PHYSICAL REVIEW B **96**, 064437 (2017)

Resistance-area product and size dependence of spin-torque switching efficiency in CoFeB-MgO based magnetic tunnel junctions

J. Z. Sun*

IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA

(Received 19 May 2017; revised manuscript received 8 August 2017; published 30 August 2017)

We report a systematic dependence of the spin-transfer torque (STT) switching threshold on junction size and its resistance-area product (RA, or r_{Amtj}) for the CoFeB-MgO-CoFeB type of magnetic tunnel juctions. The RA dependence of the switching efficiency is seen to become stronger for junctions of larger size (in the range of 15 to 50 nms). Here, the STT switching efficiency is defined as the ratio of the nanojunction free-layer energy barrier for thermal activation reversal to its STT switching threshold current. That is, the efficiency $\kappa = E_b/I_{c0}$, and it is seen to follow the junction r_{Amtj} in the form of $1/\kappa = A_0 + A_1/r_{\mathrm{Amtj}}$, with $A_1 \propto a - a_0$, where a is the device diameter and a_0 is of the order of 10 nm. The $1/r_{\mathrm{Amtj}}$ dependence is consistent with a tunnel-conductance-dependent spin-pumping-like action, although the spin-pumping mixing conductance originating from the r_{Amtj} of a magnetic tunnel interface is orders of magnitudes below that of the natural damping of the free layer. We postulate that the spin-torque-induced dynamics in these magnetic tunnel junctions must involve high-frequency spin dynamics near the tunnel interface beyond the average free-layer dynamics frequencies. Such interface high-energy process gives rise to a two-step process for spin-current transmission into the free layer and causes visible spin-pumping spin-current loss across a tunnel barrier.

DOI: 10.1103/PhysRevB.96.064437

I. INTRODUCTION

Spin-torque-driven magnetic reversal in MgO-based, perpendicularly magnetized magnetic tunnel juctions (MTJs) exhibits nonmacrospin behaviors from device sizes of 11–15 nm and up in diameter, with the apparent switching efficiency $\kappa = E_b/I_{c0}$ exhibiting an inverse diameter dependence [1–4]. This behavior of κ originates from a linear diameter scaling of E_b, which is consistent with an edge-demagnetization reduction's contribution, and from a quadratic diameter scaling of I_{c0}, whose origin beyond macrospin limit is not completely understood.

Another observation is a dependence of the efficiency κ on the MTJ's resistance-area (RA) product (r_{Amtj}). This observation has been reported before [5], although its origin is not clear. On one hand, the observed dependence resembles spin-pumping-like action across an interface with a mixing conductance related to the MTJ's RA product r_{Amtj}. On the other hand, the magnitude of r_{Amtj} in such MTJs is typically two to three orders of magnitudes too high to give noticeable change to the ferromagnetic free-layer damping of the order of $\alpha \sim 0.005$ or there about.

More recently, in closer examination of our MTJs with perpendicularly magnetized anisotropy (PMA), it is further noticed that the dependence of κ on junction r_{Amtj} is junction size dependent, with an empirical relationship of $1/\kappa = A_0 + A_1/r_{\mathrm{Amtj}}$, and $A_1 \propto a - a_0$, with $a_0 \sim 10$ nm.

A size-dependent κ is understood as a consequence of the nanomagnet being switched via dynamically nonmacrospin intermediate states. An RA dependence of κ is not expected from macrospin dynamics either, and its connection to dynamic inhomogeneity has not been established.

In this paper, we report our experimental findings on an MTJ's STT switching efficiency κ as it depends on the junction size *and* its RA. These observations point to the involvement of the tunnel-interface magnetic moment in high-energy dynamics. These interface-related high-energy processes are the leading cause for the observed RA dependence of κ.

II. EXPERIMENT

Tunnel junctions with MgO barriers and CoFeB as free layers with PMA were produced in ways similar to what has been reported earlier [6]. These include the sputter deposition, at ambient temperature, a CoFeB-based MTJ materials thin-film stack. The wafers are then post-deposition annealed in vacuum at 300 °C for 1 h prior to being lithographically patterned down to sizes ranging from about 15 nm to > 100 nm in diameter for circular-shaped devices in this series of samples. A reactive ion etch is used for the main junction etching step, followed by a low-energy (∼200 eV) grazing incidence Ar ion beam etch for trimming the junction sides to the desired dimensions. The finished structures are characterized for their spin-torque switching properties using methods described in Ref. [2]. The free-layer (FL) CoFeB is about 1.7 nm thick.

These junctions are used for measuring a set of size-dependent quantities including thermal activation energy E_b, spin-torque switching threshold voltage V_{c0}, spin-torque switching threshold current I_{c0}, as evaluated from V_{c0}/R_p for its being most representative of the associated spin current [7,8] and the switching efficiency $\kappa = E_b/I_{c0}$. The median junction sizes for each size bin, corresponding to a group of devices of identical mask dimension design, are obtained on the same wafer and same location using scanning electron microscope (SEM) on test sites prior to insulator encapsulation.

First, we concentrate on the average threshold I_{c0} and E_b behavior. Such I_{c0} values are measured using a time-dependent threshold evaluation method, namely, by measuring $I_{c\tau} = (|V_{c\tau,\mathrm{AP\text{-}P}}| + |V_{c\tau,\mathrm{P\text{-}AP}}|)/R_p$, where τ designates the voltage pulse width in time at which the threshold is measured. Here "P-AP" designates the direction for a parallel-to-antiparallel

*Corresponding author: jonsun@us.ibm.com

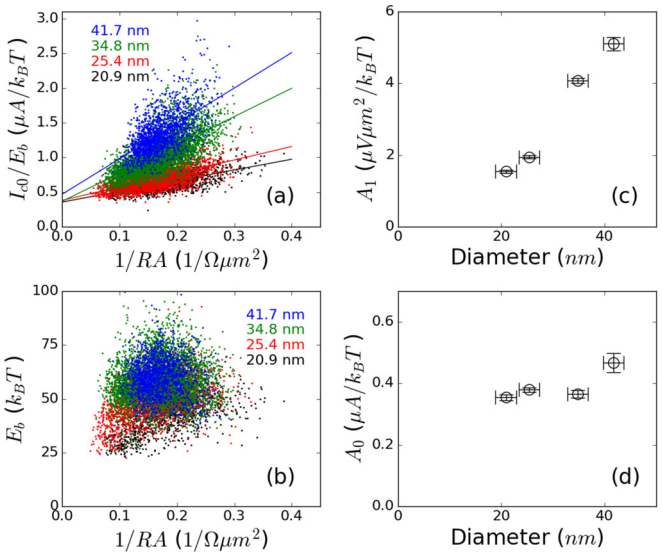

FIG. 1. Measured I_{c0}/E_b vs $1/r_{\mathrm{Amtj}}$ relationship of CoFeB-MgO tunnel junctions. (a) The dependence for junctions of four different size groups, whose median sizes are shown with same-color labels. Each data point is from one junction. Each size bin contains over 1000 devices. Lines are linear fits to the groups in the stated size bin. (b) The r_{Amtj} dependence comes primarily from I_{c0}, as E_b is shown here for the same data set to be not strongly dependent on r_{Amtj}. (c) A strong device diameter dependence is seen in the slope of I_{c0}/E_b vs $1/r_{\mathrm{Amtj}}$ from (a). (d) The intercept of I_{c0}/E_b vs $1/r_{\mathrm{Amtj}}$ fit is not as strongly size dependent. Error bars for A_0 and A_1 represent linear regression standard errors. Those for diameter reflect the bin width of data selection.

switch (in reference to the relative magnetic orientation of the free layer and the reference layer across the tunnel barrier), and "AP-P" for antiparallel-to-parallel switching. With a series of such pulse-width-dependent measurements, one could make a log-linear regression, from which the P-AP and AP-P averaged quantities of I_{c0} and E_b are deduced. The averaging of AP-P and P-AP results ensures the cancellation of the leading-order effects of external field bias on switching threshold V_{c0} and thermal-activation energy E_b because a bias-field-induced effect on I_{c0} and E_b would be antisymmetric for AP-P and P-AP. A detailed description of I_{c0} and E_b measurements in this experiment can be found in our earlier publications [1,2], based on a linearized, low-bias, symmetric MTJ model that associates the bias voltage to a spin current [7–10].

A set of experimentally measured switching efficiency I_{c0}/E_b vs $1/r_{\mathrm{Amtj}}$ data is shown in Fig. 1(a). Each data point represents a single junction's result. Different colors group junctions into different diameter size bins with a 4-nm bin width. For all devices included in this figure, the tunnel magnetoresistance (TMR) is above 80%, ensuring devices as being representative of "good" quality junctions free of gross defects in barriers, and with no significant degradation from fabrication. The TMR value also agrees with extended film's TMR estimate from current-in-plane tunnel (CIPT) measurements [11].

A sizable portion of raw data scatter in Fig. 1(a) is due to the measurement methodology. The pulse width versus switching threshold voltage for each individual pulse width was the average of only five pulse-height sweep repeats. This insufficient average resulted in a variation of reported E_b value with a 1σ variance greater than 6% for E_b values around 40 to 60 k_BT. This is simply due to temporal threshold variation of *individual switching events*. The limitation of the number of pulse repeats is so as to allow a reasonable throughput for a large number of junctions to be sampled in a development environment. The large number of junctions tested (over 1000 per size bin) as represented here gives another level of ensemble average, resulting in less net variance of the trend for size and RA dependence. This added averaging, however, is not perfect, as it mixes temporal and device-to-device variation. It is a compromise that grew out of our technology development environment. A more focused single-device detailed measurement is underway on selective wafers. These results when complete will be reported elsewhere.

From these data shown in Fig. 1, an empirical form can be established for I_{c0}/E_b vs $1/r_{\mathrm{Amtj}}$:

$$\frac{I_{c0}}{E_b} = A_0 + \frac{A_1}{r_{\mathrm{Amtj}}}. \tag{1}$$

As illustrated in Figs. 1(c) and 1(d), the value of the I_{c0}/E_b vs $1/r_{\mathrm{Amtj}}$ slope is here measured to be of the order of $A_1 \approx 4 \pm 0.07$ μV μm$^2/k_BT$ at a diameter around 35 ± 2 nm. The corresponding intercept is of the order $A_0 \approx 0.4 \pm 0.01$ μA$/k_BT$. Here, error-bar estimates are from standard error values of linear regression shown in Fig. 1(a).

Should the switching efficiency be in the macrospin limit [which it is not here for our data and for Eq. (1)], the expression I_{c0}/E_b could be further related to the fundamental materials properties in the form of [12]

$$\frac{1}{\kappa} = \frac{I_{c0}}{E_b} = \left(\frac{4e}{\hbar\eta}\right)\alpha, \tag{2}$$

where $\eta = \sqrt{m_r(m_r+2)}/2(m_r+1)$ is the spin-polarization factor[1] for a symmetric tunnel junction, $m_r = (R_{\mathrm{AP}} - R_{\mathrm{P}})/R_{\mathrm{P}}$ is the tunnel magnetoresistance (TMR).

In the context of Eq. (2), our experimental observation summarized in Eq. (1) could be viewed as a phenomenological damping coefficient of

$$\alpha_{\mathrm{eff}} = \alpha_0\left(1 + \frac{A_1}{A_0 r_{\mathrm{Amtj}}}\right) \tag{3}$$

which is exactly the form it would take if the added damping originates from a spin-pumping effect across a conducting interface [13,14]. In fact, an intercept $A_0 \approx 0.4$ μA$/k_BT$ would translate to an $\alpha_0 \sim 0.006$ according to Eq. (2) and assuming magnetoresistance $m_r \sim 1$, essentially agreeing with the pure materials damping as determined from coupon-film level FMR measurements, to well within a factor of 2.

[1]Note here η is the total spin polarization including both electrodes of the MTJ. It differs from assuming an incident charge current with fixed spin polarization independent of the relative angle between polarization of the current and the receiving electrode's magnetization. An in-depth discussion and the derivation of the η factor can be found in Refs. [7,9,10].

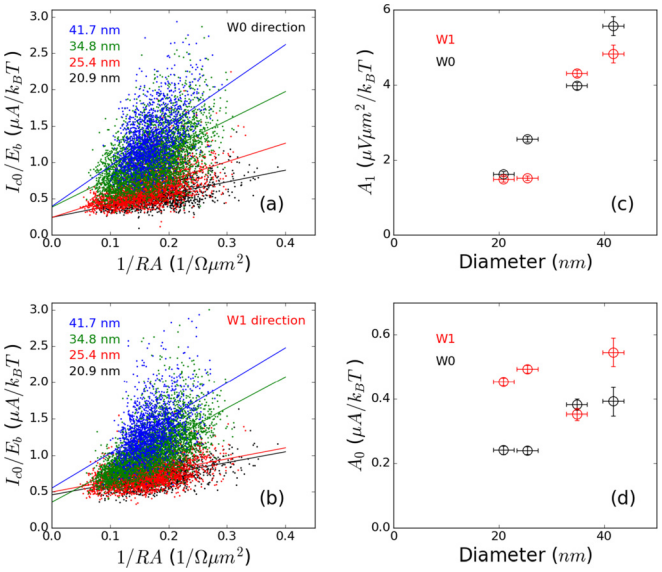

FIG. 2. The rotationally symmetric behavior of switching efficiency's r_{Amtj} and size dependence. (a), (b) The write-0 [corresponding to an antiparallel initial state between FL and RL (reference layer)] and write-1 (parallel initial state) direction's dependence of I_{c0}/E_b vs $1/r_{\mathrm{Amtj}}$. (c), (d) The resulting linear fit's slope and intercept from (a) and (b). This demonstrates the lack of dependence of A_0 and A_1 on relative orientations between the FL and RL. Note: in both directions, the data for I_{c0} plotted are obtained by V_{c0} across the MTJ divided by the *parallel state junction resistance*. This corresponds to the spin-current values in both directions, as the magnitude of a symmetric MTJ's spin current is directly proportional to bias voltage but not to the charge current [7–10].

The values of E_b and I_{c0} reported in Fig. 1 are averages of P-AP, or "write 1" (W1) in memory technology language, and AP-P (or W0) values. For reasons shown in Sec. IV B (reasons concerning the chirality-dependent dynamic spin-current coupling between the FL and its environment), it is important to also examine the individual direction's switching E_b and I_{c0}. This is shown in Fig. 2.

Data in Fig. 2 suggest the size dependence of A_1 in the above observation is essentially independent from the free layer and reference layer (FL-RL) relative orientation, whereas for A_0 there might be some orientation dependence, although the contrast over this size range is comparable to size-dependent data scatter. We consider the observed A_0 variation as inconclusive for precession-direction dependence at present. This is because, in addition to precession chirality dependence, the FL is also under an incompletely canceled dipolar field from the RL, whose magnitude as judged from the MTJ's resistance-field hysteresis loop is in the range of 100 to 800 Oe for this data set, depending on size and with significant scatter. A change of dipolar bias-field strength would also affect the switching threshold for the AP-P and P-AP branches. Given the lack of systematic size or switching-direction dependence of A_0, we conclude that, to the leading order, the precession-direction dependence of A_0 and A_1 in our data as described by Eq. (1) is not dominant.

To summarize, there are three main experimental findings, namely, (1) a linear dependence of inverse switching efficiency $1/\kappa$ on MTJ's specific tunnel conductance $1/r_{\mathrm{Amtj}}$, (2) a linear dependence of the slope A_1 on junction diameter a, and (3) the observation of (2) is independent from the relative orientation of FL-RL magnetic moment alignment. These three attributes are what we will focus on next to develop an understanding.

III. INTRODUCTION TO MODEL DEVELOPMENT

A large portion of the perpendicular magnetic anisotropy energy in this class of MTJs is believed to originate from the interface between the tunnel barrier MgO and the CoFeB "free layer" [15]. Magnetic moments situated at the MgO-CoFeB interface have a high concentration of anisotropy energy, most likely in areas coinciding with good electronic structure coupling into the MgO interface, and consequently high tunnel conductance. These same interface moments may also have a possible weakening of their exchange coupling to the rest of the FL ferromagnetic moments. Such weakening could occur either because of their partial oxidation state or because of the complex exchange interaction for interface moments. These interface moments have a high PMA energy density and a relatively weak exchange to the rest of the FL. This combination is likely to cause nonmacrospin dynamics at the interface [16–18]. It could, for example, cause the appearance of an apparent fourth-order magnetic anisotropy, reflecting the relative rotation between the magnetic moments in the interior of the FL and those with dominant PMA at the interface [16].

To distinguish the dynamics of these high-energy interface magnetic moments from the volume of the FL, a simplified model concept beyond macrospin is needed. We start with the next simplest possible construct of two coupled macrospins, or what we call an m_1-m_2 model. Here, two magnetic moments m_1 and m_2 are exchange coupled together with a finite exchange energy. Each is treated as a macrospin. In our specific situation, m_1 emulates the interface moment, whose magnitude is small, and it is situated in a strong anisotropy potential. m_2 emulates the rest of the magnetic moments in the FL, containing the majority of magnetic moments of the FL, but is with a relatively weak PMA, or even with an easy-plane anisotropy, if the demagnetization energy dominates. These two moments m_1 and m_2 couple via an exchangelike interaction. This exchange energy's relative strength to the anisotropy energies experienced by m_1 and m_2 leads to a wide range of dynamic behaviors. All dynamics are treated in the classical limit, by two coupled Landau-Lifshitz-Gilbert equations (the LLG equations).

By assuming m_1 representing the interface moment and its dynamics operating on a time scale much shorter than that of the rest of the FL, represented by m_2, it becomes possible to at least approximately use a separation of time-scales approach, and treat m_1's fast dynamics as adiabatic for the slow dynamics of m_2 and the full FL switching process.

In the following sections, we will first introduce a few topics related, some directly and others somewhat indirectly, to the m_1-m_2 two-macrospin model. These form the building blocks we then use to address the issue of spin-torque-driven magnetic instabilities and magnetic switching in such a system, and consequences brought on by the interface-concentrated PMA. This is an approach similar to what has been taken in Ref. [16].

IV. THREE MAIN FORMS OF SPIN CURRENT FOR A MACROSPIN NANOMAGNET

For a nanomagnet coupled to an external environment both magnetically (via an exchangelike or magnetic-field-like, force) and electronically (involving a spin-carrying detail-balanced transport current of mobile electrons), there are three main classes of interactions relevant to the dissipative dynamics of the magnet. These are (1) magnetic damping, (2) time-varying magnetic field (this could include a time-varying fieldlike exchange interaction), and (3) a spin-carrying electronic current. Below we give a brief description of these three types of spin currents.

For convenience of comparing to electrical transport-related experimental results, the spin-current expressions here are mostly written in equivalent charge-current unit. That is, a $\hbar/2$ per unit time spin current is converted into a charge current of one electron per unit time. This allows for the use of convenient units such as amperes and volts for easy comparison with transport variables. This is particularly helpful for describing spin-torque phenomena in magnetic tunnel junctions, spin valves, and in other electrical-transport-related situations.

A macrospin is defined as the limiting case of a small ferromagnetic body whose dimensions are well below that of the exchange lengths in the problem. Conveniently, it can be thought of as the limiting case where the ferromagnet's exchange energy $A_{\rm ex}$ is above any of the energy scales of concern in the problem.

The dynamics of a macrospin moment vector $\mathbf{m}$ can be described by the Landau-Lifshitz-Gilbert equation (the LLG equation) as

$$\frac{1}{\gamma}\frac{d\mathbf{n}_m}{dt} = \mathbf{H}_{\rm eff}\times\mathbf{n}_m + \frac{\alpha}{\gamma}\mathbf{n}_m\times\frac{d\mathbf{n}_m}{dt}, \tag{4}$$

where $\mathbf{n}_m = \mathbf{m}/|\mathbf{m}|$ is the unit vector of the moment direction, $\mathbf{H}_{\rm eff}$ is the total effective field including the anisotropy and applied field. α is the phenomenological LLG damping coefficient. $\gamma = |g|\mu_B/\hbar \approx 2\mu_B/\hbar$ is the magnitude of the electronic gyromagnetic ratio.

The left-hand side of Eq. (4) is an expression of torque Γ normalized by $m = |\mathbf{m}|$ in the form of Γ/m. This torque is composed of two terms shown on the right-hand side. The first term is energy conserving if $\mathbf{H}_{\rm eff}$ is static in time. The second term is energy nonconserving, and represents the spin current associated with the dissipative interaction via damping. In addition to damping, a time varying $\mathbf{H}_{\rm eff}$ can also bring on energy-nonconserving interactions with $\mathbf{m}$.

Magnitude wise, for most practical cases the energy-conserving part of the torque, associated with applied field or anisotropy field, is about a factor of $1/\alpha \sim 10^2$ larger than the damping torque Γ_α, i.e., the second term in Eq. (4).

A. Spin current related to damping

A torque on the macrospin is a spin-current flow. From Eq. (4), the damping-related spin current (torque), in unit of energy (such as [erg]), can be written as

$$\mathbf{I}_{s\alpha} = \frac{\alpha m}{\gamma}\mathbf{n}_m\times\frac{d\mathbf{n}_m}{dt} = \left(\frac{m\alpha}{g\mu_B}\right)\hbar\mathbf{n}_m\times\frac{d\mathbf{n}_m}{dt}, \tag{5}$$

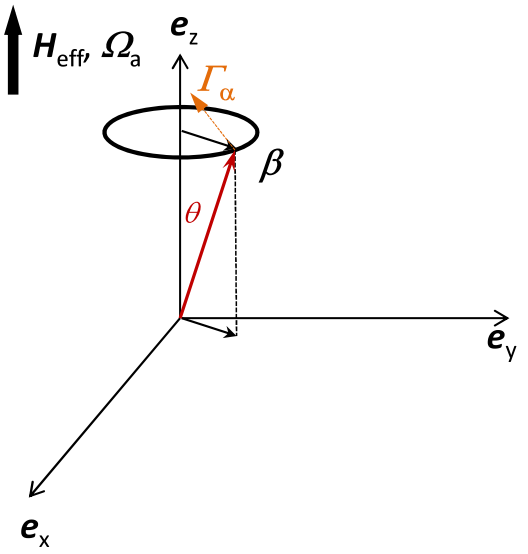

FIG. 3. An illustration to the vector definitions of the LLG equation (4). The magnetic moment unit vector makes a polar angle θ with the $\mathbf{e}_z$ axis. β is the projection of $\mathbf{n}_m$ onto the $x-y$ plane. The damping torque Γ_α is perpendicular to $\mathbf{n}_m$ and points to a direction that either increases or decreases the cone angle θ, depending on the signs of the spin and spin-current direction. The total effective field $\mathbf{H}_{\rm eff}$ is assumed to align in $\mathbf{e}_z$ direction in our models, same is the related angular velocity vector $\Omega_a = \gamma\mathbf{H}_{\rm eff}$ if the moment is in free motion.

where μ_B is the Bohr magneton and $g \sim 2$ is the magnitude of electronic Landé g factor.

Consider a moment in a uniaxial field $\mathbf{H}_{\rm eff} \parallel \mathbf{e}_z$ as in Fig. 3, the moment having an instantaneous cone angle of θ, and an angular veolcity $\Omega_a \parallel \mathbf{e}_z$, one gets a nonzero cycle-averaged spin current with the polarization direction along $\mathbf{e}_z$ as

$$I_{s\alpha} = |\langle\mathbf{I}_{s\alpha}\rangle| = \left(\frac{m\alpha}{g\mu_B}\right)(\hbar\Omega_a)\sin^2\theta. \tag{6}$$

For free damping motion with a static $\mathbf{H}_{\rm eff}$, one has $\Omega_a \to \Omega_H = \gamma|\mathbf{H}_{\rm eff}|$ as the moment-field system's intrinsic dynamics frequency, i.e., the moment's precession frequency.

Define a "voltage" $V_B = \hbar\Omega_H/e$ where e is the electron charge, one can rewrite Eq. (6) in a form similar to charge-current transport with a charge-current-equivalent unit describing the spin current. Using the conversion of $I_{s,\rm cg} = (2e/\hbar)I_s$, where I_s is in angular momentum flow unit (i.e., torque or energy), and $I_{s,\rm cg}$ is the same current but in charge-current unit (such as amperes) as indicated by the "cg" subscript, one has, for the on-resonance condition, the damping α-related spin current as

$$\begin{aligned} I_{s\alpha,\rm cg} &= \left[\left(\frac{e^2\epsilon_j}{\hbar}\right)\left(\frac{\alpha m}{\mu_B}\right)\right]V_B\sin^2\theta \\ &= \left(\frac{2e}{\hbar}\right)(\alpha m H_{\rm eff})\sin^2\theta \end{aligned} \tag{7}$$

or a spin-mixing conductance in charge unit of $g_\alpha = (e^2\epsilon_j/\hbar)(\alpha m/\mu_B)$, such that $I_{s\alpha,\rm cg} = g_\alpha V_B\sin^2\theta$. This is the charge-equivalent conductance of spin-current loss due to damping, for a macrospin with a precessing cone angle θ.

If without other dynamic excitation, this usually small damping spin current will result in a gradual cone-angle reduction, eventually bring the moment to its lowest-energy state of $\theta = 0$ along the field direction.

A word on the practical unit system used in these expressions. Here, $\epsilon_j = 10^7$ erg/J is an energy unit conversion factor. For the rest of quantities, electron charge $e = 1.602 \times 10^{-19}$ Coulomb, $\hbar = 1.0547 \times 10^{-27}$erg sec. Moment m and field H are in emu and Oe, respectively.

B. Spin current on a macrospin from a rotating field

1. Spin current from a rotating magnetic field

A common method for countering damping-related cone-angle reduction is to drive the moment with a small rotating magnetic field in the $x-y$ plane of Fig. 3, with angular velocity Ω_a. This is the basic idea of ferromagnetic resonance (FMR). This process can be readily described by Eq. (4) as

$$\frac{1}{\gamma}\frac{d\mathbf{n}_m}{dt} - \frac{\alpha}{\gamma}\mathbf{n}_m \times \frac{d\mathbf{n}_m}{dt} = (\mathbf{H}_{\rm eff} + \mathbf{H}_{\rm ac}) \times \mathbf{n}_m \tag{8}$$

and explicitly we write

$$\mathbf{H}_{\rm ac} = H_{\rm ac}[\cos(\Omega_a t)\mathbf{e}_x + \sin(\Omega_a t)\mathbf{e}_y] \tag{9}$$

usually with $H_{\rm ac} \ll H_{\rm eff}$ being the amplitude of the ac magnetic field, rotating in the (x,y) plane. For the case of uniaxial anisotropy along $\mathbf{e}_z$ direction collinear with applied field $\mathbf{H}_a$, one has $\mathbf{H}_{\rm eff} = (H_a + H_k\mathbf{e}_z \cdot \mathbf{n}_m)\mathbf{e}_z$, where H_k is the uniaxial anisotropy field for moment $\mathbf{m}$.

For small cone-angle excitation near the energy minimum, Eq. (8) with $\mathbf{n}_m \approx (\beta_x, \beta_y, 1)$ in linearized form for $\beta_{x,y} \ll 1$ and $H_{\rm ac} \ll H_{\rm eff}$ gives

$$\frac{1}{\gamma}\frac{d}{dt}\begin{bmatrix}\beta_x\\ \beta_y\end{bmatrix} = \frac{H_{\rm eff}}{1+\alpha^2}\begin{bmatrix}-\alpha & -1\\ 1 & -\alpha\end{bmatrix}\begin{bmatrix}\beta_x\\ \beta_y\end{bmatrix} + \frac{H_{\rm ac}}{1+\alpha^2}\begin{bmatrix}\alpha & 1\\ -1 & \alpha\end{bmatrix}\begin{bmatrix}\cos(\Omega_a t)\\ \sin(\Omega_a t)\end{bmatrix} \tag{10}$$

whose steady-state solution is a well-defined problem of eigensystem under driven motion, yielding

$$\begin{bmatrix}\beta_x\\ \beta_y\end{bmatrix} = \frac{H_{\rm ac}}{(H_{\rm eff} - \Omega_a/\gamma)^2 + (\Omega_a/\gamma)^2\alpha^2} \times \begin{bmatrix}H_{\rm eff} - \Omega_a/\gamma & \alpha\Omega_a/\gamma\\ -\alpha\Omega_a/\gamma & H_{\rm eff} - \Omega_a/\gamma\end{bmatrix}\begin{bmatrix}\cos(\Omega_a t)\\ \sin(\Omega_a t)\end{bmatrix} \tag{11}$$

with a power-frequency relationship of

$$\theta^2 = \beta_x^2 + \beta_y^2 = \frac{H_{\rm ac}^2}{(H_{\rm eff} - \Omega_a/\gamma)^2 + (\Omega_a/\gamma)^2\alpha^2}. \tag{12}$$

This amplitude response to the ac-driving field is the essence of FMR. When $\Omega_a \to \gamma H_{\rm eff}$, the maximum amplitude of $\theta^2 = (H_{\rm ac}/\alpha H_{\rm eff})^2$ is obtained as the on-resonance cone-angle power magnitude.

A steady-state-driven precession is present even if off resonance, when the amount of spin current pumped into the macrospin via rotating magnetic field is balanced by the macrospin's damping. That is, *in steady state* for a macrospin driven at angular veolcity Ω_a, the amount of spin current pumped into the system is the same as described by Eq. (6) with its cone angle described by Eq. (12). That is,

$$I_{s,\rm cg} = \left(\frac{2e}{\hbar}\right)(\alpha m \Omega_a/\gamma)\frac{H_{\rm ac}^2}{(H_{\rm eff} - \Omega_a/\gamma)^2 + (\Omega_a/\gamma)^2\alpha^2} \tag{13}$$

for an arbitrary rotation field frequency Ω_a. This follows from a direct time average of the driven-state solution (11). The resonance condition gives the maximum spin current into the damping channel of $I_{s\,\max,\rm cg} = (2e/\hbar)(m H_{\rm ac}^2/\alpha H_{\rm eff})$.

Equation (13) relates a rotating magnetic field's amplitude $H_{\rm ac}$ and its rotation frequency Ω_a to the amount of spin current it couples into a macrospin's damping loss *in steady-state motion*. This is a result we will use later to estimate spin-current transport in the m_1-m_2 model.

In Eq. (13), note the corresponding spin current is dependent on the sign of Ω_a. That is, the rotation direction is important in determining the resonance condition, and hence the net amount of average spin current coupled across.

The ac-field excitation discussed here is assumed to be a full rotational field. A linear oscillatory field can be decomposed into two counter-rotating fields at the same frequency. The net result would be a reduction on the θ^2 expression of Eq. (12) by approximately a factor of 4 when near resonance. Correspondingly, Eq. (13) near resonance would pick up an additional factor of $\sim\frac{1}{4}$. The resulting resonance condition will then be symmetric to the sign of Ω_a.

In our MTJ FL environment, a dipolar coupling could affect dynamics by manifesting itself as a rotational magnetic field. The direction of rotation of such a field would affect its ability to couple spin current into adjacent ferromagnets, such as the reference layer. This is the reason we need to experimentally examine the orientation dependence of the observed spin-current switching threshold, as shown in discussions surrounding Fig. 2. The lack of a systematic FL orientation dependence in our experiment therefore rules out dynamic dipolar coupling as the main mechanism for the observed additional damping behavior as reflected by the STT switching threshold change. These discussions are very approximate because the spin current versus precessional magnetic field derived above is only strictly valid for steady-state dynamics, which does not precisely describe the STT reversal dynamics.

2. Moment precession and exchange-coupling-related spin current

Now consider our two-coupled-macrospin model, i.e., the m_1-m_2 model, where moment m_1 represents the interface-based magnetic moment in an MgO-MTJ system, which is relatively weakly coupled to the rest of the free-layer film designated as m_2. The interface moment m_1 has a small total moment compared to the total of the FL, i.e., $m_1 \ll m_2$, and m_1 is situated in a strong interface uniaxial anisotropy potential whose energy density K_s is on par with the total demagnetization energy of the FL. That is $K_s \sim 2\pi M_{s2}^2 t_2$, where M_{s2} is the magnetization of m_2 in thin-film form, and t_2 its thickness. Because of the vast difference in energy density, the natural time scale for the dynamics of m_1 is much shorter

than that of m_2. For a semiquantitative understanding of our complex system, one could employ an effective adiabatic approximation, assuming that to describe the m_1 dynamics, one could treat m_2 as stationary, whereas when treating m_2 dynamics, one could assume m_1 to be always in its steady-state energy minimum with m_2.

The interface moment m_1 is then driven by an external spin current to precess at its natural frequency. One can use the adiabatic approximation described above to estimate the amount of spin current such a driven motion of m_1 delivers into the exchange coupled m_2. For simplicity, one assumes for this m_1-m_2 system the magnetic axes are all collinear, as is the incoming spin current's polarization. While this assumption may become problematic for describing large cone-angle movements of m_2, it is at least suitable for capturing the initial development of instability when m_2's cone angle is small.

Start with Eq. (13) as the expression for spin current delivered into m_2 via a driven and precessing m_1 at a polar angle of $\theta_1 \ll 1$ (and assume for this step the cone angle of m_2 is small compared to θ_1 and can be treated as on axis). One replaces a precession frequency for m_1 as $\Omega_a \to \Omega_1 = \gamma H_{\text{eff1}} = \gamma(2E_{\text{b1}} + E_{\text{ex}})/m_1$; one treats the exchange coupling between m_1-m_2 as the source of the rotating ac magnetic field that drives m_2 by writing $H_{\text{ac}} \to (E_{\text{ex}}/m_2)\sin\theta_1 \approx H_{\text{ex12}}\theta_1$, with $H_{\text{ex12}} = E_{\text{ex}}/m_2$, and the effective uniaxial anisotropy field on m_2 as $H_{\text{eff2}} \approx 2E_{\text{beff}}/m_2 \ll H_{\text{eff1}}$. Now, one can write the spin current delivered into m_2 via this rotational field as $I_{s12,\text{cg}}$, which in charge unit reads as, according to Eq. (13),

$$I_{s12,\text{cg}} = \left(\frac{2e}{\hbar}\right)(\alpha_2 m_2 H_{\text{eff1}})\frac{H_{\text{ex12}}^2\theta_1^2}{(H_{\text{eff2}} - H_{\text{eff1}})^2 + \alpha_2^2 H_{\text{eff1}}^2}. \quad (14)$$

The intrinsic damping-related spin current within moment m_1 for such a precession motion would be, from Eq. (7),

$$I_{\alpha_1,\text{cg}} = \left(\frac{2e}{\hbar}\right)(\alpha_1 m_1 H_{\text{eff1}})\theta_1^2. \quad (15)$$

Therefore, the total spin-current loss of m_1 precessing at amplitude θ_1 is $I_{\alpha_1,\text{cg}} + I_{s12,\text{cg}}$, and the effect of the precession-induced spin current on m_1 is an additional apparent damping of the form $\alpha_1 \to \alpha_{\text{eff1}}$, with

$$\alpha_{\text{eff1}} = \alpha_1 + \alpha_2\left(\frac{m_2}{m_1}\right)\frac{H_{\text{ex12}}^2}{(H_{\text{eff1}} - H_{\text{eff2}})^2 + \alpha_2^2 H_{\text{eff1}}^2}. \quad (16)$$

C. Spin pumping across an interface in contact with mobile carriers

A spin-pumping-induced spin current across a specified interface can be written as the so-called spin-mixing conductance [13,14]. Here, we only note that phenomenologically for most situations, the spin-mixing conductance associated with dissipative spin current across a conducting interface without significant interface-localized spin-flip excitation has an approximate upper bound around the corresponding charge conductance of the same interface. This spin-mixing conductance, denoted below as g_α, is equivalent to an effective spin-pumping-induced damping coefficient α_{sp}. In charge-equivalent unit, this relationship is of the form

$$g_\alpha = \left(\frac{e^2\epsilon_j}{\hbar}\right)\left(\frac{m}{\mu_B}\right)\alpha_{\text{sp}}, \quad (17)$$

where $e = 1.602 \times 10^{-19}$ C is the electron charge, $\hbar = h/2\pi$ is the reduced Planck constant in erg sec, and $\epsilon_j = 10^7$ erg/Joule. The resulting g_α is in $1/[\Omega]$, and α_{sp} is the spin-pumping-induced additional damping for the macrospin **m**'s dynamics. By the same reasoning, the bulk damping coefficient of **m** can be translated into an equivalent interface conductance using Eq. (17). This allows for the comparison of different paths of spin-current loss to see which one is the most significant.

If the FL interface to an outside spin-shunted environment is limited by an electrical transport conductivity of $1/r_{\text{Amtj}}$ per unit area, such as in the case of FM | tunnel barrier | FM structure, one has, on a per-unit-area basis, $m \to M_s t$, and the g_α-equivalent interface resistance of $r_{\text{A}\alpha} = (\hbar/e^2\epsilon_j)(\mu_B/\alpha M_s t)$, where $1/r_{\text{A}\alpha} = g_\alpha$.

For a typical materials set used for spin-torque-related MTJs, the FM is usually a CoFeB-based alloy, with $M_s \approx 10^3$ emu/cm^3, thickness $t \sim 1$ nm, and a materials $\alpha \approx 0.004$ [19,20]. These give the damping equivalent $r_{\text{A}\alpha} \sim 9\ \text{m}\Omega\,\mu\text{m}^2$. This is much lower than an MTJ's tunnel current related $r_{\text{Amtj}} \gtrsim 2\ \Omega\mu\text{m}^2$. Thus, within a single macrospin picture, one would not expect a significant amount of spin-current loss through the tunnel-current-related spin pumping because its RA is too high compared to the intrinsic damping by a factor of about 100× or more.

Experimentally, however, one does observe a systematic dependence of the spin-torque switching threshold current on the tunnel junction's RA value, as shown in Sec. II. This observation appears to be systematic and robust, but could not be readily explained within the simple macrospin spin-torque excitation model, suggesting the presence of other significant processes beyond macrospin.

Below, we discuss a "serial-spin-current divider model" based on the separation of the interface moment's high-energy dynamics from the rest of the FL. This approach within the context of the extremely simplified two-macrospin m_1-m_2 model can account for the order of magnitude of the spin-pumping loss of spin current that occurs in practical MTJs we examined in Sec. II.

V. SERIAL-SPIN-CURRENT DIVIDER MODEL

A. Model structure and material parameter assumptions

Consider a model MTJ free-layer disk of diameter a and thickness t, with $t/a \ll 1$. Assume for now that the disk is in macrospin state, with one side of the disk contacting a tunnel barrier. One can use the methodology developed in Sec. IV to keep track of the spin current induced by the precessional moment of this magnetic disk. For simplicity, we drop the factor of $\pi/4$ in area calculation and simply write the junction interface total area as a^2.

We will use notations of our two-macrospin m_1-m_2 model construct, in the form of $m_1 = M_{s1}t_1a^2$ as the interface layer

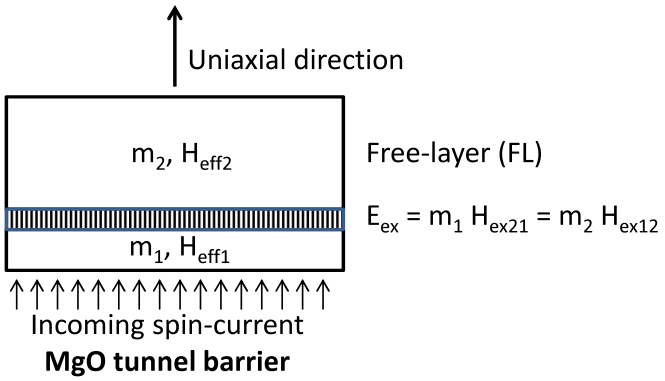

FIG. 4. An illustration of the m_1-m_2 model for the FL of an MTJ. The topology of the sketch serves only to distinguish interface anchored vs the rest of the FL moment, and does not necessarily imply the moments are smoothly lined up in cleanly layered fashion, although a layered construct could be one possible configuration of such an m_1-m_2 system. See Sec. VII for more discussions.

of moment, and $m_2 = M_{s2}t_2a^2$ the interior moment of FL, as illustrated in Fig. 4.

The interface moment m_1 is situated in a high-anisotropy energy well, with energy density per unit area K_s to provide sufficient perpendicular anisotropy for the entire FL to stay perpendicular. m_1 also has much less net total moment, i.e., $M_{s1}t_1a_1^2 \ll M_{s2}t_2a_2^2$. Here, the device FL is modeled as a simple cylindrical disk, $a_1 = a_2 = a$, and the thicknesses add up to the total FL value, $t_1 + t_2 = t$. In addition, we use an interface exchange energy E_{ex} to describe the magnetic coupling between m_1 and m_2. The relevant exchange fields on each moment are $H_{ex21} = E_{ex}/m_1$ and $H_{ex12} = E_{ex}/m_2$. We assume $K_s \sim 2\pi M_{s2}^2 t_2$, $t_1 \ll t_2$ in our specific example.

It follows then m_1 dynamics would occur at a much higher-frequency range than m_2. This allows for an adiabatic assumption for m_1 dynamics where m_2 is assumed stationary, and an approximate description for the m_2 dynamics with the same adiabatic assumption that m_1 is always in its fully relaxed steady-state energy minimum.

B. A serial-spin-current propagation scenario

In this scenario, we assume that, to the leading order, the spin-transfer torque (STT) from the tunnel carriers is completely absorbed within m_1. In steady state, the m_1 dynamics under STT influence couples some spin current out of m_1 and into other channels that we describe below.

Write the total STT spin current arriving at m_1 as I_s. Define $V_{B1} = 2\mu_B H_{eff1}/e$, with $H_{eff1} = H_{k1} + H_{ex21} = (2\xi_{b1} + \xi_{ex})k_BT/m_1$, where the dimensionless parameters ξ_{b1} and ξ_{ex} describe the anisotropy and exchange energy of m_1 in k_BT unit with k_BT being the ambient thermal energy (of ~26 meV at room temperature). For a total cone-angle amplitude of θ_1 for m_1 dynamics in steady state under I_s, the balance of spin current reads as

$$\begin{aligned} I_s &= I_{s1} + I_{s2} = V_{B1}\theta_1^2(g_{TB} + g_{\alpha 1} + g_{BB}),\\ I_{s1} &= V_{B1}\theta_1^2(g_{TB} + g_{\alpha 1}),\\ I_{s2} &= V_{B1}\theta_1^2 g_{BB}, \end{aligned} \tag{18}$$

where I_s is the total spin current leaving m_1 if it precesses at θ_1. At constant θ_1, this must equal to the total incoming STT spin current from tunnel current. I_{s1} is the spin current leaving m_1 by bulk damping and via tunnel-barrier interface spin pumping. I_{s2} is the spin current leaving m_1 and enter m_2, mostly through exchange-coupled precession dynamics. g_{TB} is the spin conductance for the tunnel-barrier facing interface of m_1, with r_{Amtj} as its RA product. $g_{\alpha 1}$ is the spin conductance for m_1's intrinsic damping α_1. g_{BB} is the spin conductance due to exchange coupling between m_1 and m_2. We assume the mobile electron's contribution to spin-current transport between m_1 and m_2 is small compared to g_{BB}.

From Eqs. (16) and (17) one may write, with $C_1 = (\frac{e^2\epsilon_j}{\hbar})(\frac{m_1}{\mu_B})$,

$$\begin{aligned} g_{TB} &= \frac{a^2}{r_{Amtj}} \triangleq C_1\alpha_{TB},\\ g_{\alpha 1} &= C_1\alpha_1,\\ g_{BB} \triangleq \frac{a^2}{r_{Amm}} &\triangleq C_1\alpha_{BB} = a^2\left(\frac{e^2\epsilon_j}{\hbar}\right)\left(\frac{\alpha_2 M_{s2}t_2}{\mu_B}\right)\\ &\quad\times \frac{H_{ex12}^2}{(H_{eff1} - H_{eff2})^2 + \alpha_2^2 H_{eff1}^2}. \end{aligned} \tag{19}$$

Here, we defined three additional parameters: α_{TB} and α_{BB} are the effective damping coefficients for m_1 of its MTJ barrier spin pumping and its exchange-induced spin-current loss from m_1 into m_2, respectively; and r_{Amm} which, related to g_{BB}, is a fictitious RA product describing the strength of interface spin-current loss from m_1 to m_2. Equations (18) and (19) solve to give the spin current divide as

$$\begin{aligned} I_{s1} &= \frac{\alpha_1 + \alpha_{TB}}{\alpha_1 + \alpha_{TB} + \alpha_{BB}} I_s,\\ I_{s2} &= \frac{\alpha_{BB}}{\alpha_1 + \alpha_{TB} + \alpha_{BB}} I_s. \end{aligned} \tag{20}$$

C. STT threshold current estimate for I_s, and RA dependence of switching efficiency

Under the adiabatic assumption discussed above, one may write the thermal-activation energy defined threshold instability current for m_1 for total spin-current input I_s as I_{cfs}:

$$I_{cfs} = (2\xi_{b1} + \xi_{ex})(\alpha_1 + \alpha_{TB} + \alpha_{BB}). \tag{21}$$

Here, for simplicity of the expressions, thresholds such as I_{cfs} are expressed using a reduced unit of $I_0 = (2e/\hbar)k_BT$ with k_BT being the ambient thermal energy.

The overall magnetic switching threshold for the combined m_1-m_2 system, as dictated by the slow dynamics of m_2, then, is approximately

$$I_{sc2} = \alpha_2\xi_{aeff} = \alpha_2\left(\xi_{b1}\Gamma_T\left[\frac{\xi_{ex}}{N_f}\right] + \xi_{b2}\right), \tag{22}$$

where ξ_{aeff} is the effective thermal-activation energy barrier height for the slow dynamics of the m_1-m_2 system, $\Gamma_T[x] =$

$3\coth^3 x - \frac{3}{x^2} - 2$ is a thermal-smearing related function[2] that reduces the amount of PMA m_1 couples into m_2 due to m_1's thermal fluctuation. N_f is the number of regions within m_1 that fluctuates independently at finite temperature k_BT.

The effective thermal activation barrier height $\xi_{\rm aeff}$'s fluctuation-dependent reduction (or equivalently the apparent increase of fluctuation temperature) due to STT current I_s can thus be approximately written as

$$\xi_{\rm aeff}[I_s] \sim \left\{\xi_{\rm b1}\Gamma_{\rm T}\left[\frac{\xi_{\rm ex}}{N_f}\left(1-\frac{I_{s1}}{I_{\rm cfs}}\right)\right]+\xi_{\rm b2}\right\}\left(1-\frac{I_{s2}}{I_{\rm sc2}}\right). \tag{23}$$

A low I_s expansion of $\xi_{\rm aeff}$ in Eq. (23) with $I_{s1,s2}$ substituted by I_s from Eq. (20) gives a leading-order result of the effective thermal activation barrier height at zero spin-current bias. Assuming $\xi_{\rm ex}/N_f \gg 1$, i.e., exchange is sufficient to keep m_1 with a sufficiently small thermal fluctuation cone angle compared to the dynamics of θ_1 for each individual region, one has

$$\begin{aligned}\xi_{\rm beff} \triangleq \xi_{\rm aeff}[I_s\to 0] &\approx \xi_{\rm b1}\left[1-3\left(\frac{N_f}{\xi_{\rm ex}}\right)^2\right]+\xi_{\rm b2},\\ \frac{I_{\rm sc}}{\xi_{\rm beff}} &\approx \alpha_2\left[1+\left(\frac{e^2\epsilon_j}{\hbar\mu_B}\right)(M_{s1}t_1\alpha_1)r_{\rm Amm}\right.\\ &\quad\left.+\left(\frac{r_{\rm Amm}}{r_{\rm Amtj}}\right)\right],\end{aligned} \tag{24}$$

where $\xi_{\rm beff}$ is the effective zero applied spin-current thermal activation barrier height for the combined m_1-m_2 system, and $I_{\rm sc}$ is the instability threshold current for I_s according to the zero barrier height intercept from linearized expansion of Eq. (23), in reduced unit of I_0 [as defined when discussing Eq. (21)]. In the limit of $\xi_{\rm ex}\to+\infty$, the first line of Eq. (24) recovers the simple macrospinlike form of $\xi_{\rm beff}\to\xi_{\rm b1}+\xi_{\rm b2}$, as it should.

Restoring real-life unit I_0 for the inverse efficiency, and including an MTJ-related spin-polarization efficiency factor $\eta = \frac{\sqrt{m_r(m_r+2)}}{2(m_r+1)}$, one writes

$$\begin{aligned}\frac{I_{\rm sc}}{\xi_{\rm beff}} \triangleq \frac{1}{\kappa} &\approx \left(\frac{2e}{\hbar}\right)k_BT\left(\frac{\alpha_2}{\eta}\right)\left[1+\left(\frac{e^2\epsilon_j}{\hbar\mu_B}\right)\right.\\ &\quad\left.\times(M_{s1}t_1\alpha_1)r_{\rm Amm}+\left(\frac{r_{\rm Amm}}{r_{\rm Amtj}}\right)\right]\end{aligned} \tag{25}$$

which gives rise to an MTJ RA-dependent inverse-efficiency slope and intercept (A_0 and A_1 in our experimental observations, Figs. 1 and 2) of

$$\begin{aligned}\text{slope} &= \frac{\delta(1/\kappa)}{\delta\left(1/r_{\rm Amtj}\right)} \approx \left(\frac{2e}{\hbar}\right)k_BT\left(\frac{\alpha_2}{\eta}\right)r_{\rm Amm},\\ \text{intercept} &= \left.\frac{I_{\rm sc}}{\xi_{\rm beff}}\right|_{r_{\rm Amtj}\to+\infty} = \left(\frac{2e}{\hbar}\right)k_BT\left(\frac{\alpha_2}{\eta}\right)\\ &\quad\times\left[1+\left(\frac{e^2\epsilon_j}{\hbar\mu_B}\right)(M_{s1}t_1\alpha_1)r_{\rm Amm}\right]\end{aligned} \tag{26}$$

with

$$r_{\rm Amm} = \left(\frac{\hbar}{e^2\epsilon_j}\right)\left(\frac{\mu_B}{\alpha_2 M_{s2}t_2}\right)\frac{(H_{\rm eff1}-H_{\rm eff2})^2+\alpha_2^2H_{\rm eff1}^2}{H_{\rm ex12}^2}, \tag{27}$$

where the definition of $r_{\rm Amm}$ follows from the third line in Eq. (19).

Equation (26) gives a model-derived slope and intercept of the inverse switching efficiency versus MTJ RA, based on a two-body, serial-spin-current transport process. This relationship can be compared with experiment.

Within the confines of the m_1-m_2 model, this predicts a size-independent slope and intercept of the switching efficiency versus junction RA (i.e., $r_{\rm Amtj}$). Experimentally, however, one observes a significant size dependence in these quantities, and one is well aware that in our size range of 15 nm and up, the MTJ does not behave in a simple macrospin fashion, but involves dynamic inhomogeneities that reveal additional internal degrees of freedoms, even after one separates the interface dynamics from that of the FL interior. The phenomenological construct of this serial-spin conduction model would need to be modified to examine the effect of finite-size magnetic excitation effects inside m_1 and m_2. Before delving into these details, however, we first examine some quantitative values to see if the two-macrospin-based m_1-m_2 model (26) and experimental results are within similar orders of magnitudes.

D. Quantitative evaluation of the two-macrospin m_1–m_2 serial conduction model: Comparison with experiments

A set of possible material parameters for use in this model system is listed in Table I. These parameters are not sufficiently constrained by available observations. To proceed with estimates, one uses a reasonable range of values to check and ensure the consistency of the model mechanism with observed quantities, and of their trends of dependence on these listed variables.

In Eqs. (26) and (27), we use a set of material parameters listed in Table I, that is, assuming $\alpha_2 = 0.01$, $M_{s2} = 800$ emu/cc, $t_2 \approx 2$ nm, $H_{\rm eff1}\sim 70$ kOe, $H_{\rm eff2}\sim 3$ kOe, $H_{\rm ex12}\sim 1.0$ kOe, one has a model predicted macrospin value for slope A_1 in Eqs. (26) and (1) as slope $\to A_{1,\rm macrospin}\approx 3.1\ \mu{\rm V}\,\mu{\rm m}^2/k_BT$. Equivalently, one has a model predicted $r_{\rm Amm,macropsin}\approx 10.7\ \Omega\mu{\rm m}^2$ compared with the experimentally derived value of $r_{\rm Amm,measured}\approx 13.8\ \Omega\mu{\rm m}^2$, for devices around 35-nm diameter.

[2]Equation (22) is the finite-temperature, finite $E_{\rm ex}$, and small m_1 total energy equivalent of the coefficient in front of the $\cos^2\theta$ term in Eq. (28) of Ref. [16]. It is the result of an estimate based on a so-called "Finger Model"(unpublished). The essence of the model is that, for our practical size range of tens of nm of diameters, the material for m_1 is not single domain, but rather breaks into N_f regions of independent fluctuators at finite temperature and under STT excitation. Such fluctuations reduce the net exchange coupling between the m_1 and m_2 regions, reducing the total PMA energy. Equation (22) is a leading-order estimate for such a residual anisotropy energy after including finite-temperature-induced reduction effect.

TABLE I. A summary view of symbols and range of values of material parameters in Sec. V D's model calculations for comparison with experimental findings. The parameters are not sufficiently constrained by observations.

Variable	Unit	Typical values	Description
M_{s1}	emu/cm^3	100–200	m_1 saturation magnetization estimate from [16] and error-rate tests
t_1	nm	∼0.2	m_1 layer thickness
H_{k1}	kOe	10–80	m_1 anisotropy field, related to ξ_{b1}
H_{ex21}	kOe	10–100	m_1 exchange field at m_1-m_2 interface, related to ξ_{ex}
H_{eff1}	kOe	10–100	m_1's total effective field, $\sim 2H_{k1} + H_{ex21}$
M_{s2}	emu/cm^3	∼800	m_2 layer's saturation magnetization
t_2	nm	∼2	Essentially it is the FL thickness, model assumes $t_2 \gg t_1$
H_{k2}	kOe	−5–10	m_2's anisotropy field, including demag field
H_{eff2}	kOe	0–10	H_{k2} plus any applied field
H_{ex12}	kOe	0.1–10	Exchange field from the m_1-m_2 interface
ξ_{b1}	k_BT	10–100	m_1's total PMA energy (for a given diameter a, same below)
ξ_{ex}	k_BT	10–100	m_1-m_2 interface's total exchange energy
ξ_{b2}	k_BT	10–100	m_2's total anisotropy energy
$\alpha_{1,2,R}$	n.u.	0.0001–0.03	m_1, m_2, and RL's LLG damping
D_a	eV Å^2	0.05–0.5	In-plane direction exchange stiffness of the m_2 layer

The model expectation value therefore is within reason compared with the experimentally measured values, although precise comparison would not be attempted here since one does not know with certainty the parameters that determine the model. An obvious discrepancy of the two-macrospin model above is it does not predict any diameter dependence of the efficiency's dependence on r_{Amtj}. This needs to be addressed by examining other pathways of spin current, which we will examine more closely in the next section.

E. Other possible channels of spin conduction

Note in Eq. (18) one essentially writes two parallel spin conduction outlet channels for m_1, that of $g_{TB} = 1/r_{Amtj}$ for everything we call spin-pumping-like, and $g_{BB} = 1/r_{Amm}$, in which we lump together all spin conduction processes from m_1 into m_2.

One could in principle split g_{TB} and g_{BB} further into other possible components. First, let us examine g_{TB}, i.e., the spin current leakage on the MgO interface side. In addition to the tunnel barrier's r_{Amtj}, one could also include the "fieldlike" interaction across a magnetic tunnel junction's barrier. This term has been demonstrated to both have an antiferromagnetic exchange direction [21], and the coupling strength having a V^2 dependence on MTJ bias voltage [22], from band-structure considerations [7,23]. Early experiments placed an estimate of an antiparallel coupling of the order of 0.1 erg/cm^2 in a Fe|MgO|Fe sandwich system [24]. Later, Tiusan *et al.* arrived at a zero-bias interface exchange energy of the order 4×10^{-2} erg/cm^2 from FMR studies of Fe|MgO|Fe|V multilayers for an MgO barrier of about 1 nm thick, with the AF coupling strength increasing strongly upon reduction of MgO thickness [25].

In our samples, the strength of this "fieldlike" exchange interaction term, here below denoted as H_{exmtj} (on the FL side), is not known. Based on the work cited above, one could place a rough estimate of the low-bias exchange field on FL, in the neighborhood of $H_{exmtj} \sim -\dfrac{0.04 \text{ erg/cm}^2}{800 \text{ emu/cc} \times 2.0 \text{ nm}} \sim -250$ Oe with the negative sign denoting antiparallel coupling.

Turn now to the expression of rotating exchange-field-induced spin current [Eq. (19)]. One could use a similar relationship to define an exchange-field-like contribution to the spin conductance across the tunnel barrier.

First, consider the possible rotation-field coupling between the FL's interface moment and RL during the two-moment serial-spin conduction model. This can be written as

$$r_{Aexmtj} = \left(\frac{\hbar}{e^2\epsilon_j}\right)\left(\frac{\mu_B}{\alpha_R M_{sR} t_R}\right)\frac{(H_{eff1} - H_{effR})^2 + \alpha_R^2 H_{eff1}^2}{H_{exmtjR}^2}, \tag{28}$$

where M_{sR}, t_R, H_{effR}, α_R, and H_{exmtjR} are the quantities for the reference layer involved in driven dynamics following the FL interface's m_1 dynamics.

For parameter estimates, a rational starting point is to assume the reference layer's response is dominated by an interface CoFeB dusting layer not tremendously different from that of the FL's m_1. Assuming $\alpha_R \sim 0.03$ (so as to include possible spin-current loss into the rest of the RL), $M_{sR} \sim 800$ emu/cc, $t_R \sim 1$ nm, $H_{effR} \sim 6$ kOe, and $H_{exmtjR} \sim 250$ Oe, one has from Eq. (28) a $r_{Aexmtj} \sim 93\ \Omega\mu\text{m}^2$, which is larger than the MTJ tunnel barrier's RA product by about 10×. This, however, is the zero-bias estimate, and the strength of H_{exmtjR} is expected to increase with V^2, although the voltage dependence coefficient for our PMA MTJ is not known, either. An earlier, in-plane version of junctions with similar $r_{Amtj} \sim 10\ \Omega\mu\text{m}^2$ gave the voltage-dependent coefficient of about 70 Oe at 1 V [22].

Noting the nature of interactions as such, the matching frequency of the two bodies plays a significant role in enhancing spin conduction. That is, an on-resonance spin conduction process is far more effective than off resonance. There could be potentially multiple scenarios of on-resonance interactions, either in RL responding to the "slow" dynamics of the FL, where in Eq. (1) one substitutes H_{eff1} with H_{eff2}, or in RL's interface moment having a similar m_1-like structure as that of the FL, in which case one might need to replace H_{effR} with something closer to H_{eff1}. The combination of possibilities is quite numerous. An extreme case is to consider

the on-resonance minimum value of r_{Aexmtj} by removing the $(H_{\mathrm{eff1}} - H_{\mathrm{effR}})^2$ term. That gives

$$r_{\mathrm{Aexmtj,min}} \sim \left(\frac{\hbar}{e^2\epsilon_j}\right)\left(\frac{\mu_B}{M_{\mathrm{sR}}t_R}\right)\frac{\alpha_R H_{\mathrm{eff1}}^2}{H_{\mathrm{exmtjR}}^2}, \tag{29}$$

which for an interface moment resonance with $M_{\mathrm{sR}} \sim 200$ emu/cc, $t_R \sim 0.2$ nm, and $\alpha_R \sim 0.03$, $H_{\mathrm{eff1}} \sim 70$ kOe, one has $r_{\mathrm{Aexmtj,min}} \sim 2.2\ \Omega\mu\mathrm{m}^2$, on par with an MTJ's tunnel barrier resistance. However, if the resonance is among the slow dynamics mode, $r_{\mathrm{Aexmtj,min}} < \mathrm{m}\,\Omega\mu\mathrm{m}^2$ is easily possible with reasonable parameters for H_{eff2} and the moment-thickness product of the RL: for example, if $M_{\mathrm{sR}} \sim 50$ emu/cc (low moment value due to SAF reference layer balance), and $t_R \sim 3.0$ nm, and with slow dynamics $H_{\mathrm{eff2}} \sim 3$ kOe replacing H_{eff1}, one could easily get $r_{\mathrm{Aexmtj,min}} \sim 1.1\ \mathrm{m}\Omega\ \mu\mathrm{m}^2$. In this scenario then, the resonant coupling between RL and FL alone would be sufficient to provide a spin-current leakage on par with intrinsic damping's equivalent interface $r_{\mathrm{A}\alpha} \sim 2\ \mathrm{m}\,\Omega\mu\mathrm{m}^2$ according to Eq. (17) (with $r_{\mathrm{A}\alpha} = a^2/g_\alpha$ as its definition here), and the parameters given above for our FL.

The mechanisms discussed in this section are meant to point out the possible existence of *many* plausible mechanisms that could provide a spin-pumping-like process for the FL to lose its spin current. It does not at present uniquely identify one single mechanism as the leading contributor. That is likely a very device geometry and materials-specific problem that remains too complex and is thus quite elusive at present.

Note also that discussions here are confined to steady-state, subthreshold limit dynamics. When fast-driven dynamics is involved, the transient spin-current flow will be more complex, and such simple linearized analysis is generally not valid. They are therefore only an approximate guide to the orders of magnitudes of the various interaction contributions in the problem, and would not yield quantitatively reliable results.

One piece of evidence, however, argues strongly *against* a significant involvement of such an exchange-field-like coupling between FL rotation and RL. As discussed in Sec. IV B, the precessional-field-induced spin-current loss depends on the direction of the precession. In other words, it should behave *very* differently for AP-P (W0) vs P-AP (W1) transition, equivalent to having the precession frequency at $\pm H_{\mathrm{eff1}}$ in Eq. (1). This magnitude of asymmetry is simply not seen in our data, as discussed in Fig. 2. Thus, there is no evidence for direct involvement of exchange-field-like coupling as the main mechanism in this particular set of junctions we experimented with.

Therefore, the serial-spin-current conduction picture described in Sec. V B and quantified in Sec. V D remains the most likely scenario for the observed RA-dependent switching efficiency. Such two-macrospin model, however, is clearly insufficient, as it would not produce a *size-dependent change* of switching efficiency versus junction RA. To attempt to explain the size dependence, one needs to go beyond the simple two-macrospin construct, and include some form of size-dependent internal degrees of freedom for the FL.

In the next section, we give a semiquantitative consideration of one such mechanism, namely, the coupling of spin current into finite-wavelength standing waves of the FL.

VI. STANDING-WAVE MODES AS ADDITIONAL OSCILLATORS COUPLING THE M_1-M_2 SYSTEM

In this section, one examines in more detail the m_1 to m_2 spin-current transport process within the context of the serial spin-current divider model as described in Sec. V. We now take a broader view of the precessional exchange-field-induced spin current within the FL's m_1-m_2 system described by Eqs. (25)–(27). This process can be viewed, in a linear system (small-amplitude limit), as possible coupling terms into a superposition of more than one available mode of excitation in m_2. The assumptions of a serial-spin-current divider model in terms of its material parameters (Sec. V A) and model scenarios (Sec. V B) are preserved.

What is different is in the expression of spin conductance g_{BB} in Sec. V B and Eq. (19), instead of a single-mode macrospin resonance for m_2, one writes the total coupling to *multiple* available standing-wave resonance modes of m_2:

$$g_{\mathrm{BB}} = \frac{a^2}{r_{\mathrm{Amm}}} = a^2\left(\frac{e^2\epsilon_j}{\hbar}\right)\left(\frac{\alpha_2 M_{\mathrm{s2}} t_2}{\mu_B}\right) \times \sum_{\mathrm{n}} \frac{\Phi_n H_{\mathrm{ex12}}^2}{\left(H_{\mathrm{eff1}} - H_{\mathrm{eff2,n}}\right)^2 + \alpha_2^2 H_{\mathrm{eff1}}^2}, \tag{30}$$

where $H_{\mathrm{eff2,n}}$ represents resonant frequency conditions of active standing-wave modes of the FL's m_2, and Φ_n a numerical factor representing the junction-area averaged coupling strength for a given mode "n." The detail of such summation will depend on the size of interface moment m_1's coherent region in comparison to that of the mode wavelength, which is too complex to attempt. The simple physics behind Eq. (30) is that one considers *any and all* spin-current m_1 precession delivers into m_2 as valid spin current to affect the slow dynamics of the m_2 system. An implicit assumption also is that the mode-mode coupling in m_2 is sufficiently strong that a steady-state precession amplitude could be established on a time scale fast compared to m_2's slow dynamics.

In practice, only modes with nearly location-independent phase within the coherent length of m_1 dynamics would couple effectively with an m_1-precession-related g_{BB}. That argues for relatively low-lying, high-symmetry, long-wavelength modes. Therefore, we attempt here to semiquantitatively consider the lowest-lying standing-wave mode that is close to H_{eff1}.

For simplicity, we consider only the two lowest-lying modes: that of the macrospin mode (which we already did in Sec. V B), and that of the lowest center-peaked, edge-pinned standing-wave mode, which we now address.

The lowest-order center-peaked standing-wave mode could have an excitation frequency next to (or near) macrospin mode in our particular geometry. Thus, it is also the mode that likely has a low-lying instability spin-current threshold. To see this, one needs to examine the nature of magnetic resonance and standing-wave modes in the presence of nonuniform edge-demagnetization fields.

A. Location-dependent anisotropy field

The leading-order location-dependent perpendicular direction demagnetization factor, which is a cause of a location-dependent anisotropy field in our perpendicularly magnetized

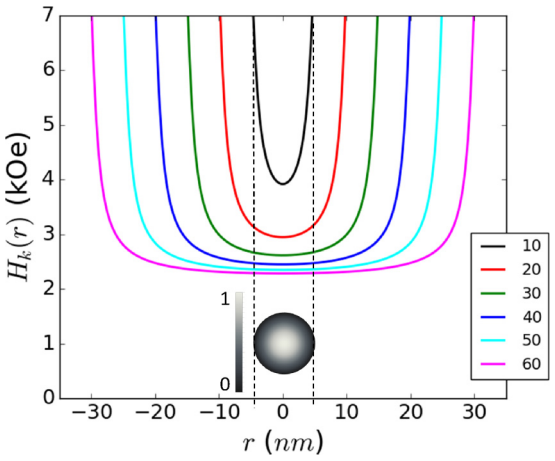

FIG. 5. An illustration of the radial location dependence of total effective anisotropy field for a 2-nm-thick disk with various diameters from $a = 10$ to 60 nm, 10 nm apart. One assumes for Eq. (32) its $M_{s2} \to M_s = 800$ emu/cm^3, and $H_{ku} = 12$ kOe. The curves are numerically generated using Eq. (32), and with the z position in the middle of the disk, $z \to t/2$. Inset: top view of the $a = 10$ nm diameter disk for the first center-peaked standing-spin-wave mode's normalized in-plane moment amplitude from Eq. (33), with $k = (s/a)$ and $s = 4.81$.

FL disk, can be expressed as [26]

$$N_{zz}^{(1)}(r,z,\zeta) = \frac{1}{4}\int_{\xi=0}^{+\infty} J_0(\xi r)J_1(\xi/2)\{\exp(-\xi z) + \exp[-\xi(\zeta - z)]\}d\xi, \quad (31)$$

where $\zeta = t/a$ is the disk's aspect ratio, t being its thickness, and a its diameter. Here, r and z are the radial and z coordinates normalized to the disk's diameter. Thus, for a point inside the FL disk, $r \in [0,1/2]$, and $z \in [0,t/a]$. The location-dependent total effective field is, therefore, expressed using real length-unit coordinates (r,z):

$$H_{eff2}(r,z) = H_{ku} - 4\pi M_s N_{zz}^{(1)}\left(\frac{r}{a},\frac{z}{a},\frac{t}{a}\right) \quad (32)$$

with H_{ku} being its effective volume anisotropy (for our FL's m_1-m_2 slow dynamics, specifically) which can also include a uniform z-direction magnetic field in this small-amplitude limit. An illustration for this radial-position-dependent effective magnetic field of a FL disk of $M_{s2} = 800$ emu/cm^3, $t = 2.0$ nm, and $H_{ku} = 12$ kOe is shown in Fig. 5 for diameters from 10 to 60 nm.

B. Standing waves in a nonuniform anisotropy field

One method of approximating a solution for a standing spin wave in a nonuniform magnetic field of similar spatial symmetry is that, to the leading order, a standing spin wave's dispersion function contains an effective anisotropy field that is weighed by the location-dependent spin-wave amplitude [27,28].

As a starting point, consider the edge-pinned standing-wave mode for a perpendicularly magnetized, circular-shaped FL disk *with uniform anisotropy field*. In small amplitude, thin film ($t \ll a$), and in linearized amplitude limit, it has the solution of [29]

$$M_x = AJ_{\nu-1}(k\rho)\cos[(\nu-1)\varphi - \Omega\tau],$$
$$M_y = -AJ_{\nu-1}(k\rho)\sin[(\nu-1)\varphi - \Omega\tau] \quad (33)$$

for a uniform (location independent in disk interior) anisotropy potential. Here, M_x, M_y denote the in-plane, location-(ρ,φ)-dependent amplitude excitation of the spin wave, with $M_x, M_y \ll M_{s2}$. They are the position-dependent equivalent of the $\beta_{x,y}$ amplitude in Eq. (10). k is the standing-wave wave vector chosen to satisfy the boundary condition, and J_ν is the Bessel-J function of νth order.

For our lowest-lying, nonuniform standing-wave mode, considering the edge-peaked nature of the anisotropy field shown in Fig. 5, it would correspond to (approximately) the standing-wave mode with a center peak, and with a first zero crossing near disk edge $\rho \approx a/2$ in Eq. (33). This sets $\nu = 1$ and $k_1 = s/a$ with $s = 4.809\,65$ representing the first node at disk edge.

The next step is to use this approximate standing-wave solution's amplitude to weigh the anisotropy field's radial dependence, and use that to estimate the leading-order correction to our spin-wave resonance frequency due to location-dependent H_{eff2} correction as defined by Eq. (32). Taking this approach, the corresponding amplitude-weighed, mode-specific demagnetization factor entering the mode-frequency-field relationship is then

$$N_{zzsw}^{(1)} \approx \frac{1}{A_b}\int_{\rho=0}^{1/2} N_{zz}^{(1)}(\rho,\zeta/2,\zeta)J_0^2(s\rho)\rho\, d\rho, \quad (34)$$

where $A_b = \int_{\rho=0}^{1/2} J_0^2(s\rho)\rho\, d\rho$ is a normalization factor. The total mode-frequency-related field H_{ksw} is now expressed as

$$H_{ksw} = \frac{\hbar\omega_{sw}}{2\mu_B} = H_{ku} - 4\pi M_{s2}N_{zzsw}^{(1)} + \frac{D_a}{2\mu_B}\left(\frac{s}{a}\right)^2 \quad (35)$$

with H_{ku} being its effective volume anisotropy, as defined in Eq. (32), D_a the exchange-stiffness constant, and $s \approx 4.81$ as defined in Eq. (34) for the first center-peaked standing-wave mode with complete edge pinning.

For comparison, the uniform-rotation (macrospinlike) mode experiences an effective field that is H_{kav} that is the simple volume average of $N_{zz}^{(1)}$ of Eq. (31), which gives

$$H_{kav} \approx K_{ku} - 4\pi M_{s2}N_z(t/a) \quad (36)$$

with the volume-averaged demagnetization factor to the leading order estimated as [30]

$$N_z(\zeta) \approx 1 - \left(\frac{\zeta}{\pi}\right)\left[2\ln\left(\frac{4}{\zeta}\right) - 1\right] + O[\zeta]^2. \quad (37)$$

The diameter dependence of the center-peaked mode (35) for a set of exchange-stiffness value D_a is compared with the volume-averaged anisotropy field in Fig. 6. Where the solid lines cross under the dashed line, the standing-wave mode becomes the lowest-frequency mode, and correspondingly with the lowest spin-torque instability threshold current.

Therefore, with these dispersion characteristics, a finite exchange-stiffness system would have a critical size, beyond which the first-order standing-wave mode's STT threshold (which is proportional to its eigenfrequency) would become

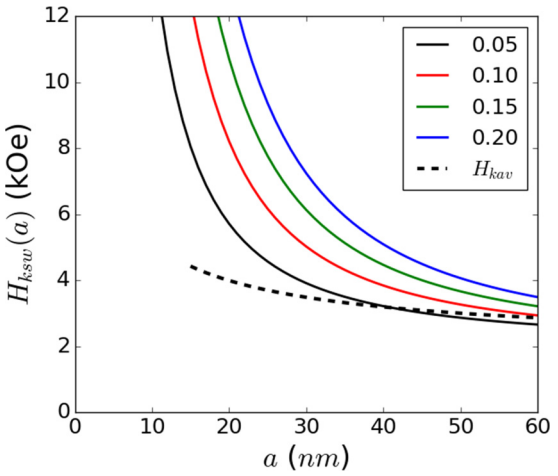

FIG. 6. The diameter dependence of the first center-peaked standing-wave mode's effective energy $H_{\rm ksw}$ [Eq. (35)] for four exchange-stiffness values of $D_a = 0.05,0.10,0.15,0.20$ eV Å^2, compared with the volume-averaged total anisotropy energy or Eq. (36) (dashed line). The same set of material parameters as in Fig. 5 was used, namely, $M_{s2} = 800\,{\rm emu/cm^3}$, $t = 2$ nm thick, and $H_{\rm ku} = 12$ kOe.

lower than the volume-averaged, uniform-rotation mode, and thus becomes the mode first to exceed threshold and grow.

This mechanism alone could imply a dimensional crossover. For devices of a diameter larger than a length of the order of the PMA-determined exchange length, which is related to the point in Fig. 6 where the dashed line crosses the solid lines, the threshold switching current would scale differently than the uniform rotation mode (i.e., macrospin), and start to scale with the first standing-wave mode, thus adding a constant-current scaling term from the exchange-stiffness part of Eq. (35). The crossover diameter due to this mechanism as illustrated in Fig. 6, however, for our representative parameters used here, tend to be relatively large, with an estimate of above 50 nm. The experimentally interesting region where our data concentrate on is already well below this size, down to 20 nm or less. Yet, our MTJ system remains nonmacrospinlike, indicating the presence of other mechanisms.

Indeed, the standing-wave mode can play a role for junction diameters well below the exchange length. Since the spin-current coupling between m_1 and m_2 is from exchange-field precession, the high-frequency m_1 dynamics would couple more effectively into the standing-wave mode which has a frequency above the uniform rotation mode, *especially in small device diameter limit*.

To see this, let us estimate the spin conductance in steady-state dynamics for the m_1-m_2 transport according to Eq. (30). For simplicity, we only evaluate this lowest standing-wave resonance mode to get an idea of the magnitude of $r_{\rm Amm}$ and its diameter a dependence. Substituting $H_{\rm ksw}$ of Eq. (35) into Eq. (30), one has

$$r_{\rm Amm} \approx \left(\frac{\hbar}{e^2\epsilon_j}\right)\left(\frac{\mu_B}{\alpha_2 M_{s2} t_2}\right)\frac{(H_{\rm eff1}-H_{\rm ksw})^2+\alpha_2^2 H_{\rm eff1}^2}{H_{\rm ex12}^2}. \tag{38}$$

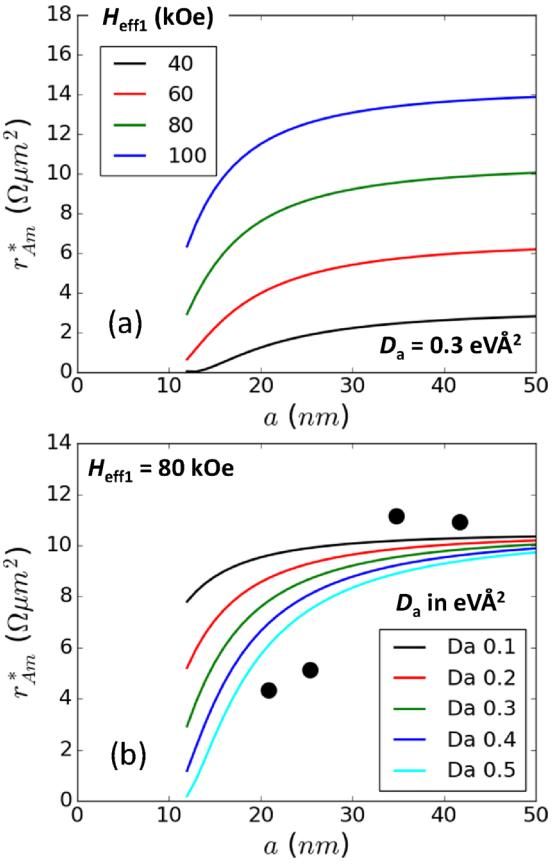

FIG. 7. The diameter dependence of $r^*_{\rm Am}$ according to Eq. (39), numerically evaluated (a) for four different values of $H_{\rm eff1}$, with $D_a = 0.3$ eV Å^2, and (b) for five different values of D_a, at $H_{\rm eff1} = 80$ kOe. Solid data points are from measurements shown in Fig. 1.

Equation (38) describes a resonance coupling process for spin-current conduction between m_1 and m_2. The interface moment m_1, under STT drive, precesses at its own frequency $H_{\rm eff1}$ which is significantly higher than the "slow dynamics" of the rest of the FL. For a given mode, $H_{\rm ksw}$ as defined by Eq. (35) is a function of diameter a as well as the mode structure. For the lowest-lying nonuniform standing-wave with edge-pinning boundary, $s \approx 4.8$ in Eq. (35). This mode's eigenfrequency depends on diameter a. As it approaches m_1's $H_{\rm eff1}$ at small device size, the resonance condition improves, and the conductance $(1/r_{\rm Amm})$ increases, making it more efficient for the interface moments to couple their spin current into the rest of the FL, resulting in less spin-pumping loss.

More conveniently for comparison with experiments, one can examine the intercept-to-slope ratio in Eqs. (25) and (1), which we call $r^*_{\rm Am}$:

$$r^*_{\rm Am} = \frac{r_{\rm Amm}}{1+\left(\frac{e^2\epsilon_j}{\hbar\mu_B}\right)(M_{s1}t_1\alpha_1)r_{\rm Amm}} = \frac{A_1}{A_0}. \tag{39}$$

Assuming the same materials parameter set as used in Sec. V D, together with an exchange-stiffness value of $D_a = 0.3$ eV Å^2, one has the resulting numerical results for $r^*_{\rm Am}$ versus device diameter a for a set of different $H_{\rm eff1}$ assumptions plotted in Fig. 7. These illustrate the effect of small

diameter or large exchange stiffness D_a as increasing effective precessional spin-current coupling, reducing r^*_{Am}.

Model results in Fig. 7(b) can be compared with our experimental data of Fig. 1(c). These are shown in Fig. 7(b) as solid symbols. The measurement data are sparse, while the model is extremely simplistic without including any distributions of quantities such as H_{eff1} and D_a, which would affect the shape of the size dependence. Nevertheless, an agreement of the general magnitude of the RA dependence and the trend direction is seen.

One may also estimate from the observed intercept Fig. 1(d) the values of the product of $M_{s1}t_1\alpha_1$ according to Eqs. (25) and (39). The fact one observes no significant dependence of the intercept term A_0 on that of A_1 (which is shown to be strongly size dependent) implies

$$\left(\frac{e^2\epsilon_j}{\hbar\mu_B}\right)(M_{s1}t_1\alpha_1)r_{\mathrm{Amm}} \ll 1. \tag{40}$$

This is qualitatively consistent with our assumption of $M_{s1}t_1$ being small. One may further deduce an upper limit for the interface moment's damping coefficient to be of the order 10^{-4} to be consistent with our other assumptions of $M_{s1}t_1$ values. Bear in mind, however, all such estimates are only an order-of-magnitude discussion, and one should not overinterpret values too quantitatively.

VII. DISCUSSIONS

In principle, there are many other possible coupling terms in Eq. (30) into higher-order spin-wave modes of the FL. The task of quantifying contributions from those terms is beyond the scope of this paper. These higher-order modes are likely to have more complex location-dependent symmetries, making conduction of m_1 exchange-driven spin current less efficient when integrated over the entire junction area. Also, higher-order modes are likely to have more losses before their spin contribution is finally coupled back into the low-energy uniform rotation mode that describes the macroscopic reversal. A more accurate analytical description will also face challenges of large-amplitude nonlinear and non-steady-state processes. Those will be extremely sensitive to details of material inhomogeneity length scales and strengths.

Fundamentally, our experiments and model discussions based on such observations presented above only require the moment for m_1 to be (a) of high-energy (high-frequency) nature, and (b) directly coupled to the interface tunnel conductance. This will not uniquely determine m_1's spatial distribution. These moments can be somewhat uniformly distributed at the junction interface, i.e., an "interface-layer" picture, as illustrated by Fig. 4. Alternatively, they may exist in the form of simultaneous "hot spots" at the MgO | CoFeB FL interface [31], where the anisotropy and tunnel conductance are concentrated within small volumes of the FL, while everywhere else, the rest of the FL covering the junction area only provide limited magnetic exchange coupling, and no coupling to tunnel conductance nor significant interface PMA. To resolve the materials nature of such interface moment will require more structural characterization and feedback on correlations between film structure and growth condition, and the resulting physical properties of the end device.

The uncertainty of physical structure also means a lack of direct knowledge for many of the model parameters used in this work, parameters such as those listed in Table I. The values used are not unreasonable based on our understandings of these materials and their behaviors, such as switching error distributions, and details of their FMR [32]. However, those observations together can only at best provide a "reasonable range" of values, and are insufficient to completely constrain the model parameters. This is likely an ongoing challenge to bridge our understanding and the physical nature of these MTJs.

Because of these factors, the aim of this paper is only to provide some extremely oversimplified models for the grasping of the leading-order cause and effects. The real-life quantitative behavior is more complex. Beyond this qualitative understanding presented here, a full micromagnetics simulation may be a tool better suited to reveal quantitative behaviors for any specific-materials constructs and parameters.

VIII. CONCLUSION

Our CoFeB | MgO | CoFeB-based MTJ samples exhibit a spin-torque switching efficiency E_b/I_{c0} that depends both on junction size and tunnel barrier RA. An empirical relationship from experiment is obtained that states $I_{c0}/E_b = A_0 + A_1/r_{\mathrm{Amtj}}$, with $A_1 \propto a - a_0$, where a is the junction diameter, and $a_0 \sim 10$ nm in this set of our samples.

The form of I_{c0}/E_b being proportional with $1/r_{\mathrm{Amtj}}$ is consistent with a spin-pumping-like spin-current loss. But, these MTJs typically have RA values $r_{\mathrm{Amtj}} \gtrsim 2\ \Omega\mu\mathrm{m}^2$. These are at least $100\times$ larger than what would give rise to significant spin-pumping loss compared to intrinsic damping of the FL in a macrospinlike uniform rotation process.

These data presented here, together with earlier experimental observations in CoFeB PMA thin-film MTJ structures, such as an abnormal hard-axis magnetoresistance, and an apparent high-order PMA coefficient [16,31], and the knowledge of an MgO-CoFeB interface with concentrated PMA energy density [15,33], these all point to the existence of interface magnetic moments responsible for spin-polarized tunneling that are under high PMA energy density and are with dynamic time scales much shorter than the whole of the FL.

A toy model, the next simplest model to that of macrospin rotation, is developed to represent such time-scale separation between the interface and bulk moments of the FL. It is done by constructing a coupled two-macrospin system, or we call it the m_1-m_2 model. The m_1-m_2 model can adequately describe the magnitude of the spin-pumping-like spin-current loss due to the tunnel conductance determined by the MTJ's RA. This is fundamentally due to the high-frequency interface moment dynamics, and the resulting division of spin-current transport between the exchange-coupled path from m_1 to m_2, and the spin-pumping path of m_1 across tunnel barrier. The m_1-m_2 model, however, would not explain the observed MTJ size dependence of the switching efficiency nor the size dependence of the efficiency's RA dependence.

The junction-size-dependent STT switching efficiency can be semiquantitatively described by a modified m_1-m_2 model, taking into account approximately the coupling between m_1's high-frequency dynamics and m_2's standing-wave modes in addition to its macrospin uniform rotation mode. The treatment

is only semiquantitative: it accounts for the leading-order effects only. The result gives a correct order-of-magnitude estimate as well as the direction for the STT switching efficiency's size dependence.

This work highlights the role the interface-moment high-frequency dynamics plays in STT excitation and switching. These high-energy processes originate from the presence of interface-concentrated PMA, and a relatively weak exchange coupling between these interface-located moments and the rest of the FL.

There are other mechanisms at play that could give rise to a small apparent RA dependence of the switching efficiency, mechanisms that we did not address in this paper. Chief among them is measurement-related junction Joule heating. This would result in a slight nonconstant temperature threshold mapping for I_{c0}, and would also affect the extracted value of E_b. A full evaluation of the role Joule heating could play in such measurements is beyond the scope of this paper. The methods one chose to evaluate E_b may also affect the resulting value of switching efficiency. We are aware of at least one published result [34] where for two values of $r_{\mathrm{Amtj}} = 6$ and $12\ \Omega\mu\mathrm{m}^2$, the efficiency of devices of similar sizes was nearly the same.

ACKNOWLEDGMENTS

Work was done with the MRAM group at IBM T. J. Watson Research Center in Yorktown Heights, New York, and is supported in part by partnership with Samsung Electronics.

[1] 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, Phys. Rev. B **84**, 064413 (2011).

[2] J. Z. Sun, P. L. Trouilloud, M. J. Gajek, J. Nowak, R. P. Robertazzi, G. Hu, D. W. Abraham, M. C. Gaidis, S. L. Brown, E. J. O'Sullivan *et al.*, J. Appl. Phys. **111**, 07C711 (2012).

[3] H. Sato, M. Yamanouchi, K. Miura, S. Ikeda, H. D. Gan, K. Mizunuma, R. Koizumi, F. Matsukura, and H. Ohno, Appl. Phys. Lett. **99**, 042501 (2011).

[4] H. Sato, M. Yamanouchi, K. Miura, S. Ikeda, R. Koizumi, F. Matsukura, and H. Ohno, IEEE Magn. Lett. **3**, 3000204 (2012).

[5] X. Wang, W. Zhu, Y. Zheng, Z. Gao, and H. Xi, IEEE Trans. Magn. **45**, 3414 (2009).

[6] M. Gajek, J. J. Nowak, J. Z. Sun, P. L. Trouilloud, E. J. O'Sullivan, D. W. Abraham, M. C. Gaidis, G. Hu, S. Brown, Y. Zhu *et al.*, Appl. Phys. Lett. **100**, 132408 (2012).

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

[8] J. C. Slonczewski and J. Z. Sun, J. Magn. Magn. Mater. **310**, 169 (2007).

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

[10] J. Z. Sun, *Handbook of Spintronics* (Springer, Dordrecht, 2014).

[11] D. C. Worledge and P. L. Trouilloud, Appl. Phys. Lett. **83**, 84 (2003).

[12] J. Z. Sun, S. L. Brown, W. Chen, E. A. Delenia, M. C. Gaidis, J. Harms, G. Hu, X. Jiang, R. Kilaru, W. Kula *et al.*, Phys. Rev. B **88**, 104426 (2013).

[13] B. Heinrich, Y. Tserkovnyak, G. Woltersdorf, A. Brataas, R. Urban, and G. E. W. Bauer, Phys. Rev. Lett. **90**, 187601 (2003).

[14] Y. Tserkovnyak, A. Brataas, and G. E. W. Bauer, Phys. Rev. Lett. **88**, 117601 (2002).

[15] 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).

[16] J. Z. Sun, Phys. Rev. B **91**, 174429 (2015).

[17] I. Barsukov, Y. Fu, A. M. Gonçalves, M. Spasova, M. Farle, L. C. Sampaio, R. E. Arias, and I. N. Krivorotov, Appl. Phys. Lett. **105**, 152403 (2014).

[18] Y. Fan, K. J. Smith, G. Lüpke, A. T. Hanbicki, R. Goswami, C. H. Li, H. B. Zhao, and B. T. Jonker, Nat. Nanotechnol. **8**, 438 (2013).

[19] M. Konoto, H. Imaura, T. Taniguchi, K. Yakushiji, H. Kubota, A. Fukusima, K. Ando, and S. Yuasa, Appl. Phys. Exp. **6**, 073002 (2013).

[20] E. C. I. Enobio, H. Sato, S. Fukami, F. Matsukura, and H. Ohno, IEEE Magn. Lett. **6**, 5700303 (2015).

[21] C. Tiusan, M. Sicot, J. Faure-Vincent, M. Hehn, C. Bellouard, F. Montaigne, S. Andrieu, and A. Schuhl, J. Phys.: Condens. Matter **18**, 941 (2006).

[22] J. C. Sankey, Y. T. Cui, J. Z. Sun, J. C. Slonczewski, R. A. Buhrman, and D. C. Ralph, Nat. Phys. **4**, 67 (2008).

[23] J. C. Slonczewski, Phys. Rev. B **39**, 6995 (1989). .

[24] J. Faure-Vincent, C. Tiusan, C. Bellouard, E. Popova, M. Hehn, F.Montaigne, and A. Schuhl, Phys. Rev. Lett. **89**, 107206 (2002).

[25] E. Popova, C. Tiusan, A. Schuhl, F. Gendron, and N. A. Lesnik, Phys. Rev. B **74**, 224415 (2006).

[26] R. I. Joseph and E. Schömann, J. Appl. Phys. **36**, 1579 (1965).

[27] G. N. Kakazei, P. E. Wigen, K. Y. Guslienko, V. Novosad, A. N. Slavin, V. O. Golub, N. A. Lesnik, and Y. Otani, Appl. Phys. Lett. **85**, 443 (2004).

[28] K. Y. Guslienko, S. O. Demokritov, B. Hillebrands, and A. N. Slavin, Phys. Rev. B **66**, 132402 (2002).

[29] R. E. Arias and D. L. Mills, Phys. Rev. B **75**, 214404 (2007).

[30] M. Beleggia, M. D. Graef, and Y. T. Millev, J. Phys. D: Appl. Phys. **39**, 891 (2006).

[31] A. A. Timopheev, R. Sousa, M. Chshiev, H. T. Nguyen, and B. Dieny, Sci. Rep. **6**, 26877 (2016).

[32] C. J. Safranski, Y.-J. Chen, I. N. Krivorotov, and J. Z. Sun, Appl. Phys. Lett. **109**, 132408 (2016).

[33] S. Baumann, F. Donati, S. Stepanow, S. Rusponi, W. Paul, S. Gangopadhyay, I. G. Rau, G. E. Pacchioni, L. Gragnaniello, M. Pivetta *et al.*, Phys. Rev. Lett. **115**, 237202 (2015).

[34] G. Jan, Y.-J. Wang, T. Moriyama, Y.-J. Lee, M. Lin, T. Zhong, R.-Y. Tong, T. Torng, and P.-K. Wang, Appl. Phys. Express **5**, 093008 (2012).