 - \frac{a_J}{\lambda})\Delta E/\epsilon}$$

where $\Delta E = E_b - E_0$ and E_0 solves $-\alpha A(E_0) + a_J B(E_0) = 0$.

On the other hand when $x_0 = E_a$, the expansion of $-(\Phi_1(E_b) - \Phi_1(x))$ includes a large term near E_a . From the scalings worked out in Sec. C.1, we know $2C(x)/A(x) \sim 1/(x - E_a)$,

which produces a large term, $\log(x - E_a)$, in the expansion of $-(\Phi_1(E_b) - \Phi_1(x))$. Together with the scaling $A(x) \sim 2(\beta_y + \beta_z + h_x)(x - E_a)$ near E_a , we have

$$\begin{aligned} -\frac{1}{\alpha} \int_{E_a}^{E_b} \frac{1}{A(y)} e^{-(\Phi(E_b) - \Phi(y))/\epsilon} dy &\approx -\frac{1}{\alpha} \int_{E_a}^{E_b} \frac{e^{-(\Phi_0(E_b) - \Phi_0(y))/\epsilon + \log(x - E_a)}}{2(\beta_y + \beta_z + h_x)(x - E_a)} dy \\ &\approx -\frac{1}{\alpha} \int_{-\infty}^0 \frac{e^{-(1 - \frac{\alpha I}{\lambda})\xi}}{2(\beta_y + \beta_z + h_x)} \epsilon d\xi = \frac{\epsilon}{\alpha} \frac{1}{2(\beta_y + \beta_z + h_x)(1 - \frac{\alpha I}{\lambda})}. \end{aligned}$$

For the term in (47) involving $\tau'(E_a)$, we return to Eq. (43), where for $x \ll \epsilon$ we have

$$2\alpha\epsilon C(x)\tau'(x) = -1$$

to leading order and therefore

$$\tau'(E_a) = \frac{-1}{2\alpha\epsilon C(E_a)}.$$

We then have

$$\tau'(E_a) e^{-(\Phi(E_b) - \Phi(E_a))/\epsilon} \approx \lim_{x \rightarrow E_a} \frac{1}{2\alpha\epsilon C(x)} e^{-(\Phi_0(E_b) - \Phi_0(x))/\epsilon + \log(x - E_a)}$$

to leading order in the exponent, which goes to zero due to the $\log(x - E_a)$ term. Thus, the $\tau'(E_a)$ term does not contribute to the solution. Combining with the switching probability factor, we have Eq. (8) in the text,

$$(49) \quad \tau(E_a) \sim \frac{\gamma_1}{\gamma_1 + \gamma_2} \frac{\epsilon}{\alpha} \frac{1}{2(\beta_y + \beta_z + h_x)(1 - \frac{\alpha I}{\lambda})^2} e^{(1 - \frac{\alpha I}{\lambda})\Delta E/\epsilon}$$

where $\Delta E = E_b - E_a$. We compare this approximation to the exit time obtained by numerically solving Eq. (43) in Fig. 5.

REFERENCES

- [1] D. M. Apalkov and P. B. Visscher. Spin-torque switching: Fokker-plank rate calculation. *Phys. Rev. B*, 72:180405(R), 2005.
- [2] Charles Augustine, Niladri Narayan Mojumder, Xuanyao Fong, Sri Harsha Choday, Sang Phill Park, and Kaushik Roy. Spin-transfer torque MRAMs for low power memories: Perspective and prospective. *IEEE Sensors J.*, 12(4):756–766, 2012.
- [3] Ya. B. Bazaliy. Precession states in planar spin-transfer devices: The effective one-dimensional approximation. *Phys. Rev. B*, 76:140402(R), 2007.
- [4] Giorgio Bertotti, Isaak Mayergoyz, and Claudio Serpico. *Nonlinear Magnetization Dynamics in Nanosystems*. Elsevier, 2009.
- [5] William Fuller Brown. Thermal fluctuations of a single-domain particle. *Phys. Rev.*, 130(5):1677–1686, 1963.
- [6] Gabriel D. Chaves-O’Flynn, Daniel L. Stein, Andrew D. Kent, and Eric Vanden-Eijnden. Minimum action paths for spin-torque assisted thermally induced magnetization reversal. *J. of Appl. Phys.*, 109:07C918, 2011.
- [7] William Feller. Diffusion processes in one dimension. *T. American Mathematical Society*, 77(1):1–31, 1954.
- [8] Mark I. Freidlin and Wenqing Hu. On perturbations of generalized landau-lifshitz dynamics. *J. Stat. Phys.*, 144:978–1008, 2011.
- [9] Mark I. Freidlin and Alexander D. Wentzell. *Random Perturbations of Hamiltonian Systems*, volume 109 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 1994.
- [10] P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory - 50 years after Kramers. *Rev. Mod. Phys.*, 2:251, 1990.

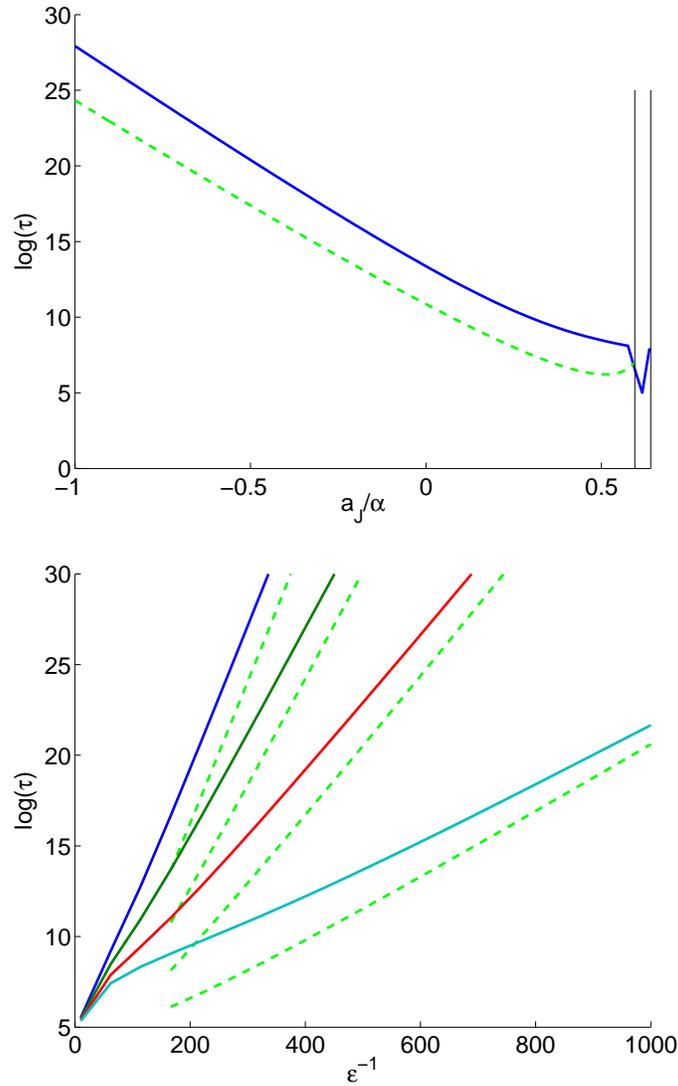


FIGURE 5. Comparison of the mean first passage time on energy branch 1 obtained by numerically solving Eq. (43) (solid lines) to the approximation, Eq. (49) (green dashed lines), plotted as a function of spin-polarized current strength, a_J (top) and as a function of inverse temperature, ϵ^{-1} (bottom). (top) Between the two thin black lines, the meta-stable starting point is no longer the minimum. The constant values $h_x = 0.0296$ and $\epsilon = 0.01$ have been used. (bottom) We have taken, $h_x = 0$ and the spin-polarized currents for the lines from top to bottom are $a_J/\alpha = -0.2, 0.0, 0.2$ and 0.4 . In both plots, the values $\beta_y = 0.06$, $\beta_z = 0.5$, and $\alpha = 0.003$ have been used.

- [11] Matthias Heymann and Eric Vanden-Eijnden. The geometric minimum action method: A least action principle on the space of curves. *Commun Pur Appl Math*, 61(8):1051–1117, 2008.
- [12] S. I. Kiselev, J. C. Sankey, I. N. Krivorotov, N. C. Emley, R. J. Schoelkopf, R. A. Buhrman, and D. C. Ralph. Microwave oscillations of a nanomagnet driven by a spin-polarized current. *Nature*, 425:380–383, 2003.

- [13] R. V. Kohn, M. G. Reznikoff, and Eric Vanden-Eijnden. Magnetic elements at finite temperature and the large deviation theory. *J. Nonlinear Science*, 15:223–253, 2005.
- [14] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.
- [15] I. N. Krivorotov, N. C. Emley, A. G. F. Garcia, J. C. Sankey, S. I. Kiselev, D. C. Ralph, and R. A. Buhrman. Temperature dependence of spin-transfer-induced switching of nanomagnets. *Phys. Rev. Lett.*, 93(16):166603, 2004.
- [16] Z. Li and S. Zhang. Magnetization dynamics with spin-transfer torque. *Phys. Rev. B*, 68:024404, 2003.
- [17] Z. Li and S. Zhang. Thermally assisted magnetization reversal in the presence of a spin-transfer torque. *Phys. Rev. B*, 69:134416, 2004.
- [18] E. B. Myers, F. J. Albert, J. C. Sankey, E. Bonet, R. A. Buhrman, and D. C. Ralph. Thermally activated magnetic reversal induced by a spin-polarized current. *Phys. Rev. Lett.*, 89(19):196801, 2002.
- [19] B. Özyilmaz, Andrew D. Kent, D. Monsma, J. Z. Sun, M. J. Rooks, and R. H. Koch. Current-induced magnetization reversal in high magnetic fields in co/cu/co nanopillars. *Phys. Rev. Lett.*, 91(6):067203, 2003.
- [20] Grigoris Pavliotis and Andrew Stuart. *Multiscale Methods: Averaging and Homogenization*. Springer, 2008.
- [21] D. C. Ralph and M. D. Stiles. Spin transfer torques. *J. of Magnetism and Magnetic Materials*, 320:1190–1216, 2008.