PHYSICAL REVIEW B **72**, 180405(R) (2005)

Spin-torque switching: Fokker-Planck rate calculation

D. M. Apalkov* and P. B. Visscher†
MINT Center and Department of Physics and Astronomy, University of Alabama, Tuscaloosa, Alabama 35487-0324, USA
(Received 19 October 2005; published 23 November 2005)

We describe a general Fokker-Planck approach to understanding and calculating magnetization switching rates and noise in the recently observed phenomenon of spin-torque switching. In this phenomenon, which has possible applications to information storage, a large current passing from a pinned ferromagnetic (FM) layer to a free FM layer switches the free layer. Beginning with Brown [Phys. Rev. **130**, 1677 (1963)], switching rates in magnetic systems have been calculated using the Fokker-Planck equation. In the small-oscillation limit, the equations have been solved analytically, giving a first-principles justification for phenomenological effective temperature theories: the spin-torque effect increases the Arrhenius factor $\exp(-E/kT)$ in the switching rate by raising the effective spin temperature T. In the present Rapid Communication we generalize the nonlinear Fokker-Planck equation to the case of a Slonczewski spin torque. As an example, we use a linear approximation to calculate telegraph noise rates, leading to good qualitative agreement with recent experiments. However, our nonlinear formulation is also valid for large precessional oscillations. The method also allows the calculation of current-induced magnetic noise in current perpendicular to plane spin valve read heads.

DOI: 10.1103/PhysRevB.72.180405 PACS number(s): 75.70.−i, 75.10.−b, 75.47.−m, 85.75.−d

Recently it has been demonstrated that the magnetization of a thin ferromagnetic film can be switched by passing a current between it and a pinned layer.[1] This "spin-torque switching" phenomenon is of interest for possible information storage applications. Except at very high currents, the switching appears to be thermal in nature. Previous theoretical treatments of thermal spin-torque switching[2–4] have been based on the idea that the spin torque increases the rate by lowering the effective potential energy barrier, and have encountered a fundamental problem: the common Slonczewski[5,6] model for the spin torque is not conservative, so it cannot be described by a potential energy. The effects of the Slonczewski torque on the Landau-Lifshitz (LL) equation for the magnetization dynamics are similar to those of the LL damping, so in the Fokker-Planck approach it makes a contribution to the effective damping. When this contribution is negative, the effective temperature is raised. The notion of an elevated effective temperature during spin-torque switching has been discussed previously;[7–9] the Fokker-Planck formulation allows the precise definition and calculation of the effective temperature, which we will refer to as the Boltzmann temperature [Eq. (16)] and clarifies the relation between it and the (lower) LL noise temperature. (This relation has recently been described in the small-oscillation limit by Li and Zhang;[10] in the present Rapid Communication we give a general nonlinear formulation applicable also to large-amplitude precession and switching.) The Fokker-Planck equation gives the time evolution of a phase-space probability density. It was first applied to chemical rate problems in 1940 by Kramers,[11] who observed that except for very large or very small damping constants, the escape rate is well described by an earlier transition state theory (TST),[12] in which the rate of barrier crossing in a nonequilibrium system is assumed to be the same as that in an equilibrium system. Although corrections to TST have been extensively studied,[13,14] TST has been found to be the most useful starting point for rate calculations. In this Rapid Communication we will use a TST-like approximation, differing from the usual TST in that the system is not in a true thermal equilibrium, but a nonequilibrium steady state. We will write the magnetic Fokker-Planck (FP) equation of Brown,[15] generalized to include the Slonczewski torque, but following Kramers[11] convert it to describe diffusion in energy rather than magnetization; to the best of our knowledge this has not been done previously except for systems with azimuthal symmetry.[15,16]

The LL equation[17] for the evolution of a uniform magnetization $\mathbf{M}(t)$ has a deterministic and a random part:

$$\dot{\mathbf{M}} \equiv \frac{d\mathbf{M}}{dt} = \dot{\mathbf{M}}_{\mathrm{det}} + \dot{\mathbf{M}}_{\mathrm{rand}}. \tag{1}$$

The deterministic part is divided into a conservative precession term and the dissipative LL damping, and we will include also the Slonczewski current-induced torque:

$$\dot{\mathbf{M}}_{\mathrm{det}} = \dot{\mathbf{M}}_{\mathrm{cons}} + \dot{\mathbf{M}}_{\mathrm{LL}} + \dot{\mathbf{M}}_{\mathrm{Slon}}. \tag{2}$$

We will first specify the precession torque

$$\dot{\mathbf{M}}_{\mathrm{cons}} = -\gamma \mathbf{M} \times \mathbf{H}_{\mathrm{cons}} \tag{3}$$

where γ is the gyromagnetic ratio. We refer to the field $\mathbf{H}_{\mathrm{cons}}$ about which $\mathbf{M}$ precesses as "conservative" because it can be written as the gradient with respect to $\mathbf{M}$ of an energy density,

$$\mu_0 \mathbf{H}_{\mathrm{cons}} = -\nabla E(\mathbf{M}). \tag{4}$$

[This is a two-dimensional 2D gradient on the $\mathbf{M}$-sphere; see Eq. (A1) of Ref. 18.]

Our derivation of the FP equation is valid for a system with arbitrary anisotropy, but for specificity we will consider the case of a thin-film element (Fig. 1) for which the energy density is given (in SI units) by[17]

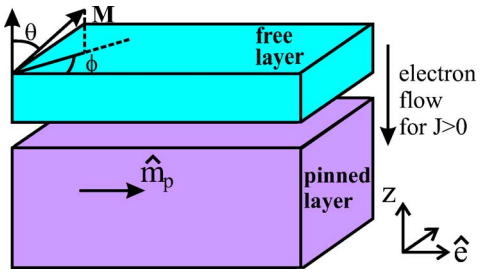

FIG. 1. (Color online) Geometry of thin-film element, for the case where the magnetization $\hat{\mathbf{m}}_p$ of the "pinned" layer is along the easy axis $\hat{\mathbf{e}}$ of the free layer.

$$E(\mathbf{M})/\mu_0 = -\frac{1}{2}H_K M_s(\hat{\mathbf{m}}\cdot\hat{\mathbf{e}})^2 + \frac{1}{2}M_s^2(\hat{\mathbf{m}}\cdot\hat{\mathbf{z}})^2 - \mathbf{H}_{\text{ext}}\cdot\mathbf{M}. \tag{5}$$

Here $\mathbf{H}_{\text{ext}}$ is an external field, H_K is the uniaxial anisotropy field, $\hat{\mathbf{m}}\equiv\mathbf{M}/\mathbf{M}_s$, $\hat{\mathbf{e}}$, and $\hat{\mathbf{z}}$ are unit vectors along the magnetization, easy axis, and z axis (perpendicular to the film), respectively, and M_s is the saturation magnetization.

The nonconservative LL damping torque [Eq. (2)] is[17]

$$\dot{\mathbf{M}}_{\text{LL}} = -\gamma\alpha M_s\hat{\mathbf{m}}\times(\hat{\mathbf{m}}\times\mathbf{H}_{\text{cons}}) = \gamma\alpha M_s\mathbf{H}_{\text{cons}} \tag{6}$$

where α is the dimensionless LL damping constant ($\mathbf{H}_{\text{cons}}$ is perpendicular to $\hat{\mathbf{m}}$).

The Slonczewski spin torque[3,6] is

$$\dot{\mathbf{M}}_{\text{Slon}} = -\gamma JM_s\hat{\mathbf{m}}\times(\hat{\mathbf{m}}\times\hat{\mathbf{m}}_{\text{p}}) \tag{7}$$

where J is an empirical constant with units of magnetic field, proportional to the current density, and $\hat{\mathbf{m}}_p$ is the magnetization direction in the thicker (often pinned) layer from (or to) which the current flows.

The effect of the random torque $\dot{\mathbf{M}}_{\text{rand}}$ is to produce a diffusive random walk on the surface of the $\mathbf{M}$-sphere. We will relate this to a diffusivity D [Eq. (18)] by giving the mean square value of the increment $\Delta\mathbf{M}_{\text{rand}}=\dot{\mathbf{M}}_{\text{rand}}\Delta t$:

$$\langle\Delta\mathbf{M}_{\text{rand}}^2\rangle = 4D\Delta t. \tag{8}$$

The directions of these torques are shown in insets to Fig. 2, from which the basic mechanism of spin-torque switching can be seen: the Slonczewski torque pulls the magnetization out of the left well (number 1) and allows it to jump to well 2.

The Fokker-Planck equation describes the evolution of a probability density $\rho(\mathbf{M},t)$ on the $\mathbf{M}$-sphere. It can be written in the form of a continuity equation[15] for $\rho(\mathbf{M},t)$:

$$\frac{\partial\rho(\mathbf{M},t)}{\partial t} = -\boldsymbol{\nabla}\cdot\mathbf{j}(\mathbf{M},t) \tag{9}$$

where the probability current $\mathbf{j}$ along the sphere has a convective and a diffusive part:

$$\mathbf{j}(\mathbf{M},t) \equiv \rho(\mathbf{M},t)\dot{\mathbf{M}}_{\text{det}}(\mathbf{M}) - D\,\boldsymbol{\nabla}\,\rho(\mathbf{M},t) \tag{10}$$

(note that both the divergence and the gradient are two dimensional here). Inserting Eq. (10) into Eq. (9) gives the FP

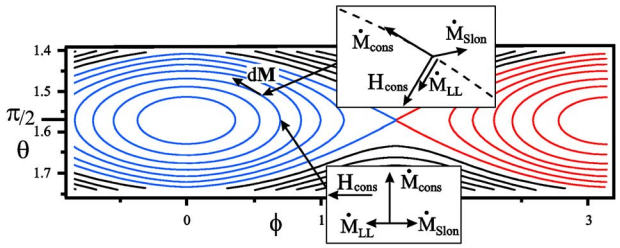

FIG. 2. (Color online) Energy contours (Stoner-Wohlfarth orbits) for a thin film, plotted in terms of the coordinates θ and ϕ defined in Fig. 1, for the case $H_K/M_s=0.028$, $\mathbf{H}_{\text{ext}}=\mathbf{0}$. The regions $i=1,2,3$ are blue, red, and black respectively. The vertical scale is exaggerated for clarity. Lower inset: contributions to the rate of change of magnetization for a magnetization in the film plane. The insets show the tangent plane: magnetization points out of the paper. Upper inset: the same for an arbitrarily chosen direction of $\mathbf{M}$.

equation [Eq. (A2) of Ref. 18] first derived (without the spin torque term) in 1963 by Brown.[15]

Frequently, the probability density ρ depends mostly on energy, depending weakly on phase around a constant-energy orbit. In a thermal equilibrium system, even with damping, ρ is exactly independent of phase. This has often been assumed to be approximately true except near the barrier, to compute nonequilibrium switching rates.[11,15,16] In a nonequilibrium steady state, such as the telegraph-noise state considered below, if the damping α and the current J vanish the system follows a closed constant-energy orbit and ρ depends only on energy. If α and J are small, as in all experiments of interest here, the variation of ρ along an orbit is of first order in α and J. The effect of α and J on a deterministic orbit is (a) to shrink or enlarge the orbit (drift in energy) and (b) to cause the orbit to drift (horizontally or vertically in Fig. 2). We will assume (as has most experimental and theoretical work in this area) that the pinned magnetization m_p is parallel to the easy axis; in this case the orbit drift will vanish by symmetry, and only the energy drift need be considered.

Kramers derived a Fokker-Planck equation in energy for a particle in a well, but we are not aware of any previous derivation for the magnetic case so we will derive it here. Formally, we need only assume that the probability density depends only on energy. However, the energy dependence may be different in different regions of the sphere (for example, different energy wells), so we will define a density $\rho_i'(E,t)$, where the region (well) index $i=1$ for the $\phi=0$ well (for $E\leqslant E_{\text{sad}}$), $i=2$ for the $\phi=\pi$ well (Fig. 2), and $i=3$ for $E\geqslant E_{\text{sad}}$. (The three will be equal at the saddle point, where all three regions touch.) This density ρ' is related to ρ_i by

$$\rho(\mathbf{M},t) = \rho_i'(E(\mathbf{M}),t). \tag{11}$$

The FP equation in energy takes the form of a continuity equation

$$\frac{\gamma M_s P_i(E)}{\mu_0}\frac{\partial\rho_i'(E,t)}{\partial t} = -\frac{\partial}{\partial E}j_i^E(E,t) \tag{12}$$

where the current $j_i^E(E,t)$ is the number of systems per unit time crossing a constant-energy contour. There is a factor on the left-hand side involving the orbital period $P_i(E)$ because

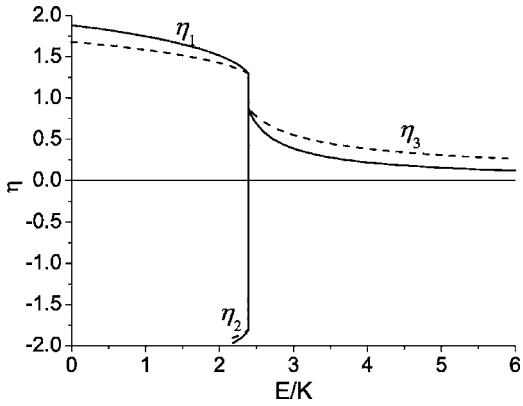

FIG. 3. The current-damping coefficients η_1 and η_2 in the two potential wells, and η_3 for the region above the saddle-point energy. The values we actually use (near the bottoms of the wells) depend only weakly on the parameters $H_{\text{ext}}=-120$ Oe and $H_K=220$ Oe, which were estimated from the fit to experimental data [Eq. (A28) of Ref. 18]. The averages $\bar{\eta}_i$ [Eq. (20)] are also shown, as dashed lines.

ρ' is not the probability per unit energy but per unit area on the M-sphere [see Eq. (A23) of Ref. 18]. The current in energy can be obtained from the current on the M-sphere [Eq. (10); see Eqs. (A4)–(A13) of Ref. 18 for details]:

$$j_i^E(E,t) = -\gamma\alpha M_s \rho_i'(E,t) I_i^E(E) + \gamma J \rho_i'(E,t)\hat{\mathbf{m}}_{\text{p}} \cdot \mathbf{I}_i^M - D\frac{\partial \rho_i'(E,t)}{\partial E} I_i(E) \tag{13}$$

in terms of a damping term involving an energy integral over an orbit in the ith well

$$I_i^E(E) \equiv \oint H_{\text{cons}} dM, \tag{14}$$

a Slonczewski torque term involving a magnetization integral

$$\mathbf{I}_i^M(E,t) = \oint d\mathbf{M} \times \mathbf{M}, \tag{15}$$

and a diffusion term.

For the telegraph-noise problem we require the steady-state form obtained by setting $j_i^E=0$:

$$\frac{\partial \ln \rho_i'(E)}{\partial E} = \frac{\gamma}{D}[-\alpha M_s + \eta_i(E) J] \equiv -V\beta_i(E), \tag{16}$$

where the right-hand side defines an effective inverse "Boltzmann" temperature $\beta_i(E)$, and V is the volume of the switching element. We have also defined a dimensionless spin-torque-damping ratio $\eta(E)$ (Fig. 3) as the ratio of the work of the Slonczewski torque [Eq. (7)] to that of the LL damping [Eq. (6)]

$$\eta_i(E) = \frac{\hat{\mathbf{m}}_{\text{p}} \cdot \mathbf{I}_i^M(E)}{I_i^E(E)}. \tag{17}$$

Equation (16) shows clearly that the Slonczewski torque acts like a correction to the LL damping α. Because η has opposite signs in the two wells, the damping contribution is negative in one well and positive in the other. A similar result has been suggested previously[9] for the special case of an isotropic material in which $\hat{\mathbf{m}}_p$ is parallel to $\mathbf{H}$.

If $J=0$, we get the expected Boltzmann distribution with $\beta=1/k_BT$ only if

$$D = \gamma M_s \alpha k_B T/V; \tag{18}$$

this is the fluctuation-dissipation theorem. Like previous authors,[3,4,9] we assume that the noise temperature T is unchanged by the current J.

If we integrate Eq. (16) downward from the saddle point into well $i=1$ or 2, we get

$$\rho_i'(E) = \rho'(E_{\text{sad}}) \exp\left(\frac{V}{k_BT}[1-\bar{\eta}_i(E)J/\alpha M_s](E_{\text{sad}}-E)\right) \tag{19}$$

where the average η is

$$\bar{\eta}_i(E) \equiv \frac{1}{[E_{\text{sad}}-E]}\int_E^{E_{\text{sad}}} \eta_i(E')dE'. \tag{20}$$

In region $i=3$ we must integrate upward from E_{sad} [see Eq. (A20) of Ref. 18]. The ratio η and its average $\bar{\eta}$ are weakly dependent on E (Fig. 3), so the distribution is nearly a Boltzmann distribution with an effective temperature[10] $T/[1-\bar{\eta}_i(E)J/\alpha M_s]$.

We now compute the switching rates using transition state theory. The TST rate is the steady-state probability per unit time of crossing a vertical line ($\phi=\pi/2$) through the saddle point in Fig. 2. This gives[18]

$$j_{\text{TST}} = \frac{\gamma M_s k_B T}{\mu_0 V[1-\bar{\eta}_3(E_{\text{sad}})J/\alpha M_s]}\rho'(E_{\text{sad}}). \tag{21}$$

From the ρ_i''s [Eq. (19)] it is straightforward to obtain the total probability p_i of being in each well [Eq. (A25) of Ref. 18]. With the transition-state-theory current j_{TST} [Eq. (21)] these determine the dwell times τ_1 and τ_2. We will write these in terms of a stability factor $S_i \equiv VE_i^b/k_BT$ (E_i^b is the barrier height $E_{\text{sad}}-E_i$, where E_i is the bottom of well $i=1$ or 2) and a critical current at which the exponent in Eq. (19) vanishes, $J_{ci} \equiv \alpha M_s/\bar{\eta}_i(E_i)$. Since we do not know the exact proportionality factor between the parameter J and the actual physical current I we can write J/J_{ci} as I/I_{ci}, where the critical currents I_{ci} should be related by

$$I_{c1}\bar{\eta}_1(E_1) = I_{c2}\bar{\eta}_2(E_2) = I_{c3}\bar{\eta}_3(E_{\text{sad}}). \tag{22}$$

Then the dwell times are given by[18]

$$\tau_i = \frac{p_i}{2j_{\text{TST}}} = \frac{1}{2}P_i(E_i)\frac{1-I/I_{c3}}{1-I/I_{ci}}(e^{S_i(1-I/I_{ci})}-1) \tag{23}$$

in the "high-barrier" limit in which the exponent is large (at the crossing point in Fig. 4 it is about 10). We define an Arrhenius-Néel approximation by neglecting I in the prefactor and the -1:

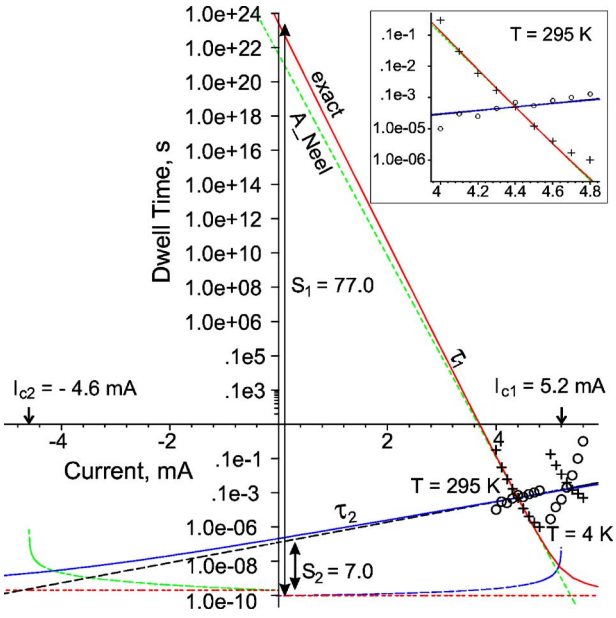

FIG. 4. (Color online) Dwell times as functions of current, illustrating the fitting of the room-temperature data of Urazhdin *et al.* (Ref. 7). Inset (upper right) shows the experimental points (+ for τ_1, ○ for τ_2) more clearly. The exact results are solid lines; the Arrhenius-Néel approximation is dashed. Dotted (red, online) horizontal straight lines at the bottom are the orbit periods (P_1 on the right, P_2 on the left). Curves that approach them at $I=0$ are the prefactor $P_i(1-I/I_{c3})/2(1-I/I_{ci})$, which intersects the exact τ_i at I_{ci}. We use $I_{c3}=I_{c1}\overline{\eta_1}/\overline{\eta_3}$. Low-temperature (4 K) data cannot be fitted well; possible reasons (Joule heating) are discussed in Ref. 19.

$$\tau_i^{A\text{-}N}=\frac{1}{2}P_i(E_i)e^{S_i(1-I/I_{ci})}, \tag{24}$$

so that the dwell time is just a straight line on a logarithmic plot of τ (Fig. 4). We adjust the two parameters S_1 and I_{c1} to match the slope and value of the measured[7] dwell time at the current $I=4.4$ mA at which τ_1 and τ_2 cross. In the Arrhenius-Néel approximation, these constants have simple graphical interpretations: I_{c1} is the current at which τ_1 intersects the horizontal line at the prefactor P_i (the orbit period), and S_1 is the (logarithmic) height of the dwell time above this prefactor at zero current.

The experimental data determine I_{c1} within a few percent because the extrapolation to the prefactor curve is very short. The long extrapolation needed for I_{c2} would produce large uncertainties, so we instead determine I_{c2} from Eq. (22), and adjust S_2 to fit τ_2 at the crossing point—this gives good semiquantitative agreement with the experimental data. Although we forced the slope of τ_1 to agree, the fact that the slope of τ_2 is much smaller is a true prediction of the theory. The effective temperatures $T/(1-I/I_{ci})$ are 2135 and 158 K.

The theory developed here is also applicable to the calculation of magnetic noise in read heads;[20] the large fluctuations seen in simulations[21] are predicted by the present theory as $I\to I_{c1}$.

This work was partially supported by NSF Grants No. ECS-0085340 and No. DMR-MRSEC-0213985, and by the DOE Computational Materials Sciences Network.

*Electronic address: apalk001@ua.edu

†Electronic address: visscher@ua.edu; URL: http://bama.ua.edu/~visscher/mumag

[1] F. J. Albert, J. A. Katine, R. A. Buhrman, and D. C. Ralph, Appl. Phys. Lett. **77**, 3809 (2000).

[2] E. B. Myers, F. J. Albert, J. C. Sankey, E. Bonet, R. A. Buhrman, and D. C. Ralph, Phys. Rev. Lett. **89**, 196801 (2002).

[3] Z. Li and S. Zhang, Phys. Rev. B **68**, 024404 (2003).

[4] Z. Li and S. Zhang, cond-mat/0302339v1 (unpublished).

[5] J. Slonczewski, J. Magn. Magn. Mater. **159**, L1 (1996).

[6] J. Slonczewski, J. Magn. Magn. Mater. **195**, L261 (1999).

[7] S. Urazhdin, N. O. Birge, W. P. Pratt, and J. Bass, Phys. Rev. Lett. **91**, 146803 (2003).

[8] J.-E. Wegrowe *et al.*, J. Appl. Phys. **91**, 8606 (2002).

[9] R. H. Koch, J. A. Katine, and J. Z. Sun, Phys. Rev. Lett. **92**, 088302 (2004).

[10] After the completion of this work, we learned that a Fokker-Planck description of the linear (small-amplitude oscillation) case has been recently treated by Z. Li and S. Zhang, Phys. Rev. B **69**, 134416 (2004).

[11] H. A. Kramers, Physica (Amsterdam) **7**, 284 (1940).

[12] H. Eyring, J. Chem. Phys. **3**, 107 (1935).

[13] D. Chandler, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, Singapore, 1998), p. 3.

[14] P. B. Visscher, Phys. Rev. B **14**, 347 (1976).

[15] W. F. Brown, Phys. Rev. **130**, 1677 (1963).

[16] W. T. Coffey, D. S. F. Crothers, J. L. Dormann, Y. P. Kalmykov, E. C. Kennedy, and W. Wernsdorfer, Phys. Rev. Lett. **80**, 5655 (1998).

[17] J. Fidler and T. Schrefl, J. Phys. D **33**, R153 (2000).

[18] See EPAPS Document No. E-PRBMDO-72-R10542 for some details of derivations for which there was not space in this Rapid Communication. This document can be reached via a direct link in the online article's HTML reference section or via the EPAPS homepage (http:www.aip.org/pubservs/epaps.html).

[19] I. N. Krivorotov, N. C. Emley, A. G. F. Garcia, J. C. Sankey, S. I. Kiselev, D. C. Ralph, and R. A. Buhrman, Phys. Rev. Lett. **93**, 166603 (2004).

[20] S. I. Kiselev, J. C. Sankey, I. N. Krivorotov, N. C. Emley, R. J. Schoelkopf, R. A. Buhrman, and D. C. Ralph, Nature (London) **425**, 380 (2003).

[21] J. G. Zhu and X. Zhu, IEEE Trans. Magn. **40**, 182 (2004).