Effect of subvolume excitation and spin-torque efficiency on magnetic switching

J. Z. Sun, R. P. Robertazzi, J. Nowak, P. L. Trouilloud, G. Hu, D. W. Abraham, M. C. Gaidis, S. L. Brown, E. J. O'Sullivan, W. J. Gallagher, and D. C. Worledge

IBM-MagIC MRAM Development Alliance, IBM T. J. Watson Research Center, P. O. Box 218, Yorktown Heights, New York 10598, USA

(Received 16 March 2011; revised manuscript received 11 May 2011; published 19 August 2011)

Recently developed magnetic tunnel junctions with full perpendicular magnetization that are spin-torque switchable allow for quantitative comparison of spin-torque switching statistics with a macrospin model. For typical devices above 50 nm in lateral size, the comparison suggests the presence of subvolume magnetic excitations which often dominate the switching process and which degrade the spin-torque switching efficiency. A simple model of subvolume spin-torque-driven magnetic switching is presented to account for the experimental observations. The origin of the subvolume thermal excitation is traced to a competition between the macrospin fluctuation within a simple uniaxial anisotropy potential and that of thermal magnon excitation. The subvolume excitation problem highlights the importance of improving the magnetic exchange stiffness of the junction free layer, and the reduction of junction lateral sizes below 50 nm where an improved spin-torque efficiency is seen as the switching dynamics cross over to a more macrospin-like process.

DOI: 10.1103/PhysRevB.84.064413 PACS number(s): 75.78.Jp, 85.75.Dd, 75.76.+j, 73.40.−c

I. INTRODUCTION

Spin-torque-induced magnetic switching has been of interest for the manipulation of nanomagnets in solid-state memory technologies[1–6] and for nanomagnetic oscillator applications. The simplest model system of a spin-torque-driven magnetic switch is that of a macrospin under spin torque and in a simple uniaxial anisotropy energy potential.[7] A set of quantitative predictions has been developed for such a model system at finite temperatures based on the Landau-Lifshitz-Gilbert (LLG) equation, its finite temperature probabilistic dynamics,[8,9] and the corresponding Fokker-Planck equation description.[10–14] Most of these predictions, however, have not been experimentally tested quantitatively due to the difficulties of preparing an experimental materials system sufficiently simple so as to warrant a direct comparison with these analytical model results. Recent materials and device technologies[3,15–21] have brought forth a new generation of spin-torque-based magnetic switches[3,20–23] that are based on sub-100-nm magnetic tunnel junctions and employ perpendicular magnetic anisotropy (PMA) thin film elements, so that the anisotropy energy now closely resembles that of the simplest possible geometry. Early dynamic measurements in PMA-based switching in spin valves show a thermal activation volume smaller than free-layer size.[18,19] More recently, quantitative measurements of switching probabilities on PMA-based magnetic tunnel junctions (MTJs) with great statistical precision have been accomplished,[21,22] thanks, in part, to a complementary metal-oxide semiconductor (CMOS) integrated device test environment that allows for rapid counting of individual switching events.[21,22,24,25] Here we offer a quantitative comparison of this new class of MTJ spin-torque devices with the macrospin model. The comparison reveals substantial agreement. It also highlights important deviations of devices from the ideal model, most notably, the presence of subvolume thermal and spin-torque-induced activation and reversal. The subvolume thermal activation process is consistent with earlier reports of similar behaviors in nanomagnet systems (see, for example, Refs. 26–28). The subvolume excitation originates from thermal agitation of a thin film nanomagnet's internal magnetic degrees of freedom. This, in our thin film case, leads to a prediction that the maximum thermal activation volume and activation energy should be approximately proportional to the product of the exchange-stiffness constant and the film thickness. In addition to a limited activation energy, such thermal agitation also modifies the dynamic behavior and size scaling of a nanomagnet's reversal under spin torque.

II. EXPERIMENT

The MTJs studied here are of the full PMA type,[20,21] based on an interface- and/or strain-induced perpendicular anisotropy drive force induced in the CoFeB free layer.[20,21,29,30] Details of our materials and device preparation were reported earlier.[21] To briefly repeat, the junction thin film stack consists of ‖ Substrate ‖ 5 RuCoFe | 2 Ta | 0.8 CFB | 0.9 MgO | 0.5 Fe | 0.8 CFB | 0.3 Ta | 0.25 Co | 0.8 Pt | [0.25 Co | 0.8 Pd] × 4 | 0.3 Co | 0.9 Ru | [0.25 Co | 0.8 Pd] × 14 | 20 Ru ‖, where the numbers preceding layer composition are thicknesses in nanometers, and where CFB = $Co_{0.60}Fe_{0.20}B_{0.20}$. The CFB layer below the MgO barrier acts as the magnetic free layer. The samples were annealed at 240 °C for 1 h and then patterned into circles with diameters around 70 to 120 nm. Our knowledge of such values is limited to an accuracy of perhaps around 20% when comparing large numbers of devices fabricated at different times due to various fabrication process variations. The MTJs shown here have typical tunnel barrier resistance-area products (RA) around 5–10 Ω μm^2, giving a final junction resistance on the order of 1 kΩ. The magnetic easy axis of the film stack is perpendicular, with the MTJ's resistance versus field loop showing square easy-axis behavior, as shown in Fig. 1(a), where the MTJ's resistance versus easy-axis applied magnetic field was measured at 50-mV bias. The MTJ free-layer's magnetic state can be switched at a bias voltage well below 1 V. In Fig. 1(b) the MTJ's current-voltage (IV) characteristics are shown for its parallel (P), antiparallel (AP), and switching states. The P and AP states in this

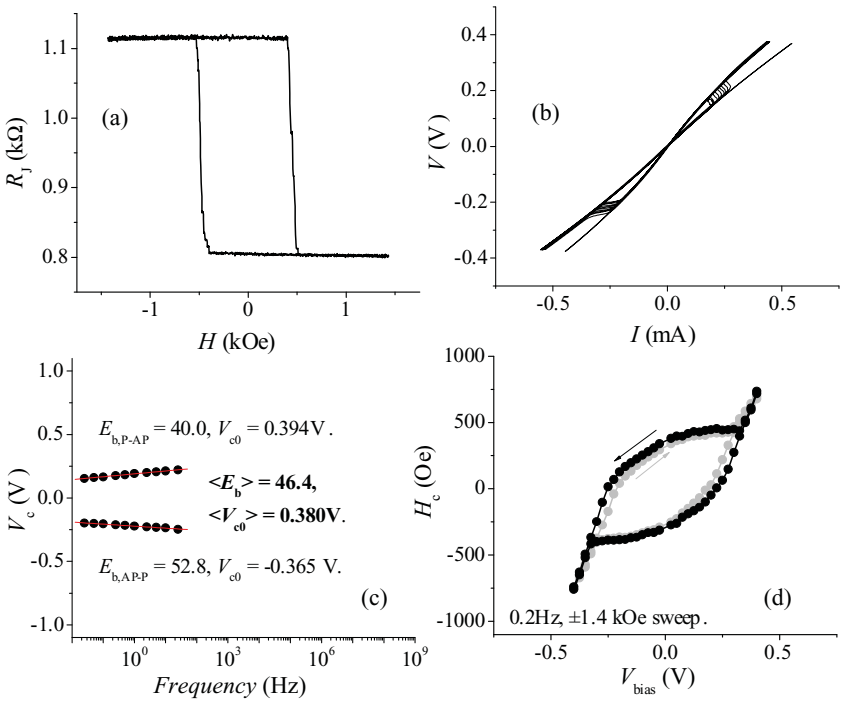

FIG. 1. (Color online) A representative MTJ spin-torque switch. (a) The MTJ's resistance vs easy-axis-applied magnetic field. (b) The MTJ's current-voltage (IV) characteristics in its parallel (P), antiparallel (AP), and switching states. (c) The frequency dependence of the switching voltage position from (b). Values of calculated E_b shown are in units of k_BT with $T = 300$ K. (d) The switching threshold boundary in (V,H) phase space, measured by sweeping $R(H)$ loops such as shown in (a) at different junction bias voltages.

measurement are stabilized by externally applied easy-axis magnetic fields. The switching branches are measured with an easy-axis magnetic field bias at the center of the $R(H)$ hysteresis loop shown in Fig. 1(a), and by using a triangular wave sweep of the bias voltage with a peak-to-peak amplitude of $\pm 0.4 \sim 0.6$ V at frequencies from 0.025 to 25 Hz. The IVs shown in Fig. 1(b) are 25 to 100 times trace averaged for each sweep frequency. Figure 1(c) shows the sweep-frequency dependence of the switching voltage from Fig. 1(b). Such rate-dependent switching threshold is an indication of the existence of a characteristic switching time that follows the thermally activated switching probability as discussed in the Appendix. This can be used to estimate the thermal activation energy E_b of the junction free layer. For this particular device representative of the discussion in this paper, its E_b thus measured is about $46k_BT$ with $T \approx 300$ K. Figure 1(d) shows the switching threshold boundary in (V,H) phase space, measured here by sweeping $R(H)$ loops such as shown in Fig. 1(a) at different junction bias voltages. The magnetic field was swept at 0.2 Hz.

Using such junctions one then studies in detail the statistical behavior of the spin-torque-induced magnetic switching for a variety of bias voltage amplitudes and time durations. A large number of repeated switchings were performed for an individual junction to build good statistics for the temporal distribution of the switching behavior. This was done by exercising the CMOS memory circuit to digitally discriminate individual switching events as a function of the switching pulse width and height.[31,32] The probability of a junction *not* switching (hence the name of a bit-error rate or BER), E_r, as a function of the voltage pulse height at five different pulse widths is shown in Fig. 2. Each point on this plot represents about 10^6 repeated attempts of switching, from which E_r was calculated. The vertical quantity plotted here is S_r, the normal-quantile value which is $E_r \in [0,1]$ mapped through an inverse Gaussian-Normal distribution function for easy comparison with experiments over many decades of S_r. The mapping relationship is

$$S_r = \sqrt{2}\,\mathrm{inverf}(2E_r - 1), \tag{1}$$

where "inverf" is the inverse of the cumulative normal distribution (erf) function. $S_r \in [-\infty,+\infty]$. Note that a Gaussian cumulative distribution of the error probability E_r would result in a straight line on such plots, with a slope of $\pm 1/\sigma$ with σ being the standard deviation of the Gaussian.[33] In reality, for spin-torque switching the switching error probability does not follow a Gaussian distribution, hence the data in Fig. 2 are not linear.

Data such as those shown in Fig. 2 can then be used to extract the dependence of the switching speed and distribution width as a function of the drive pulse amplitude and pulse width, as shown in Fig. 3. Results from five different junction's BER versus V_{pulse} such as Fig. 2 were used to extract the data shown in Fig. 3. The threshold voltage V_s in Fig. 3 is defined as $V_s = V_{\mathrm{pulse}}(S_r = -2.5)$ (i.e., the middle of the three horizontal dashed lines shown in Fig. 2). The reason to chose such a small error probability corresponding to $E_r \sim 0.006$ for the threshold studies will become clear in the discussion below. The time t in Fig. 3 is that of the pulse width used in obtaining data such as in Fig. 2. The distribution width for V_s is σ_s, calculated using the inverse of the slope in S_r versus V_{pulse} between $S_r = -2$ and -3 as shown in Fig. 2. Figure 3(b) shows the relative switching threshold voltage distribution width σ_s/V_s, for the same five devices as in Fig. 2. The thick solid lines in Fig. 3(b) refer to two different model scenarios, and will be discussed below.

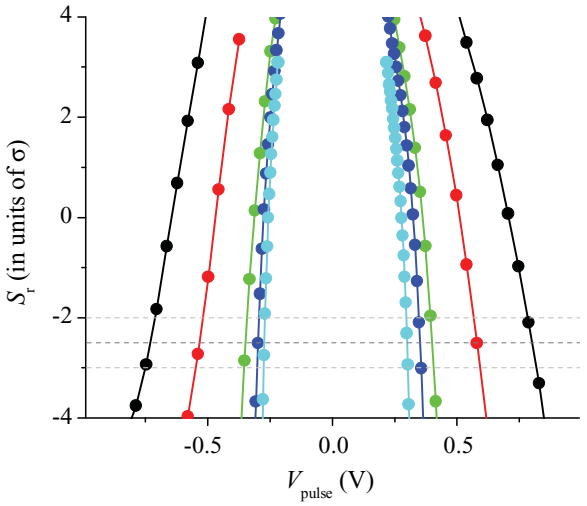

FIG. 2. (Color online) The normal quantile values S_r for the bit error rate (BER) of a spin-torque-driven MTJ for pulse widths of 2 (in black), 5, 50, 400, and 3200 (in cyan) ns. Shorter pulses require a higher voltage to achieve the same switching probability. Points are data, lines are guide to the eye. BER (bit-error-rate) represents cumulative nonswitching probability in unit of the apparent standard deviation σ, with zero corresponding to 50% nonswitching probability. Positive voltage is for parallel to antiparallel switching. The dashed lines indicate the location where the switching threshold and transition width are analyzed, as discussed in the text.

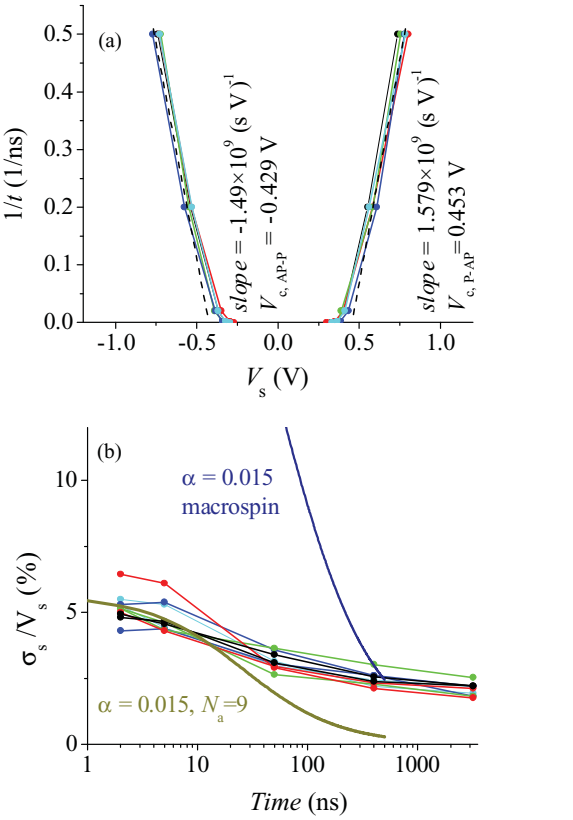

FIG. 3. (Color online) (a) The threshold voltage V_s and switching speed (defined as 1/switching time) for five MTJ devices. The threshold was defined at a probability for NOT switching of 0.006 (BER at -2.5σ), as shown in Fig. 2. (b) the relative switching width, as defined by the ratio of the inverse slope within -2 to -3 σ as shown in Fig. 2 normalized to the threshold voltage at -2.5σ. Heavy solid lines in (b) represent the macrospin and the simple N_a region subvolume excitation model's behavior in relative distribution width compared with measurement data. For both models the curves are generated using short-time asymptotes. Thus model curves should only be compared with data at the short-time limit.

III. COMPARISON TO A MACROSPIN MODEL

The dynamics of a macrospin under spin-torque-driven magnetic reversal has been extensively studied. In the limit of uniaxial anisotropy only and with finite temperature at large drive amplitude $I \gg I_{c0}$, with I being the current passing through the junction, I_{c0} the zero-temperature spin-torque current instability threshold [$I_{c0} = V_{c0}/R_p$ where R_p is the MTJ's parallel state resistance,[34] and V_{c0} the corresponding bias-voltage instability threshold as described below by Eq. (6)], the "long-time" super-threshold asymptotic form for the probability of *not* switching at time t can be expressed as[35,36]

$$E_r(t) \approx \left(\frac{\pi^2\xi}{4}\right)\exp\left(-\frac{2t}{\tau_I}\right) + O\left[\exp\left(-\frac{\pi^2\xi}{4}\right)\right], \quad (\text{for } I \gg I_{c0},\ E_r \ll 1 \text{ and } \xi \gg 1), \tag{2}$$

where $\xi = mH_k/2k_BT$ is the normalized thermal activation energy barrier height, m the total magnetic moment of the free layer, and H_k the uniaxial anisotropy field. $\tau_I = \tau_0/(I/I_{c0} - 1)$ is the characteristic time scale for spin-torque-induced reversal, $\tau_0 = 1/\gamma H_k\alpha$ is the natural unit of time with $\gamma \approx 2\mu_B/\hbar$ as the gyromagnetic constant, and α the LLG damping coefficient. A corresponding asymptote for the subthreshold regime ($I \ll I_{c0}$) can also be found[37]

$$E_r(t) \approx \exp\left\{-\left(\frac{t}{\tau_A}\right)\exp\left[-\xi\left(1-\frac{I}{I_{c0}}\right)\right]\right\}, \quad (\text{for } I \ll I_{c0},\ E_r \sim 1, \text{ and } \xi \gg 1), \tag{3}$$

where $\tau_A \approx \pi\hbar/\mu_B H_k$ is the inverse attempt frequency. A quantitative assessment of the subthreshold thermal activation process related to Eq. (3) is more complex than that of Eq. (2) for several reasons. First, the asymptotic region for Eq. (3) is more difficult to reach in experiments and for numerical simulations. In practice, due to the limited time of the measurement or numerical integration length, one is forced to work in the region where E_r is not sufficiently close to 1, and the condition $I \ll I_{c0}$ is usually not satisfied for a linear expansion. Second, the general shape of the energy landscape, especially that near the top of the energy potential during an activated reversal, could affect the details of the spin-torque dependence Eq. (3). Without precise knowledge of this it is difficult to model experimental systems with sufficient accuracy. Third, the time scale involved in measurements reported in Fig. 2 only extends out to about 3.2 μs, which might not be long enough to establish a steady-state initial condition upon which a simple thermal activation type of model is formulated—a concern especially when I approaches I_{c0} as the apparent damping constant diminishes. In consideration of these factors, for comparison between experimental observations and results from macrospin models, one concentrates

below on the supercritical threshold statistics where $I \gg I_{c0}$, corresponding to Eq. (2), which involves relatively short time scales of a few nanoseconds.

Equations (2) and (3) are asymptotes, and do not faithfully reproduce the macrospin behavior when $E_r \sim 1/2$ where most experimental results were examined in the literature. It would be more reliable if one compared the switching distribution function with these asymptotes within their respective limits of applicability. This is done for data shown in Fig. 2 near $E_r \sim 0.006$, or on normal quantile y scale at $S_r \sim -2.5$, as shown by the center dashed line in Fig. 2. The threshold voltage and distribution thus extracted are shown in Fig. 3. A linear expansion of Eqs. (1) and (2) at -2.5σ (in reference to plots such as Fig. 2) gives this slope of $1/t$ versus threshold V_s shown in Fig. 3 as

$$\frac{1}{t} = \left[\frac{1}{(m/\mu_B)(e/\eta)}\right]\left(\frac{2}{\ln(397.3\xi)}\right)\frac{V_s - V_{c0}}{R_p}, \tag{4}$$

where t is the pulse width. The relative distribution width for V_s can be obtained from the local slope near the expansion point

$$\frac{\sigma_s}{V_s} = \frac{2.8227}{2\alpha\gamma H_k t + \ln(397.3\xi)}. \tag{5}$$

For an MTJ involving symmetric electrodes and at low bias (a somewhat crude approximation but sufficient to provide a semi-quantitative guidance to our problem at hand)[34,38]

$$V_{c0} = \left(\frac{4e}{\hbar}\right)\left(\frac{\alpha}{\eta}\right)R_p(\xi k_B T), \tag{6}$$

with R_p the MTJ's parallel state low-bias resistance, and $\eta = \sqrt{m_r(m_r+2)}/2(m_r+1)$, with $m_r = (R_{ap} - R_p)/R_p$ as the magnetoresistance of the MTJ.

The selection of $E_r \sim 0.006$ is somewhat arbitrary, the consideration being it should be small enough to allow a reasonable comparison with Eq. (2), but sufficiently large to avoid excessive amounts of averaging at the expense of increased measurement and/or simulation time. At $E_r \sim 0.006$, a direct comparison with numerical simulation gave the accuracy for slope of $1/t$ versus V_s and its relative width estimate to better than ~5% in the short time limit for a macrospin model with similar parameter sets as in these experiments.

Using such a parameter extraction procedure, one estimates that for junction devices shown in Fig. 3(a), the slope of $1/t$ versus V_s on average is about 1.53×10^9 $(\mathrm{s}\cdot V)^{-1}$, and the threshold voltage $V_{c0} \sim 0.441$ V. For the data shown in Fig. 3(b) at the shortest time scale of 2 ns, one gets for the relative distribution width $\sigma_s/V_s \approx 0.055$. From transport tunnel magneto-resistance (TMR) measurements $\eta \sim 0.36$, and $R_p \sim 850\ \Omega$. One also knows film thickness $d \sim 0.8$ nm, and a junction lateral size $a = b \sim 120$ nm.

With these values and with Eq. (4), one gets for the free layer $M_s \sim 277$ emu/cm^3. This value of M_s is slightly below the separately measured single-layer magnetization value for a PMA CoFeB film of this composition and thickness of about 400 emu/cc (Ref. 21). It appears macrospin model in the supercritical limit reasonably well accounts for the dependence of an MTJ's spin-switching speed on drive voltage amplitude, to within a factor of 2 of the corresponding M_s etsimate.

The threshold voltage comparison, however, yields significant discrepancy. The data shown in Fig. 3(a) (the intercept of the zero-speed line) give $V_{c0} \sim 0.44$ V. From separate time-dependent switching threshold measurements similar to those shown in Fig. 1(b), one has the junction's thermal activation barrier height $\xi \sim 45$ to 60. Given such V_{c0} and ξ, one gets an effective damping coefficient from Eq. (6) of the order $\alpha \sim 0.15$. That is five to ten times larger than independent measurements on CoFeB thin films would suggest.[20,39,40]

The distribution width of the threshold appears to contain even more discrepancy. Taking the shortest time limit within this experimental dataset in Fig. 3(b) so as to best satisfy the asymptote expression Eq. (2), one estimates the relative one-σ variation of the threshold voltage at the point of $E_r \sim 0.006$ to be about 5.5%. Comparing with the macrospin prediction of Eq. (5), and taking the experimentally estimated values of $H_k \sim 400$ Oe, the macrospin model would predict a much broader threshold distribution at 2 ns. Indeed it would require a nonphysical $\alpha \sim 1.5$ to give a 5.5% relative distribution width as shown in Figs. 2 and 3(b).

To summarize, the experiment agrees with macrospin in terms of switching speed versus drive voltage, Eq. (4), to well within a factor of 2, essentially confirming the relationship of charge-to-spin conversion as described by the low-bias tunnel junction model.[34,38] The effective spin-polarization, η, is within a factor of 2 in agreement with model expectations. The threshold voltage comparison of Eq. (6) is less satisfying, with the experiment giving ten times too high a threshold voltage as macrospin would predict for the observed thermal activation barrier height. The switching threshold distribution width at 2 ns gives the least satisfying agreement, with the experiment giving a significantly narrower distribution than the macrospin model prediction, requiring an unphysical 100 times increase in the effective damping coefficient in the macrospin model.

The deviation of experimental threshold voltage from macrospin can be further quantified by taking the ratio of the threshold voltage, or more directly the threshold current, and that of the thermal activation barrier height. For the macrospin model this ratio depends on very few things. From Eq. (6)

$$\frac{E_b}{I_{c0}} = \left(\frac{\hbar}{4e}\right)\left(\frac{\eta}{\alpha}\right), \tag{7}$$

where $E_b = \xi k_B T$ is the thermal activation energy, $I_{c0} = V_{c0}/R_p$ (Ref. 34). This ratio, for the typical values of $\eta \sim 0.36$ and $\alpha \sim 0.015$ would give a macrospin expectation value of about 1 $k_B T/\mu A$ at $T = 300$ K. Experimental ratios are typically smaller, although the results appear to approach that of the macrospin prediction as junction size is reduced,[41] as shown in Fig. 4. From this point of view, the apparent ten times overestimate of the damping constant based on the threshold Eq. (6) is equivalent to saying the efficiency ratio of E_b/I_{c0} is ten times below that of a simple macrospin prediction. Note the examination of switching speed relating to Fig. 3(a) indicates the η factor is approximately right. Therefore, this apparent inefficiency of spin torque in the expression of V_{c0} is not because of poor charge-to-spin conversion.

One could separately examine the junction-size dependence of the threshold current I_{c0} and that of E_b. For macrospin both quantities should scale linearly with junction area. However,

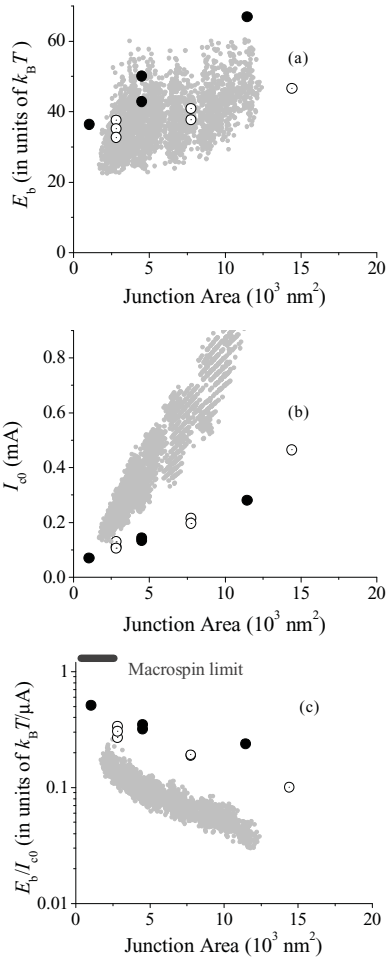

FIG. 4. Junction size dependence of E_b(a) and I_{c0}(b). All shapes are approximately circular with an aspect ratio equal to 1. Open and closed circles are data from two wafers measured manually using the protocol shown in Fig. 1. One sees here that while I_{c0} does appear to scale roughly with junction area, E_b shows a significant zero-area intercept. (c) The spin-torque efficiency factor plotted against junction area. Most of the large size devices have spin-torque efficiency factors well below that of the macrospin model value. The deviation becomes smaller, however, as junction sizes shrink. Also shown here in gray are measurements from a much larger set of junctions on a separate wafer with about two times lower junction RA and using a different measurement protocol involving fully automated on-chip CMOS-circuit-driven tests. Aside from a slope change attributable to the difference in junction RA, the basic scaling behavior remains the same. Part of the difference between the macrospin limit and the observed efficiency in (c) may be due to uncertainties in the estimate of E_b in absolute terms, as the measurements for E_b do not strictly satisfy $I/I_{c0} \ll 1$, and will suffer the problems discussed in Refs. 37 and 41. The CMOS-based measurement for E_b extends only to data in millisecond duration which is shorter than that of the manually measured data, consistent with the CMOS-based data giving a lower E_b and efficiency estimate on average.

experiments on these MTJs show only I_{c0} having a roughly linear scaling with junction area, while for junctions larger than about 40 nm in size, E_b is only weakly dependent on junction size, as shown in Fig. 4. This saturating size dependence of thermal activation energy is similar to what was reported by Rohart *et al.*[28] in their Co/Au PMA nano-islands and in other nanomagnetic systems.[26,27] The presence of such subvolume excitation suggests the apparent inefficiency of spin-torque originates from the involvement of nonmacrospin excitation and reversal. The sections below give a semi-quantitative subvolume excitation model that accounts for these observations.

IV. SIMPLE MODEL FOR SUBVOLUME SPIN-TORQUE-DRIVEN REVERSAL IN MTJS

Consider an MTJ whose thin free layer consists of N_a magnetic regions in the film plane that are fluctuating independently, as illustrated in Fig. 5. Assume that at large drive amplitude the reversal of the whole free layer follows that of the first region that develops a large amplitude instability for reversal. Assume further that all regions are identical and variation only results from temporal initial condition distribution due to thermal fluctuation. The subvolumes shown in Fig. 5 may be dynamic in their locations as well. Then, the various quantities for subvolume region would read

$$\begin{cases} m_i = m/N_a, \\ \xi_i = \xi/N_a, \\ \eta_i = \eta, \\ \alpha_i = \alpha, \\ R_{pi} = R_p N_a, \\ H_{ki} = H_k, \\ \tau_{0i} = \tau_0 = 1/\gamma H_k \alpha, \end{cases} \tag{8}$$

where the subscripts "i" denote the ith region. Assume the probability of an individual region's reversal is governed by the same probability function as a macrospin, Eq. (2), namely the probability for not switching, is E_{ri} with

$$E_{ri}(I_{si}) = \left(\frac{\pi^2 \xi_i}{4}\right) \exp\left[\left(-\frac{2t}{\tau_{0i}}\right)(I_{si}/I_{ci} - 1)\right], \tag{9}$$

where I_{si} is the current passing through the subvolume, $I_{ci} = V_{ci}/R_p$ represents the characteristic threshold current for the individual subvolume. The total BER for the whole free layer therefore would be

$$E_m = E_{ri}^{N_a}. \tag{10}$$

Consider the threshold position and distribution at the -2.5σ point, that is, for $E_m \to E_{mc} = \frac{1}{2}[1 - \mathrm{erf}(5/2\sqrt{2})] \sim 0.006$. The threshold I_{si} is simply the solution to Eqs. (9) and (10) when setting $E_{ri} = (E_{mc})^{1/N_a}$. A linear expansion of the normal quantile function Eq. (1) for $E_m(I_{si})$ versus I_{si} gives the one-σ switching width at $S_m = -2.5$ as $\sigma_{Isi} = -dI_{si}/dS_m$, or $1/\sigma_{Isi} = a_0(2N_a t/I_{ci}\tau_{0i})$ where the numerical prefactor is $a_0 = \sqrt{\pi/2}\ \mathrm{efrc}(5/2\sqrt{2}) \exp[(5/2\sqrt{2})^2] \approx 0.354\,265$. Note that for individual subregions $V_{si} = V_s = I_{si}R_{pi} = I_{si}(R_p N_a) = (I_{si}N_a)R_p = I_s R_p$, and $\sigma_{si} = \sigma_s = \sigma_{Isi}R_{pi} = \sigma_{Isi}(R_p N_a)$, and therefore

$$\begin{cases} V_s \approx V_c\left\{1 + \left(\frac{\tau_0}{2t}\right)[\ln((\pi/2)^2\xi_i) + 5.0814/N_a]\right\}, \\ \dfrac{\sigma_s}{V_s} \approx \dfrac{1}{N_a} \dfrac{2.82274}{2\gamma H_k \alpha t + \ln((\pi/2)^2\xi_i) + 5.0814/N_a}, \end{cases} \tag{11}$$

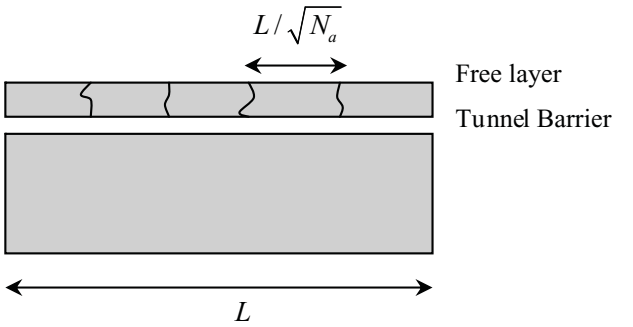

FIG. 5. A side-view illustration of the N_a unit subvolume excitation model structure. The free layer is assumed to fluctuate magnetically independently with N_a regions, with the average size of such region being $L/\sqrt{N_a}$. The film is assumed to be sufficiently thin so the magnons perpendicular to the film surface do not participate in the dynamics of interest here. This structure gives the scaling relationships of Eq. (8).

where

$$V_c = V_{ci} = \left(\frac{2e}{\hbar}\right)\left(\frac{\alpha_i}{\eta_i}\right) R_{pi}\,(2\xi_i k_B T) = \left(\frac{2e}{\hbar}\right)\left(\frac{\alpha N_a}{\eta}\right) R_p(2\xi_i k_B T). \tag{12}$$

The first relation in Eq. (11) results in switching speed versus write voltage amplitude relation as

$$\left(\frac{1}{t}\right) \approx \frac{\eta}{e(m/2\mu_B)[\ln((\pi/2)^2\xi_i) + 5.0814/N_a]} \frac{V_s - V_c}{R_p}. \tag{13}$$

Equation (12) would rescale the apparent spin-torque efficiency factor Eq. (7) for the whole MTJ by a factor of $1/N_a$

$$\frac{E_b}{I_{c0}} = \frac{1}{N_a}\left(\frac{\hbar}{4e}\right)\left(\frac{\eta}{\alpha}\right). \tag{14}$$

From the data shown in Figs. 3(a) and 3(b), $V_c \approx 0.44V$, $\delta(1/t)/\delta V_s \approx 1.53 \times 10^9\ 1/(\mathrm{s}\cdot V)$. From separate measurements one has $R_p \approx 850\ \Omega$, $\xi_i \approx 60 k_B T$, $H_k \sim 400$ Oe, and $\eta \approx 0.36$ from MR. The device size is about $120 \times 120\ \mathrm{nm}^2$. Solving for α and N_a using Eq. (12) and the second Eq. (11) at $t = 2$ ns with $\sigma_s/V_s \sim 0.055$ as read from Fig. 3(b), one obtains $N_a \approx 9$, and $\alpha \approx 0.015$. This puts the subvolume size on the order of $120/\sqrt{9} \approx 40$ nm. From this one further deduces a value of $M_s \approx 501$ emu/cc, using the experimentally measured $(1/t)$ versus V_s slope in Eq. (13). This M_s value is basically in agreement with the macroscopically measured thin film value (of about 400 emu/cm^3 at that thickness).[21]

With this simple subvolume activation model, one can account for most of the experimental observations. The subvolume analysis removes the ten times and 100 times inconsistency of the damping parameter α for matching macrospin and experimental observations of V_{c0} and relative distribution width. One also arrives at an estimate of the subvolume size on the order of 40 nm. This size is consistent with the observation of Ikeda *et al.*'s experimental observation at 40 nm,[20] and with the micromagnetic simulation results by the Toshiba group (Ref. 3), showing a gradual improvement of the spin-torque efficiency factor E_b/I_{c0} as junction sizes approach a similar lengthscale.

V. ORIGIN OF SUBVOLUME EXCITATION AND ACTIVATION

The above-detailed examination of experimental observations suggests a subvolume excitation and reversal process for spin-torque-induced magnetic switching in these thin film MTJ free layers. It poses the question of what might be the reason for such subvolume excitation, and what determines the size of such subvolume regions. In this section a semi-quantitative estimate is given to the origin of such subvolume excitations.

Knowing that most quantities discussed thus far depend sensitively on thermal fluctuation, and separate, low-bias field-swept measurement of the thermal activation energy in such MTJs reveals[42] similarly small thermal activation energy, more consistent with a subvolume thermal activation process than with macrospin, it is reasonable to start the discussion by trying to establish an adequate description of finite temperature magnetic fluctuation in such thin film structures.

For a simple macrospin at thermal equilibrium and with $\xi \gg 1$, it is straightforward to show with Boltzmann energy distribution that the self-correlation amplitude of the magnetic moment is

$$\langle m_v^2 \rangle = \frac{1}{2\xi} = \left(\frac{k_B T}{M_s H_k}\right)\frac{1}{V}, \tag{15}$$

where $v = x, y$ indicates the fluctuating component of the magnetic moment in the plane perpendicular to its easy axis direction $\mathbf{e_z}$, and V is the volume of the macrospin.

For nonmacrospin magnetic thin film at thermal equilibrium, follow the conventions of the authors of Ref. 43, one may write the correlation function of a magnetic moment varying both in time and space in the form of

$$\langle m_\mu(\mathbf{r},t) m_\nu(0,0)\rangle = \frac{2\gamma k_B T}{M_s V}\Sigma_{\mathbf{k}} \exp(-i\mathbf{k}\cdot\mathbf{r}) \int_0^\infty \frac{d\omega}{2\pi}\exp(-i\omega t) \times \left(\frac{\chi_{\mu\nu} - \chi^*_{\nu\mu}}{i\omega}\right)\frac{\hbar|\omega|/(k_B T)}{\exp[\hbar|\omega|/(k_B T)] - 1}, \tag{16}$$

where the susceptibility matrix reads

$$\chi = -\frac{1/(1+\alpha^2)}{(\omega - \omega_k^+)(\omega - \omega_k^-)}\begin{bmatrix} \omega_k - i\alpha\omega & -i\omega \\ i\omega & \omega_k - i\alpha\omega \end{bmatrix}, \tag{17}$$

with $\omega_{\mathbf{k}}^{\pm} = \mp\omega_{\mathbf{k}}/(1 \pm i\alpha)$, $\omega_{\mathbf{k}} = \omega_0 + (D/\hbar)|\mathbf{k}|^2$, and $\omega_0 = \gamma H_k$, where $D = (4\mu_B/M_s)A_{ex}$ is the exchange stiffness constant, and A_{ex} the exchange energy. The diagonal elements in Eq. (16) gives the autocorrelation amplitude

$$\langle m_v^2 \rangle = \left(\frac{k_B T}{M_s H_k}\right)\frac{1}{V_a}, \tag{18}$$

where V_a can be viewed as an effective fluctuation volume. Assume a *thin* film with finite lateral size $L \times L$. Further assume magnons with a wave vector perpendicular to the film surface are sufficiently high in energy so as to not give large thermal fluctuation amplitude in m_v, which is to say $2\pi\sqrt{D/k_B T} \gg d$, with d being the film thickness. Then for

the high temperature limit $k_BT \gg \mu_B H_k$, an evaluation of Eq. (16) gives

$$V_a \approx dL_D^2\left[\left(\frac{L_D}{L}\right)^2 + \frac{2}{2(L/L_D)^2+\pi} + \frac{8}{4(L/L_D)^2+\pi} + \ln\frac{(k_BT/2\mu_BH_k)}{1+(\pi/4)(L_D/L)^2}\right]^{-1}, \quad (19)$$

where $L_D = \sqrt{2\pi D/\mu_B H_k}$. For materials such as $Co_{72}Fe_{18}B_{10}$ similar to our free layer, $D \sim 0.37$ eVÅ^2 (Ref. 40). Assume this value also applies to our free layer, then for a range of $H_k \sim 0.5$ to 2 kOe, one has $L_D \sim 100$ to 45 nm. This lengthscale is similar to our junction size range. The V_a expression in Eq. (19) appears to underestimate the experimentally observed values by about two times, perhaps due to short-wavelength and/or large-amplitude treatments of the fluctuation-dissipation identity. Even so, it does give a good estimate of the crossover size for lateral confinement. As a function of lateral size L, the value of V_a saturates to an L-independent value when $L > L_D$.

A simple interpretation of Eq. (19) is the following. There are two sets of degrees of freedom involved in thermal agitation of a laterally confined thin film structure. One set involves the mean fluctuation identical for all moments, of the amplitude from easy axis, denoted as $\langle\theta_0^2\rangle$. The other sets the excitation amplitude $\langle\theta_L^2\rangle$ of the first nonzero wave number spin wave with the half wavelength of the order of lateral confinement L. The total local fluctuation amplitude is the sum of these two components. Each degree of freedom gets $k_BT/2$ energy (equal partition). The amplitude of each depends on the energy scale, with H_k determining that for $\langle\theta_0^2\rangle$, and $D/(2\mu_BL^2)$ determining that for the spin wave, $\langle\theta_L^2\rangle$. When size L is small, $D/(2\mu_BL^2) \gg H_k$, and the amplitude of the spin wave $\langle\theta_L^2\rangle$ is small compared to $\langle\theta_0^2\rangle$, such a system appears like a macrospin. When this relation is no longer valid, spin-wave amplitude becomes appreciable, and the system's net fluctuation appears larger than that of the macrospin. The crossover dimension is $\sim L_D$. When the lateral size is larger than L_D, thermal activation is subvolume, with the activation energy of the order

$$E_{bi} \sim \tfrac{1}{2}M_sH_kdL_D^2 \sim 4\pi A_{ex}d, \quad (20)$$

where the substitution of $D = (4\mu_B/M_s)A_{ex}$ was made. Thus for a typical set of parameters such as $A_{ex} \sim 2.84 \times 10^{-6}$ erg/cm (Ref. 40), and $d \sim 10$ Å, one has $E_{bi} \sim 86k_BT$ for $T = 300$ K, in reasonable agreement with our experimental observations to within a factor of 2. In the case of such subvolume excitation, this simple picture points out that the relevant energy scale is that of the exchange stiffness/film thickness product, and not the total anisotropy energy. This is consistent with the data summarized in Fig. 4.

More importantly, a subvolume below full sample size as described by Eq. (19) describes the enhanced amplitude of thermal agitation. This amplitude is the *initial condition* when the pulsed spin torque arrives. It determines the subsequent spin-torque-driven dynamic reversal. This is the real physical meaning of the parameter N_a in the previous section, and V_a here. These are not static subvolumes, but a dynamically enhanced fluctuation with an amplitude equivalent to a fluctuating macrospin of volume V_a. In reality, of course, there will be correlation effects and the correlation would determine the precise spatial and temporal dynamics of the reversal. Therefore, this simple model is only semi-quantitative. The boundary conditions for magnons were not treated accurately, nor were the large amplitude processes involved in thermal and especially spin-torque-driven dynamics. Also, dipolar interaction is not included in this linearized LLG treatment.

In thin film nanomagnets with in-plane magnetization, a similar crossover from macrospin-like to a vortex-like reversal mode has been experimentally observed under spin-torque drive,[44] and may involve a fundamentally similar mechanism to the subvolume excitation discussed above. However, quantitative treatments in such cases would need to include in-plane dipole contributions because shape anisotropy often is the leading order uniaxial anisotropy determining thermal activation barrier in that case. A nonuniform dynamic switching behavior has also recently been observed for PMA devices with about 100 nm in short-axis size and with elongated shape.[45] It will be interesting to further examine the relationship between the subvolume thermal excitation discussed here, and the details of such dynamic reversal modes.

The detailed reversal pathways in phase space, such as whether it involves a domain-wall structure or is via a vortex curling process[46] cannot be adequately addressed without including the full dipolar interaction, which tend to become more significant at larger sample sizes and/or high aspect ratios. These considerations would confine the applicability of this simple model to samples with small lateral dimensions not much larger than the exchange lengths of the thin film in question where exchange interaction is expected to dominate that over dipolar effects.[47,48]

Simple as it is, this model does capture the fundamental mechanisms at play that give rise to the appearance of a subvolume thermal activation. In addition to the experimental observations discussed here, this model description is also consistent with the report from other works,[26–28] and it agrees with the earlier observation that spin-wave excitation can result in subvolume thermal activation and reversal.[28] Further, from this simple model one could estimate the limit of thermal activation volume (and activation energy) as it relates to exchange A_{ex} and film thickness d through Eq. (20). This is not only important for fundamental understandings but for practical design of spintronics materials and devices as well.

VI. CONCLUSION

Examination of experimental results of spin-torque switching in MTJs suggest the presence of subvolume thermal excitation in most device sizes examined to date. The subvolume excitation picture naturally explains the departure of observed switching speed and threshold voltage from simple macrospin predictions. For a typical CoFeB-based MTJ with a lateral size ranging from 50 to 150 nm, a subvolume lateral size of the order of 40 nm was deduced by comparison of a simple subvolume activation model with experiment.

Subvolume excitation reduces the effective spin-torque efficiency because it reduces the effective thermal activation energy without correspondingly reducing the total spin-torque required for magnetic switching. For applications requiring

efficient operation of spin-torque-induced magnetic switching such as spin-torque-based magnetic memory, one would need junction free-layer materials that preserve large exchange energy A_{ex}, and the device needs to be patterned down to a size below that of the subvolume size, roughly of the order of $L_D = \sqrt{2\pi D/\mu_B H_k}$. This dimension, for $A_{ex} \sim 10^{-6}$ erg/cm, $M_s = 10^3$ emu/cm^3, and $H_k = 1$ kOe, is around 50 nm.

ACKNOWLEDGMENTS

This work was supported under a joint development program between IBM Research and MagIC Technologies, a subsidiary of TDK Corp. The authors gratefully acknowledge many fruitful discussions with our collaborators at MagIC Technologies. The teams at IBM Research's Materials Research Lab at Yorktown were instrumental in providing device fabrication support which made these experiments possible. J. Z. Sun also acknowledges many discussions with Andrew Kent, Daniel Bedau, Huanglong Liu, and many others at New York University on the subject of PMA spin-torque switching dynamics.

APPENDIX: SWEEP-CURRENT MEASUREMENT OF THRESHOLD CURRENT AND THERMAL ACTIVATION ENERGY

Assume the thermal activation transition rate can be expressed as[37]

$$\gamma_{(1,2)} = \gamma_{0,(1,2)} \exp\left[-\frac{E_b}{k_B T}\left(1 - \frac{H}{H_k}\right)^2 \left(1 - \frac{I}{I_{c(1,2)}}\right)\right], \tag{A1}$$

where (1,2) represent the forward and backward transitions. For large E_b the forward transition rate dominates, $\gamma_1 \gg \gamma_2$. Writing the forward transition probability at any given time t as $P(t)$, one has

$$\frac{dP}{dt} = -\gamma_1 P. \tag{A2}$$

Assume the current $I = I(t)$ is linearly time dependent with a current sweep rate of r_I, such that $I(t) = r_I t$. The maximum probability for switching is at the point $t = t_{\max}$ where $\frac{d^2P}{dt^2}|_{t=t_{\max}} = 0$. This gives for $t_{\max}$ an equation $(\frac{d\gamma_1}{dt}) - \gamma_1^2 = 0$, or

$$\gamma_1 = \gamma_{0,1} \exp\left[-\xi\left(1 - \frac{I}{I_{c1}}\right)\right], \tag{A3}$$

with $\xi = E_b/k_B T$ being the normalized activation energy, measured in $k_B T$. For a triangular sweeping current with constant rate, one has $r_I = \alpha_I I_p f$, where α_I is a numerical constant, I_p the peak value of the current being swept, and f the frequency of the sweep. This together with Eq. (A3) gives

$$\xi = \left(\frac{I_{c1}}{\log_{10} e}\right) \frac{1}{(\delta I/\delta \log_{10} f)}. \tag{A4}$$

Assuming I_{c1} can be estimated by the intercept of $I(t)$ versus $\log_{10} f$ data at a given frequency value, say around 10^9 Hz, one thus arrives at an estimate of the thermal activation barrier ξ. This is the method used in estimating ξ from the data shown in Fig. 1(c).

[1] M. Hosomi, H. Yamagishi, T. Yamamoto, K. Bessho, Y. Higo, K. Yamane, H. Yamada, M. Shoji, H. Hachino, C. Fukumoto, H. Nagao, and H. Kano, IEEE IEDM 2005. IEEE 0-7803-9269-8/05 (2005).

[2] T. Kawahara, R. Takemura, K. Miura, J. Hayakawa, S. Ikeda, Y. Lee, R. Sasaki, Y. Goto, K. Ito, T. Meguro, F. Matsukura, H. Takahashi, H. Matsuoka, and H. Ohno, IEEE J. Solid-State Circuits **43**, 109 (2008).

[3] T. Kishi, H. Yoda, T. Kai, T. Nagase, E. Kitagawa, M. Yoshikawa, K. Nishiyama, T. Daibou, M. Nagamine, M. Amano, S. Takahasi, M. Nakayama, N. Shimomura, H. Aikawa, S. Ikegawa, S. Yuasa, K. Yakushiji, H. Kubota, A. Fukushima, K. Oogane, T. Miyazaki, and K. Ando, in *Electron Devices Meeting (IEDM), IEEE International, Digital Object* (2008), p. 1–4.

[4] S. Matsunaga, J. Hayakawa, S. Ikeda, K. Miura, H. Hasegawa, T. Endoh, H. Ohno, and T. Hanyu, Appl. Phys. Exp. **1**, 091301 (2008).

[5] S. Fukami, T. Suzuki, K. Nagahara, N. Ohshima, Y. Ozaki, S. Saito, R. Nebashi, N. Sakimura, H. Honjo, K. Mori, C. Igarashi, S. Miura, N. Ishiwata, and T. Sugibayashi, in *Proceedings of the Symposium on VLSI Technology Digest of Tech. Papers* (IEEE, New York, 2009), p. 230.

[6] E. Chen, D. Apalkov, Z. Diao, A. Driskill-Smith, D. Druist, D. Lottis, V. Nikitin, X. Tang, S. Watts, S. Wang, S. A. Wolf, A. W. Ghosh, J. W. Lu, S. J. Poon, M. Stan, W. H. Butler, S. Gupta, C. K. A. Mewes, T. Mewes, and P. B. Visscher, IEEE Trans. Magn. **46**, 1873 (2010).

[7] J. Z. Sun, Phys. Rev. B **62**, 570 (2000).

[8] J. Z. Sun, T. S. Kuan, J. A. Katine, and R. H. Koch, Proc. SPIE **5359**, 445 (2004).

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

[10] Z. Li and S. Zhang, Phys. Rev. B **69**, 134416 (2004).

[11] Z. Li, J. He, and S. Zhang, Phys. Rev. B **72**, 212411 (2005).

[12] P. B. Visscher and D. M. Apalkov, J. Appl. Phys. **99**, 08G513 (2006).

[13] P. B. Visscher, Proc. SPIE **7036**, 70360B (2008).

[14] R. Zhu and P. B. Visscher, J. Appl. Phys. **103**, 07A722 (2008).

[15] H. Meng and J.-P. Wang, Appl. Phys. Lett. **88**, 172506 (2006).

[16] S. Mangin, D. Ravelosona, J. A. Katine, M. J. Carey, B. D. Terris, and E. E. Fullerton, Nat. Mater. **5**, 210 (2006).

[17] S. Mangin, Y. Henry, D. Ravelosona, J. A. Katine, and E. E. Fullerton, Appl. Phys. Lett. **94**, 012502 (2009).

[18] D. Bedau, H. Liu, J.-J. Bouzaglou, A. D. Kent, J. Z. Sun, J. Katine, E. E. Fullerton, and S. Mangin, Appl. Phys. Lett. **96**, 022514 (2010).

[19] D. Bedau, H. Liu, J. Z. Sun, J. A. Katine, E. E. Fullerton, S. Mangin, and A. D. Kent, Appl. Phys. Lett. **97**, 262502 (2010).

[20] S. Ikeda, K. Miura, H. Yamamoto, K. Mizunuma, H. D. Gan, M. Endo, S. Kanai, J. Hayakawa, F. Matsukura, and H. Ohno, Nat. Mater. **9**, 721 (2010).

[21]D. C. Worledge, G. Hu, D. W. Abraham, J. Z. Sun, P. L. Trouilloud, J. Nowak, S. Brown, M. C. Gaidis, E. J. O'Sullivan, and R. P. Robertazzi, Appl. Phys. Lett. **98**, 022501 (2011).

[22]D. C. Worledge, G. Hu, P. L. Trouilloud, D. W. Abraham, S. L. Brown, M. C. Gaidis, J. Nowak, E. J. O'Sullivan, R. P. Robertazzi, J. Z. Sun, and W. J. Gallagher, International Electron Devices Meeting (IEDM) 2010, IEEE International (2010), pp.12.5.1–12.5.4.

[23]K. Yakushiji, T. Saruya, H. Kubota, A. Fukushima, T. Nagahama, S. Yuasa, and K. Ando, Appl. Phys. Lett. **97**, 232508 (2010).

[24]R. Beach, T. Min, C. Horng, Q. Chen, P. Sherman, S. Le, K. Yang, H. Yu, X. Lu, W. Kula, T. Zhong, R. Xiao, A. Zhong, G. Liu, J. Ken, J. Yuan, J. Chen, R. Tong, J. Chien, T. Torng, D. Tang, P. Wang, M. Chen, S. Assefa, M. Qazi, J. DeBrosse, M. Gaidis, S. Kanakasabapathy, Y. Lu, J. Nowak, E. O'Sullivan, T. Maffitt, J. Z. Sun, and W. J. Gallagher, International Electron Devices Meeting (IEDM) 2008, IEEE International (2008), pp. 1–4.

[25]T. Min, Q. Chen, R. Beach, G. Jan, C. Horng, W. Kula, T. Torng, R. Tong, T. Zhong, D. Tang, P. Wang, M. min Chen, J. Z. Sun, J. K. Debrosse, D. C. Worledge, T. M. Maffitt, and W. J. Gallagher, IEEE Trans. Magn. **46**, 2322 (2010).

[26]M. Bode, O. Pietzsch, A. Kubetzka, and R. Wiesendanger, Phys. Rev. Lett. **92**, 067201 (2004).

[27]S. Krause, G. Herzog, T. Stapelfeldt, L. Berbil-Bautista, M. Bode, E. Y. Vedmedenko, and R. Wiesendanger, Phys. Rev. Lett. **103**, 127202 (2009).

[28]S. Rohart, P. Campiglio, V. Repain, Y. Nahas, C. Chacon, Y. Girard, J. Lagoute, A. Thiaville, and S. Rousset, Phys. Rev. Lett. **104**, 137202 (2010).

[29]Y. Guan, J. Z. Sun, X. Jiang, R. Moriya, L. Gao, and S. S. P. Parkin, Appl. Phys. Lett. **95**, 082506 (2009).

[30]M. D. Mascaro and J. Z. Sun, J. Appl. Phys. **107**, 08C701 (2010).

[31]R. P. Robertazzi and J. Nowak, (unpublished).

[32]J. J. Nowak, R. P. Robertazzi, J. Z. Sun, G. Hu, D. W. Abraham, P. L. Trouilloud, S. Brown, M. C. Gaidis, E. J. O'Sullivan, W. J. Gallagher, and D. C. Worledge, IEEE Magnetics Letters **2**, 3000204 (2011).

[33]J. Chambers, W. Cleveland, B. Kleiner, and P. Turkey, *Graphical Methods For Data Analysis*, Wadsworth Statistics/Probability Series (Duxbury Press, Pacific Grove, CA, 1983).

[34]J. Z. Sun and D. C. Ralph, J. Magn. Magn. Mater. **320**, 1227 (2008).

[35]J. Z. Sun, IBM Internal Memo (2006).

[36]J. He, J. Z. Sun, and S. Zhang, J. Appl. Phys. **101**, 09A501 (2007).

[37]The detail of the exponent for the $(1 - I/I_{c0})$ term depends sensitively on the shape of the energy potential especially near the top of the barrier. A simple temperature-scaling argument for spin-torque gives an exponent of 1. For purely uniaxial anisotropy and in strictly collinear alignment, a Fokker-Planck solution can be derived which gives an exponent of 2 for it as well as for the magnetic field-dependence term inside ξ (Refs. 8–14). In either case the expression is only approximately true and is limited to $I/I_c \ll 1$. This is often difficult to access experimentally as it would involve a very long measurement time. One routinely uses expressions related to Eq. (3) for an estimate of the thermal activation barrier height ξ which tend to underestimate ξ when the measurement is not carried out to sufficiently long time to warrant the condition $I \ll I_{c0}$. By extending measurements out to beyond many seconds in time, and by confirming the consistency between ξ values obtained from spin-current sweep as well as from magnetic field sweep, one is led to believe the ξ values reported here are generally consistent enough so as to not affect any of the conclusions drawn from this paper.

[38]J. C. Slonczewski, Phys. Rev. B **71**, 024411 (2005).

[39]D. W. Abraham, (unpublished).

[40]C. Bilzer, T. Devolder, J.-V. Kim, G. Counil, C. Chappert, S. Cardoso, and P. P. Freitas, J. Appl. Phys. **100**, 053903 (2006).

[41]Here one concentrates first on the size-dependent trend of the efficiency factor. In some devices such as shown here, even the zero-size intercept in Fig. 4(c) falls below the macrospin expectation by perhaps a factor of 2. This is an indication of additional inefficiency mechanisms not yet fully identified. Part of this may relate to the underestimate of the thermal activation barrier E_b as discussed in Ref. 37.

[42]R. S. Beach, (unpublished).

[43]J. Xiao, G. E. W. Bauer, K. C. Uchida, E. Saitoh, and S. Maekawa, Phys. Rev. B **81**, 214418 (2010).

[44]Y. Acremann, J. P. Strachan, V. Chembrolu, S. D. Andrews, T. Tyliszczak, J. A. Katine, M. J. Carey, B. M. Clemens, H. C. Siegmann, and J. Stöhr, Phys. Rev. Lett. **96**, 217202 (2006).

[45]D. P. Bernstein, B. Bräuer, R. Kukreja, J. Stöhr, T. Hauet, J. Cucchiara, S. Mangin, J. A. Katine, T. Tyliszczak, K. W. Chou, and Y. Acremann, Phys. Rev. B **83**, 180410(R) (2011).

[46]R. Skomski, H.-P. Oepen, and J. Kirschner, Phys. Rev. B **58**, 3223 (1998).

[47]C. A. F. Vaz, J. A. C. Bland, and G. Lauhoff, Rep. Prog. Phys. **71**, 056501 (2008).

[48]C. L. Dennis, R. P. Borges, L. D. Buda, U. Ebels, J. F. Gregg, M. Hehn, E. Jouguelet, K. Ounadjela, I. Petej, I. L. Prejbeanu, and M. J. Thornton, J. Phys. Condens. Matter **14**, R1175 (2002).