Observation of thermal spin-transfer torque via ferromagnetic resonance in magnetic tunnel junctions

Zhaohui Zhang,[1,*] Lihui Bai,[1] Xiaobin Chen,[2] Hong Guo,[2] X. L. Fan,[3] D. S. Xue,[3] D. Houssameddine,[4,†] and C.-M. Hu[1]

[1]*Department of Physics and Astronomy, University of Manitoba, Winnipeg R3T 2N2, Canada*

[2]*Centre for the Physics of Materials and Department of Physics, McGill University, Montréal, Québec H3A 2T8, Canada*

[3]*The Key Laboratory for Magnetism and Magnetic Materials of Ministry of Education, Lanzhou University, Lanzhou 730000, People's Republic of China*

[4]*Everspin Technologies, 1347 North Alma School Road, Chandler, Arizona 85224, USA*

(Received 11 March 2016; revised manuscript received 20 June 2016; published 11 August 2016)

The thermal spin-transfer torque (TSTT) in magnetic tunneling junctions (MTJs) was systematically studied using electrical detection of ferromagnetic resonance (FMR). Evidence for the existence of TSTT in MTJs is observed. A temperature difference was applied across an MTJ acting as a TSTT on the free layer of the MTJ. The FMR of the free layer was then excited by a microwave current and electrically detected as a dc voltage. We found that the FMR line shape was changed by the TSTT, indicated by the ratio of dispersive and Lorentz components of the FMR spectra (D/L). D/L increases by increasing the temperature difference. In addition, we analyze the magnetization orientation dependence of TSTT and provide solid evidence that this dependence differs from the magnetization orientation dependence of spin-transfer torque driven by a dc bias.

DOI: 10.1103/PhysRevB.94.064414

I. INTRODUCTION

Spin-transfer torque (STT) is an effect applied onto the magnetization of a ferromagnetic (FM) material by a spin current. It is generated by the exchange interaction of the net spin in the current and the magnetization localized in the material when a spin-polarized current or a pure spin current passes through it [1]. The first report of STT was published in 1978 by studying the effects on domain wall in ferromagnets [2], and STT was studied in spin valves by Slonczewski [3] and Berger [4] independently in 1996. After that, STT reattracted people's attention due to the improvement of nanostructure technology in the early 1990s [5–8].

A magnetic tunneling junction (MTJ) is a sandwichlike structure containing two FM layers with a tunneling layer in between. It is an ideal device to study the physics of STT. When a current passes through an MTJ, one of the FM layers will act as a spin polarizer and the spin-polarized current will apply an STT on the other FM layer. Besides the interest in fundamental research, the study of STT in MTJs also has huge potential for the industry since MTJs have been widely used in nonvolatile random memories [9–15]. It has been proven that switching the magnetization configuration via STT is efficient and beneficial to the integration of circuits [6,16–21]. However, the threshold current density for switching is still as high as $10^7 \mathrm{A/cm^2}$. The heat current, as the third flow after charge and spin flow during electron transport, has been considered recently to reduce this high current density for switching. A temperature gradient can generate a spin current carrying spin momentum, which can, in turn, interact with localized spin by applying STT on it. This STT is generated by a temperature gradient, so it is called thermal STT (TSTT).

TSTT has been explored in many theoretical works [22–26]. In 2007, Hatami *et al.* predicted that a temperature gradient would induce an STT, which could excite a magnetization [22]. Later, it was predicted that a temperature gradient of 0.2 K/nm is as efficient as a charge current density of 10^7 A/cm^2 for typical permalloy structures [24], indicating the potential of using TSTT for switching in ferromagnetic materials. In 2011, Jia *et al.* predicted TSTT in MgO-based MTJs showing that TSTT may theoretically be very large in a three monolayer MgO-based MTJ. They found that TSTT is angular dependent to the relative angle of the magnetization between two FM layers in MTJs but very skewed [27], and this TSTT in MgO strongly depends on the thickness and the roughness of the tunneling barrier [28]. Later, Leutenantsmeyer *et al.* showed the possibility of switching in MTJs via TSTT with femtosecond lasers [29].

In contrast, there are a few studies regarding the observation of TSTT in magnetic multilayer structures. The first experimental evidence of TSTT was found by Yu *et al.* on Co/Cu/Co spin valves via studying the moving of FMR resonance position [30]. The evidence of TSTT in the same system was also observed by Fitoussi *et al.* [31], via magnetoresistance change due to out-of-equilibrium magnetization under TSTT. Recently, Pushp *et al.* observed that the STT would affect the switching field of an MTJ from antiparallel to parallel configuration. They attributed the origin of the thermal torque to the asymmetry of the resistance of an MTJ at positive and negative dc voltage bias [32].

There are two main difficulties preventing the observation of TSTT. One difficulty is that it is hard to generate a local temperature difference across several nanometers. Basically, two heating methods can be used to achieve a large enough temperature difference, utilizing the Joule heating when a current passes through the sample [30,33] or using an external heating source [34]. The other difficulty is that the effects of TSTT are usually very weak. Ferromagnetic resonance (FMR) is a sensible technique to probe the magnetization of ferromagnetic materials. Around resonance positions, small

*zhaohui@physics.umanitoba.ca

†Present address: GLOBALFOUNDRIES, 60 Woodlands, Industrial Park D, Street 2, 73840 Singapore.

effects of TSTT on magnetization will be magnified significantly because of this resonance nature. Thus, the properties of TSTT can be studied by carefully analyzing the FMR spectra such as the line shape, resonance position, and linewidth. Furthermore, electrical detection was employed to enhance the signal-to-noise ratio since it is an effective way to detect the FMR based on spin-rectification effect [35–37], which can detect a voltage signal as weak as nV by using lock-in technique.

In this paper, we employed an external laser heating to generate a temperature difference across MTJs and electrical detection of FMR to enhance the signal-to-noise ratio. By measuring and analyzing the FMR spectra at various temperature differences and magnetization configurations, the TSTT was systematically studied. The rest of the paper is organized as follows: The basic theory about TSTT and electrical detection of FMR are summarized in Sec. II, the experiment details are described in Sec. III, and the results and analysis are shown in Sec. IV followed by the conclusion in Sec. V.

II. THEORY

A. Thermal spin-transfer torque

In electron transport, following Onsager's law, the interaction between charge (J_C), spin (J_S), and heat (J_Q) current densities can be represented by a 3×3 matrix [25]:

$$\begin{pmatrix} J_C \\ J_S \\ J_Q \end{pmatrix} = \sigma \begin{pmatrix} 1 & P & ST \\ P & 1 & P'ST \\ ST & P'ST & \kappa T/\sigma \end{pmatrix} \begin{pmatrix} \nabla\mu_C/e \\ \nabla\mu_S/2e \\ -\nabla T/T \end{pmatrix}, \quad (1)$$

where σ is the electrical conductivity; P and P' are spin polarization of the conductivity and its energy derivative; S is the Seebeck coefficient; T and ∇T are the temperature and temperature gradient, respectively; κ is the thermal conductivity; e is the elementary charge; μ_C is the charge electrochemical potential; and μ_S is the spin chemical potential.

The density of spin-transfer torque τ_d is defined as the opposite of the divergence of spin current density:

$$\tau_d = -\nabla \cdot J_S. \quad (2)$$

τ_d is nonzero when there is a spin source or sink. Specifically, in electron transport, if the direction of the spin carried by the spin current is noncollinear with the direction of the magnetic momentum of the material, the spin current will align to the direction of the material's magnetic momentum due to exchange interaction. Conversely, STT will be applied on the localized magnetic momentum.

From Eq. (1), one can see that J_S can be driven by not only a gradient of electrochemical potential $\nabla\mu_C$ but also a temperature gradient ∇T. To distinguish those two torques, the STT driven by $\nabla\mu_C$ is called dc biased STT, since it is always achieved by applying a dc bias voltage, and the STT driven by ∇T is called TSTT.

Hatami *et al.* derived an expression for the TSTT in spin valves showing that the STT generated by a voltage ΔV and a temperature difference ΔT between the two ends of a spin valve are additive and can be represented as [22]

$$\tau \propto (P\Delta V + P'S\Delta T). \quad (3)$$

In contrast to P, $|P'|$ is not bounded, and $P'S$ can be very large [22], which means a small ΔT may generate a torque as large as ΔV can. Although Hatami *et al.*'s expression comes from a spin accumulation model where only the in-plane torque was considered, the result that STT can be generated by ΔV and ΔT is still useful in an MTJ. It is worth noting that, in an MTJ, both the in-plane and out-of-plane torque should be considered, and the out-of-plane STT follows a quadratic dependence against ΔV without a linear term [38].

In summary, either $\nabla\mu_C$ or ∇T can drive J_S. In a magnetic system with spin source or sink, STT will be applied to the magnetic momentum. The dc biased STT has been well studied [38,39], and in this paper we will show that the angular dependence of TSTT affected FMR line-shape change is very different from the change made by dc biased STT.

B. FMR under thermal spin-transfer torque and rectification voltage V_r

Figure 1 shows a sketch of the TSTT in our MTJ. From left to right, there is the free FM layer, tunneling battier, and fixed FM layer. $\mathbf{m}$ and $\mathbf{M}$ are the magnetization of the free and the fixed FM layer, respectively. θ is the angle made by $\mathbf{m}$ and $\mathbf{M}$. Later, we will treat STT as two components, which link to the symmetrical and asymmetrical line shapes. The two components are determined by the cross product $\hat{\mathbf{m}} \times (\hat{\mathbf{M}} \times \hat{\mathbf{m}})$ or $\hat{\mathbf{M}} \times \hat{\mathbf{m}}$, respectively. Thus the relative direction between $\hat{\mathbf{m}}$ and $\hat{\mathbf{M}}$ influences the direction of the torques. Here $\hat{\mathbf{m}}$ and $\hat{\mathbf{M}}$ are the unit vectors of $\mathbf{m}$ and $\mathbf{M}$, respectively. By the coordinate system shown in Fig. 1, the film is located in the y-z plane with the easy axis along the z direction. The positive x axis is from the fixed FM layer to the free FM layer. Since both $\hat{\mathbf{M}}$ and $\hat{\mathbf{m}}$ are in the y-z plane, the cross product $\hat{\mathbf{M}} \times \hat{\mathbf{m}}$ is along the x axis. The polarities of θ were defined as the same as the polarities of $\hat{\mathbf{M}} \times \hat{\mathbf{m}}$, i.e., $\theta > 0$ when $\hat{\mathbf{M}} \times \hat{\mathbf{m}}$ along the positive x axis and $\theta < 0$ when $\hat{\mathbf{M}} \times \hat{\mathbf{m}}$ along the negative x axis.

When there is no external torque, both $\mathbf{M}$ and $\mathbf{m}$ are aligned along the directions of their effective fields, which are determined by the applied magnetic fields, the anisotropy, and the demagnetization. FMR of the free layer can be excited by a microwave (MW) when the frequency of the MW equals the

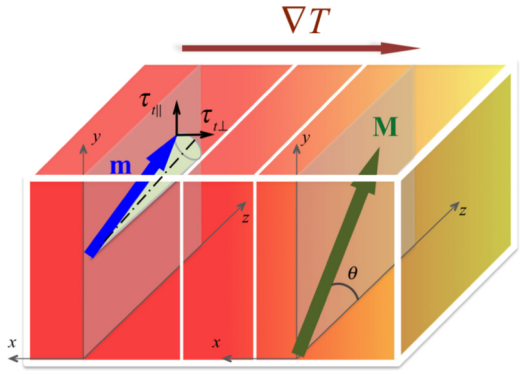

FIG. 1. Sketch of the thermal spin-transfer torque in an MTJ. $\mathbf{M}$ and $\mathbf{m}$ are the magnetization of the fixed and free layer, respectively. $\tau_{t\parallel}$ and $\tau_{t\perp}$ are the in-plane and out-of-plane torque generated by the temperature gradient ∇T.

resonance frequency determined by the effective fields. The precession of $\mathbf{m}$ under magnetic field can be described quantitatively by the Landau-Lifshitz-Gilbert (LLG) equation as

$$\frac{d\hat{\mathbf{m}}}{dt} = -\gamma\hat{\mathbf{m}} \times \mathbf{H} + \alpha\hat{\mathbf{m}} \times \frac{d\hat{\mathbf{m}}}{dt}, \tag{4}$$

where γ is the gyromagnetic ratio, α is the Gilbert damping parameter, and $\mathbf{H}$ is the total applied magnetic field. The STT has two components corresponding to two scattering mechanisms [1]. One component is the in-plane torque, $\tau_\parallel$, which is along the direction of $\hat{\mathbf{m}} \times (\hat{\mathbf{M}} \times \hat{\mathbf{m}})$, and the other component is the out-of-plane torque, $\tau_\perp$, which is along the direction of $\hat{\mathbf{M}} \times \hat{\mathbf{m}}$. Similarly as shown in Ref. [38], we can separate STT into $\tau_{\alpha\parallel}$ and $\tau_{\alpha\perp}$ where $\alpha = b$ and $\alpha = t$ denote STT generated by dc bias and temperature difference, respectively. $\tau_{b\parallel}$ and $\tau_{b\perp}$ can be added on the right side of Eq. (4) as

$$\frac{d\hat{\mathbf{m}}}{dt} = -\gamma\hat{\mathbf{m}} \times \mathbf{H} + \alpha\hat{\mathbf{m}} \times \frac{d\hat{\mathbf{m}}}{dt} - \gamma\frac{\tau_{b\parallel}(I,\theta)}{M_s}\frac{\hat{\mathbf{m}} \times (\hat{\mathbf{M}} \times \hat{\mathbf{m}})}{\sin\theta} - \gamma\frac{\tau_{b\perp}(I,\theta)}{M_s}\frac{\hat{\mathbf{M}} \times \hat{\mathbf{m}}}{\sin\theta}, \tag{5}$$

where M_s is the saturation magnetization of the free FM layer. In our experiment, we measured a dc voltage V_r rectified by the MW and the periodic resistance of the MTJ. V_r can be obtained by expending the time-dependent voltage on the two electrodes of the MTJ and taking the time average over a period (see the Appendix for detailed deduction):

$$V_r = \frac{1}{4}\frac{\partial^2 V}{\partial I^2}I_{\mathrm{RF}}^2 + \frac{I_{\mathrm{RF}}^2}{4M_s}\frac{\partial^2 V}{\partial\theta\partial I}\sqrt{1+\frac{M_0}{H_r}}\frac{1}{\Delta H} \times \left[D(H)\sqrt{1+\frac{M_0}{H_r}}\left(\left.\frac{d\tau_{b\perp}}{dI}\right|_{I=0} - \left.\frac{d\tau_{b\parallel}}{dI}\right|_{I=0}\cot\theta\cdot\delta m_x\right) - L(H)\left(\left.\frac{d\tau_{b\parallel}}{dI}\right|_{I=0} - \left.\frac{d\tau_{b\perp}}{dI}\right|_{I=0}\cot\theta\cdot\delta m_x\right)\right]. \tag{6}$$

Here I_{RF} is the amplitude of the MW current; H_r is the resonance magnetic field; $D(H) = \frac{\Delta H(H-H_r)}{(H-H_r)^2+\Delta H^2}$ and $L(H) = \frac{\Delta H^2}{(H-H_r)^2+\Delta H^2}$ are the dispersive (asymmetrical) and Lorentz (symmetrical) resonance line shape, respectively; $\frac{d\tau_{b\parallel}}{dI}$ and $\frac{d\tau_{b\perp}}{dI}$ are the in-plane and out-of-plane "torkance"; and $I = 0$ indicates there is no dc charge current;

$$\Delta H = \sqrt{1+\frac{M_0}{H_r}}\left[\alpha\left(H_r + \frac{M_0}{2}\right) - \frac{1}{2M_s}\frac{d\tau_{b\parallel}}{d\theta} - \alpha\frac{1}{2M_s}\frac{d\tau_{b\perp}}{d\theta}\right] \tag{7}$$

is the linewidth of the resonance. Here α is the damping of the material and $M_0 \approx N_x M_s$ where N_x is the demagnetization factor of the x direction. δm_x is the disturbance applied on $\mathbf{m}$ when there is a temperature difference across the sample. δm_x can be expressed as [31]

$$\delta m_x = \frac{1}{M_s}\sum_{\beta=\perp,\parallel}\chi_{x\beta}\frac{\partial\tau_{t,\beta}}{\partial T}\Delta T \tag{8}$$

where $\chi_{x\beta} = \frac{dm_x}{d\tau_{t,\beta}}$ is the magnetic susceptibility tensor of the free layer and $\tau_{t,\beta}$ is the in-plane($\beta = \parallel$) or out-of-plane($\beta = \perp$) TSTT.

The ratio of the amplitudes of dispersive and Lorentz D/L reflects the relation between FMR line shape and TSTT. At zero dc bias, D/L is to a good approximation

$$D/L = \frac{1}{M_s}\sqrt{1+\frac{M_0}{H_r}} \times \cot\theta \times \sum_{\beta=\perp,\parallel}\chi_{x\beta}\frac{\partial\tau_{t,\beta}}{\partial T}\Delta T, \tag{9}$$

where D and L are the amplitudes of dispersive and Lorentz components, respectively. In dc biased STT, this ratio is

$$D/L = \sqrt{1+\frac{M_0}{H_r}}\frac{(d\tau_{b\perp}/dI)|_{I_0=V/R}}{(d\tau_{b\parallel}/dI)|_{I_0=V/R}}, \tag{10}$$

from which we can see that the FMR line shape is adjusted by STT. Here V is the voltage bias and R is the resistance of the MTJ. The ratio D/L is a key feature which we will study later.

III. EXPERIMENTS

A. Samples and adjustment of θ

The MTJ structures we measured were fabricated by Everspin. The wafer was grown on a Si substrate covered with 200-nm SiO_2. The multilayer structures include PtMn(20)/CoFe(2.27)/Ru(0.8)/CoFeB(2.2)/CoFe(0.525)/MgO(1.2)/CoFeB(2.5) (the unit of the values in braces is nm). The MTJs have been annealed postdeposition at 300 °C for 1 h under an in-plane magnetic field to set the direction of the pinned synthetic antiferromagnets.

Measurements in this paper were performed on two samples marked as sample A and sample B with cross-sectional areas of 6.2×10^{-11} and $1.1 \times 10^{-10}\mathrm{cm}^2$, respectively. The long axis of the MTJ was parallel to the pinning direction. All measurements were performed at room temperate unless otherwise mentioned.

Figure 2(a) shows the resistance loop of sample A measured by sweeping the magnetic field at $\phi = 0°$ (gray) and $\phi = 60°$ (red), respectively. ϕ is the angle made by the external magnetic field and the easy axis of the MTJ. The TMR ratios for sample A and B are 52 and 70%, respectively.

The angle between $\mathbf{m}$ and $\mathbf{M}$ can be set by an external magnetic field H, by changing the amplitude of H and the angle ϕ made by H and the easy axis of the magnetization. The angle θ is near 0° when the resistance R is the lowest at large positive H, for example, point A in Fig. 2(a). When sweeping H from positive to negative, a switching will be observed at $H = -10.6$ mT where $\mathbf{m}$ jumps to the opposite direction and θ is near 180°; thus the resistance at that time is the highest, for example, at point B in Fig. 2(a). When H increases negatively, $\mathbf{M}$ will rotate towards H, then θ will decrease from $\theta = 180°$ resulting in the decrease of R. θ is linked with R by

$$R(\theta) = \tfrac{1}{2}(R_P + R_{\mathrm{AP}}) + \tfrac{1}{2}(R_P - R_{\mathrm{AP}})\cos\theta, \tag{11}$$

which can be deduced starting from Julliere's model [40]. Here R_P and R_{AP} are the resistance of parallel (P) and antiparallel (AP) states, respectively. θ at a certain H can be calculated as long as R is known. For example, the resistance at point C is

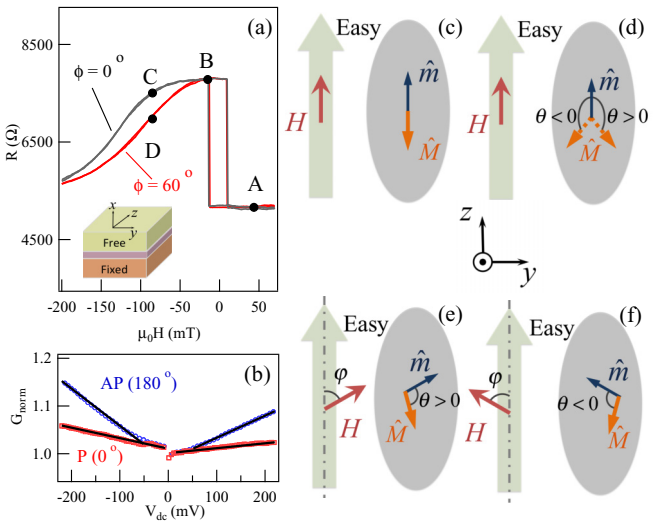

FIG. 2. (a) The resistance loops as a function of H of the MTJ at $\phi = 0^\circ$ (gray) and $\phi = 60^\circ$ (red). There are four points marked as A, B, C, and D. (b) Normalized conductance $G_{\rm norm}$ as a function of dc biased voltage $V_{\rm dc}$ at AP and P states, corresponding to $\theta = 180$ and 0°, respectively. (c) and (d) are the sketches of the configuration of **m** and **M** corresponding to point B and C, respectively. (e) and (f) are the sketches of the configuration of **m** and **M** corresponding to point D when $\theta > 0$ and $\theta < 0$, respectively.

7494 Ω, thus θ at point C is 140.8° with $R_P = 5136$ Ω and $R_{\rm AP} = 7792$ Ω.

Figure 2(b) shows the normalized conductance $G_{\rm norm}$ as a function of biased dc voltage $V_{\rm dc}$, from which the asymmetry of our conductance versus voltage can be clearly seen. The asymmetry at the AP state is more significant than that of the P state. Here, $G_{\rm norm}$ is defined as $G_{\rm norm} = G(V_{\rm dc})/G(V_{\rm dc} \to 0)$ [32], where $G(V_{\rm dc})$ is the conductance of the MTJ and $G(V_{\rm dc} \to 0)$ is the conductance when $V_{\rm dc}$ approaches zero.

Figure 2(c) shows the antiparallel configuration corresponding to point B in Fig. 2(a), where the external magnetic field H is along the direction of the easy axis. By increasing H, **M** will start rotating to the direction of H; however, the direction of rotation is random. An anticlockwise rotation will result in a positive θ, and a clockwise rotation will result in a negative θ as shown in Fig. 2(d).

To set θ with determined polarity, the relative positions of **M** and **m** are controlled by setting H in different directions to the easy axis. As shown in Fig. 2(e), when the projection of H on the y axis is positive, **M** will prefer to rotate anticlockwise, thus θ is positive by increasing H. In contrast, as shown in Fig. 2(f), when the projection of H on the y axis is negative, **M** will prefer to rotate clockwise, thus θ is negative by increasing H.

The amplitude of θ can be determined by using Eq. (11). For example, the resistance of the MTJ at point D in Fig. 2(a) is 6813 Ω. If θ is positive as shown in Fig. 2(d), the magnitude of θ can be calculated as 105.24°.

B. Building a temperature difference across an MTJ

We used a laser heating technique to heat our MTJ to establish a temperature difference, which has already been used to study the Seebeck effect on MTJs by Walter *et al.* [34]. In our measurements, we employed a diode laser with a wavelength of 671 nm and output power range from 0 to 300 mW. The laser beam was focused on the *top* electrode (near the free FM layer) of the MTJ with an area of 100 × 100 μm. The laser spot has a diameter around 50 μm. When there is a temperature difference across the MTJ established by laser heating, a Seebeck voltage V_S can be detected directly on the two ends of the MTJ. To measure the V_S by lock-in technique, the laser beam was modulated with a chopper with the frequency of 20 Hz. The measurement shows that V_S is linearly dependent on the laser power (not shown in this paper). Since V_S is proportional to ΔT ($V_S = S\Delta T$), ΔT is linked with laser power and thus ΔT at certain laser power can be determined by using $S = 50$ μV/K based on our previous work [33].

C. Electrical detection of FMR

Figure 3(a) shows the sketch of electrical detection of FMR under a TSTT. An MW was sent into the MTJ by a coaxial cable and the rectification voltage across the two ends of the MTJ was measured by sweeping the magnetic field around resonance. A bias tee was used to separate the MW and dc signal. A laser beam was employed to heat the top electrode of the MTJ, and a temperature difference ΔT was established across the MTJ, and the FMR measurements were performed at various ΔT.

To exclude the effects made by temperature rise, the temperature dependence of the electrically detected FMR was systematically studied by attaching a Peltier device on the electrode near the fixed FM layer (bottom side) of an MTJ, as shown in Fig. 3(b). The MTJ will be heated by the Peltier device, since the heating area is much larger than that in the laser heating case; after equilibrium by waiting for enough time, the temperature of the whole sample will be raised up. The temperature was detected by a thermal couple attached on the surface of the electrodes of the MTJ.

Figure 3(c) shows the measurement of dc biased STT, where a dc bias was applied to the two ends of an MTJ and the FMR was measured at various dc bias voltages.

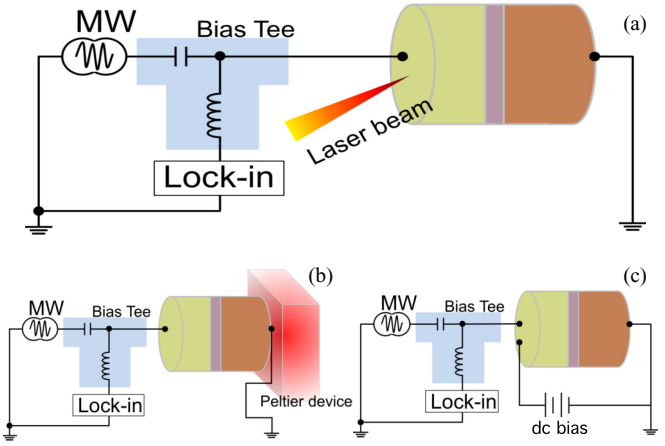

FIG. 3. The measurement setup of (a) laser heating, (b) external heating by Peltier device, and (c) dc bias STT.

IV. RESULTS AND ANALYSIS

A. The dc biased spin-transfer torque

We studied the dc biased STT first. As shown in Fig. 3(c), a dc bias and simultaneously a MW with the frequency of $\omega/2\pi = 6.91$ GHz were applied on sample A. V_r was measured while sweeping H from 82 to 140 mT by lock-in technique with a 8.33-kHz sinusoidal signal modulating the MW. θ was set to 91 and -101° as shown in Figs. 4(a) and 4(e), respectively. The amplitudes of the MW I_{RF} were 12.6 and 10.9 μA at $\theta = 91$ and -101°, respectively.

Figures 4(b)–4(d) are the results at $\theta = 91^\circ$. The gray hollow circles in Fig. 4(b) are the raw FMR spectra measured at zero (middle), positive (top), and negative (bottom) dc bias. In all three cases, there are dominated negative Lorentz components, and the sign of the dispersive components are determined by the polarity of dc voltage bias V_{dc}. All of the spectra were fit by the sum of Lorentz and dispersive and plotted as black curves in Fig. 4(b).

The ratio D/L, as discussed above, was determined by the ratio of the two components of the STT. To clearly show the proportion of the two components, the amplitudes of dispersive and Lorentz (D and L) were fit from Fig. 4(b) and both D and L were normalized by dividing by L. The normalized Lorentz and dispersive components L_{norm} and D_{norm} were plotted in Fig. 4(c), from which one can see that all Lorentz components are negative. At $V_{dc} = 248$ mV, the dispersive component is positive, and at $V_{dc} = -248$ mV the dispersive component is negative. D/L as a function of V_{dc} is shown in Fig. 4(d), from which we can see that D/L increases negatively at positive bias and positively at negative bias.

Figures 4(f)–4(h) are the results at $\theta = -101^\circ$. Similarly, as shown in Fig. 4(f), under zero (middle), positive (top), and negative (bottom) dc bias, there are dominant negative Lorentz components, and the sign of the dispersive components are determined by the polarity of dc voltage bias V_{dc}. The normalized Lorentz and dispersive components L_{norm} and D_{norm} were plotted in Fig. 4(g), from which one can see that all Lorentz components are negative. At $V_{dc} = 198$ mV, the dispersive component is positive, and at $V_{dc} = -198$ mV the dispersive component is negative. D/L as a function of V_{dc} is shown in Fig. 4(h), from which we can see that D/L increases negatively at positive bias and positively at negative bias.

The insets in Figs. 4(d) and 4(h) are the resonance position $\mu_0 H_r$ (left and black) and linewidth $\mu_0 \Delta H$ (right and red) as functions of V_{dc} at $\theta = 91$ and -101°, respectively. For both cases, $\mu_0 H_r$ and $\mu_0 \Delta H$ are nearly a constant at various V_{dc}. We will verify that the experimental results of $\mu_0 \Delta H$ agree with the theoretical calculation now.

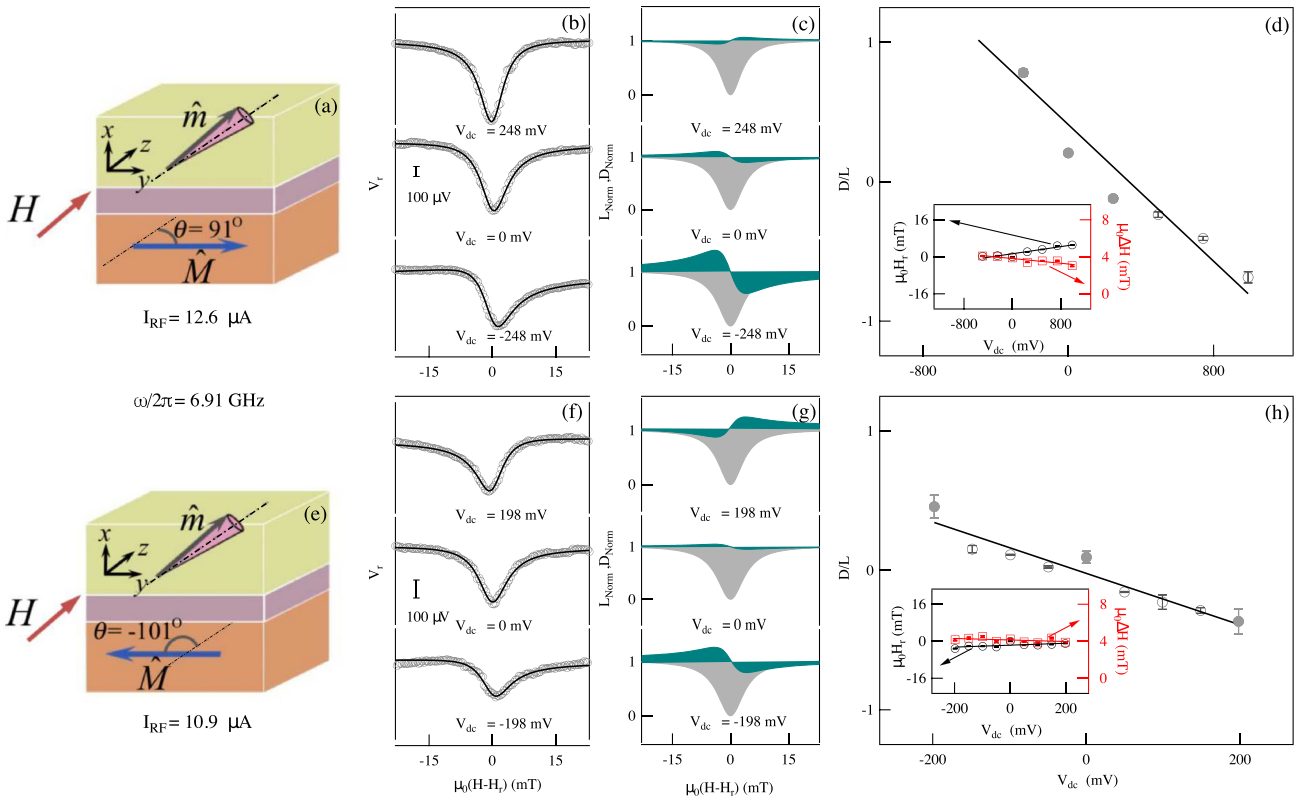

FIG. 4. The coordinates and the magnetization configurations at (a) 90° and (e) -101°. (b)–(d) are results at 91° and (f)–(h) are results at -101°. (b) and (f) are the FMR line-shape evolution with the dc bias in different polarities. The FMR spectrum with no dc bias is in the middle while the FMR spectra with positive and negative dc bias are on the top and bottom, respectively. The gray circles are the measurement results, and the black line is the fittings of the data with the combination of dispersive and Lorentz components. (c) and (g) are the Lorentz (gray) and dispersive (dark cyan) components normalized by L. The cases without a dc bias are in the middle, and the cases with positive and negative dc bias are on the top and bottom in each figure. D/L at various dc bias were plotted in (d) and (h), in which the black lines are linear fittings. The insets of (d) and (h) are the resonance positions $\mu_0 H_r$ (left, black) and the linewidth $\mu_0 \Delta H$ (right, red) of the FMR at various dc bias.

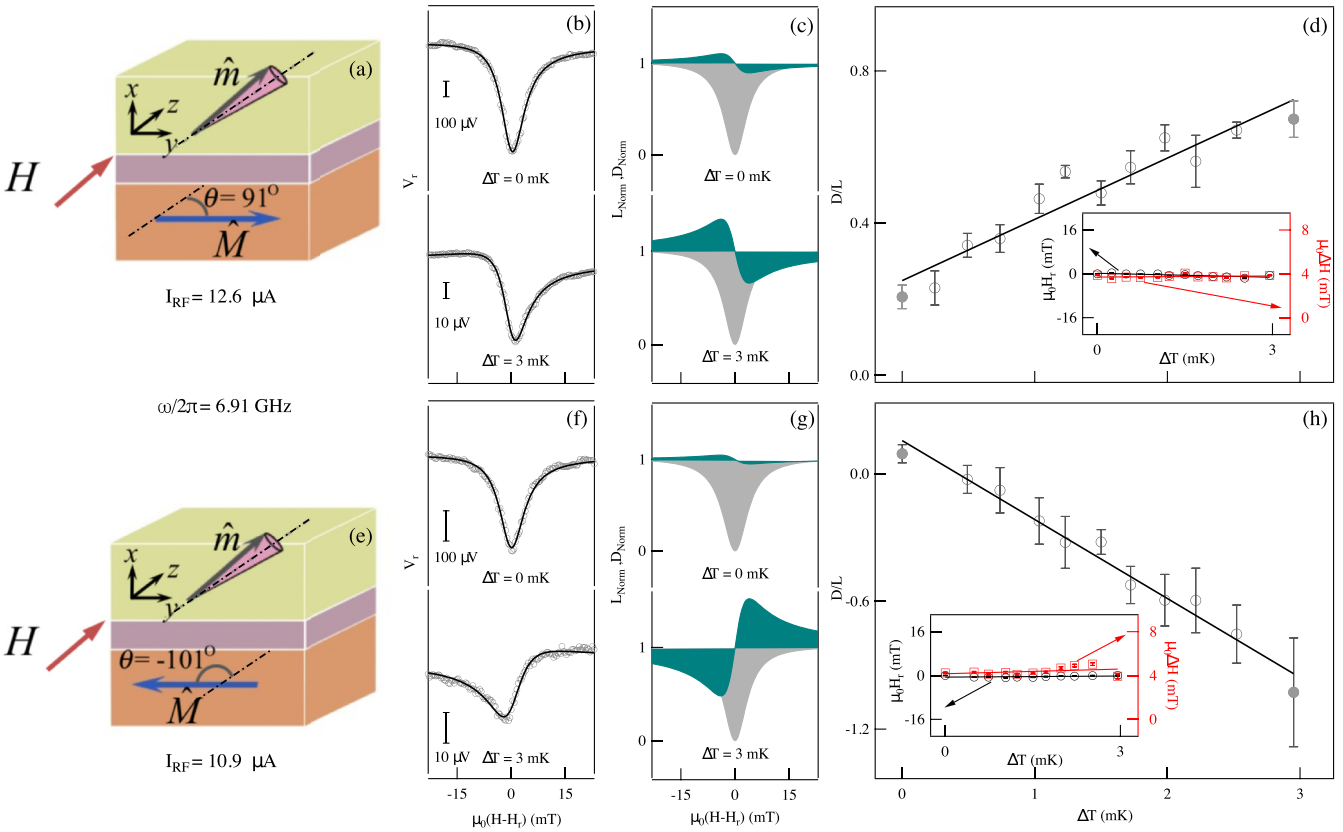

FIG. 5. The coordinates and the magnetization configurations at (a) 90 ° and (e) −101°. (b)–(d) are results at 90°, and (f)–(h) are results at −101°. The top and bottom spectra in (b) and (f) are the FMR line shapes at $\Delta T = 0$ and 3 mK, respectively. The gray circles are the measurement results and the black lines are fittings with a sum of dispersive and Lorentz components. (c) and (g) are the Lorentz (gray) and dispersive (dark cyan) components normalized by the amplitude of L. The curves on the top and bottom correspond to $\Delta T = 0$ and 3 mK, respectively. D/L at various ΔT were plotted in (d) and (h). The solid gray dots are the values corresponding to $\Delta T = 0$ and 3 mK. The insets of (d) and (h) are the resonance positions $\mu_0 H_r$ (left, black) and the linewidth $\mu_0 \Delta H$ (right, red) of the FMR at various temperature differences.

By using the model proposed by Theodonis *et al.* [41], we will have a rough estimation of $(d\tau_{b||}/d\theta)|_{I_0} = 1.3 \times 10^{-5} \cos\theta$ eV/rad and $(d\tau_{b\perp}/d\theta)|_{I_0} = 0.023 \cos\theta$ eV/rad when bias $V = 0.1$ V. Further, by using Eq. (7), with saturation magnetization $M_s/\text{Vol} = 1100$ emu/cm^3 [38], volume of the free layer Vol $= 1.4 \times 10^{-17}$ cm^3, and assuming $\alpha = 0.01$, the calculated $\mu_0 \Delta H$ is 4.6 mT, where STT has little contribution. This agrees well with experimental results of about $\mu_0 \Delta H =$ 4 mT. The slight change of $\mu_0 \Delta H$ can be attributed mainly to the change of effective damping α under various V_{dc} [38].

In summary, at $\theta = 91$ and $-101°$, when there is no dc bias, the FMR spectra are always Lorentz dominated with a small dispersive component. The sign of the dispersive component is determined by the polarity of dc voltage bias V_{dc}. When a positive dc bias is applied, D/L will increase negatively. When a negative dc bias is applied, D/L will increase positively. We will see later that the behavior of D/L under opposite θ in the dc biased STT is very different compared to the TSTT case.

B. Thermal spin-transfer torque

1. Key features of thermal spin-transfer torque

Now we are going to study the effects made by TSTT. As shown in Fig. 3(a), a MW with the frequency of $\omega/2\pi =$ 6.91 GHz was sent into sample A, and the field H was swept from 82 to 140 mT. A temperature difference was applied on the MTJ by laser heating. The MW was modulated with a sinusoidal signal with a frequency of 8.33 kHz and V_r was detected by lock-in technique. θ was set to 91 and $-101°$ as shown in Figs. 5(a) and 5(e), respectively. I_{RF} are 12.6 and 10.9 μA at $\theta = 91$ and $-101°$, respectively.

Figures 5(b)–5(d) are the results at $\theta = 91°$. The gray hollow circles in Fig. 5(b) are the raw FMR spectra, and the black lines are fittings with a sum of the dispersive and Lorentz components. When $\Delta T = 0$ mK, the FMR curve is Lorentz dominated with a small negative dispersive component. When $\Delta T = 3$ mK, the dispersive component increases significantly.

By fitting the raw FMR spectra, D and L were separated, and both components were normalized by dividing by L. The normalized dispersive and Lorentz components D_{norm} and L_{norm} were plotted against H in Fig. 5(c), from which we can see that the Lorentz components are always negative. When $\Delta T = 0$ mK, there is a negative dispersive component, and the dispersive component will become more negative when $\Delta T = 3$ mK. Figure 5(d) shows D/L at various ΔT from 0 to 3 mK. By increasing ΔT, D/L increases positively.

Figure 5(f)–5(h) are the results at $\theta = -101°$. Similarly, the raw spectra in Fig. 5(f) were fit by the sum of the

dispersive and Lorentz components. When $\Delta T = 0$ mK, the FMR curve is Lorentz dominated with a small negative dispersive component. When $\Delta T = 3$ mK, the dispersive component changed to a positive shape. The normalized dispersive and Lorentz components D_{norm} and L_{norm} were plotted against H in Fig. 5(g), from which we can see that the Lorentz components are always negative. When $\Delta T = 0$ mK, there is a small negative dispersive component and the dispersive component will become positive when $\Delta T = 3$ mK. Figure 5(d) shows D/L at various ΔT from 0 to 3 mK. By increasing ΔT, D/L increases negatively.

The change of D/L by increasing ΔT is the most interesting result here. The behavior of D/L under ΔT is very different from that under V_{dc}. In the dc bias STT case we discussed above, by increasing V_{dc} positively, D/L always increases negatively at both positive and negative θ. However, in the TSTT case, by increasing ΔT positively, D/L increases positively at $\theta > 0$ and negatively at $\theta < 0$. We will discuss this θ dependence of D/L under both dc biased STT and TSTT with more details later.

The insets in Figs. 5(d) and 5(h) are the resonance position $\mu_0 H_r$ (left and black) and $\mu_0 \Delta H$ (right and red) as functions of ΔT at $\theta = 91$ and -101°, respectively. For both cases, $\mu_0 H_r$ and $\mu_0 \Delta H$ are nearly a constant at various ΔT, which are evidences that the temperature of the whole sample did not arise significantly. $\mu_0 \Delta H$ can be calculated by the similar procedure we used in the dc biased STT case. Under a temperature difference of 1 K, $(d\tau_{b||}/d\theta)|_{I_0} = -4.4 \times 10^{-9} \cos\theta$, while $(d\tau_{b\perp}/d\theta)|_{I_0}$ does not change much due to the large portion of zero-field out-of-plane torque, we still have calculated $\mu_0 \Delta H \approx 4.6$ mT, which also well agrees with experimental results of $\mu_0 \Delta H = 4$ mT.

2. *Different angular dependence between dc bias and thermal spin-transfer torque*

In this subsection, we will study the angular dependence of D/L under dc bias and thermal STT systematically. The measurements discussed above were performed at a positive angle $\theta = 91^\circ$ and a negative angle $\theta = -101^\circ$. Similar measurements were performed at various θ and in total ten sets of measurements were performed at five positive angles and five negative angles. All the FMR spectra were fitted by the sum of Lorentz and dispersive components. Finally, D/L as a function of ΔT or V_{dc} at various θ were obtained.

Figure 6(a) shows D/L as a function of ΔT at various θ. The lines are linear fittings. It is clear that D/L has a linear dependence on ΔT, and the trend of D/L with increasing ΔT is positive and negative with positive and negative θ, respectively.

Figure 6(c) shows the angular dependence of D/L at $\Delta T = 3$ mK. The solid line in this figure is a guide to the eye. D/L has a negative value when $\theta < 0$ and a positive value when $\theta > 0$. Our result is also consistent with the work in Ref. [32], which indicates that TSTT is larger at θ near 180° than at θ near 0°.

Figure 6(b) shows D/L as a function of V_{dc} at the same θ as we chose in Fig. 6(a).The solid lines are linear fittings. At both positive and negative θ, D/L always increases negatively by increasing V_{dc} positively, and always increases positively

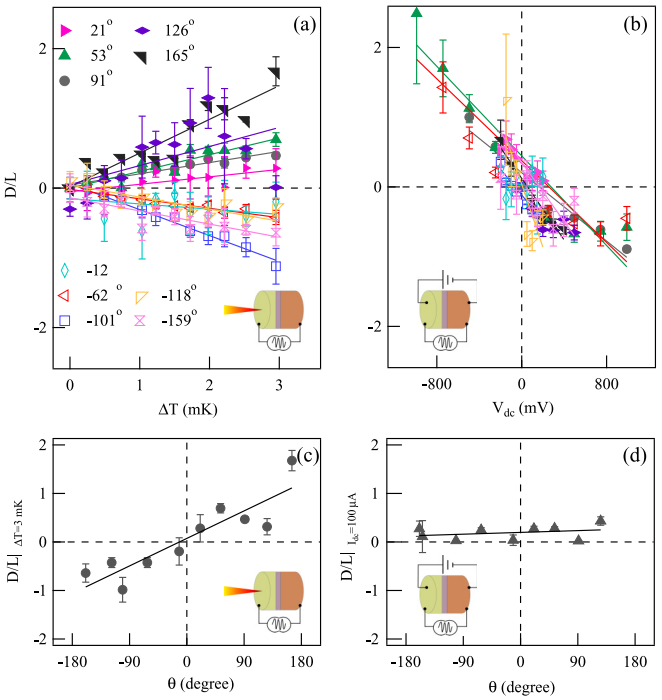

FIG. 6. (a) D/L as a function of (a) ΔT and (b) V_{dc} at various θ. The solid dots correspond to positive θ and the hollow dots correspond to negative θ. Solid lines in (a) and (b) are linear fittings. (c) D/L as a function of θ at $\Delta T = 3$ mK and (d) at $I_{\text{dc}} = 100$ μA. Solid lines in (c) and (d) are guides to the eye.

by increasing V_{dc} negatively. Figure 6(d) shows the angular dependence of D/L at a constant dc current $I_{\text{dc}} = 100$ μA. The solid line in this figure is a guide to the eye. For all the angles, D/L is always positive and nearly a constant comparing to Fig. 6(c).

The angular dependence of D/L under STT generated by V_{dc} was also observed in Ref. [38]. They performed the measurement for angles between 45 and 90° and found that the ratios of in-plane and out-of-plane "torkance" are nearly constant between 45 and 90°. Since D/L is linked to the ratio of two torkances by Eq. (10), and both torkances are $\sin\theta$ dependent, thus D/L is angular independent in the dc biased STT case.

In contrast, in this paper the unique angular dependence of D/L in TSTT has been observed. The origin of the TSTT is the divergence of the spin current and the TSTT measurements were performed in an open circuit so that there is no charge current. This indicates that the temperature difference generated a pure spin current. In Ref. [32], possible reasons for the generation of this spin current were proposed. They found that the magnitude of TSTT was related to the symmetry of the conductance of the MTJ. A larger TSTT exists in a larger asymmetry MTJ and, in our sample, this asymmetry also exists as shown in Fig. 2(b).

Essentially, D/L under TSTT is linked to the thermal torkance $d\tau_{t||}/dT$ and $d\tau_{t\perp}/dT$ by Eq. (9). The temperature difference and angular dependences of D/L are related to the temperature difference and angular dependences of $d\tau_{t||}/dT$ and $d\tau_{t\perp}/dT$. However, those dependences of $d\tau_{t||}/dT$ and

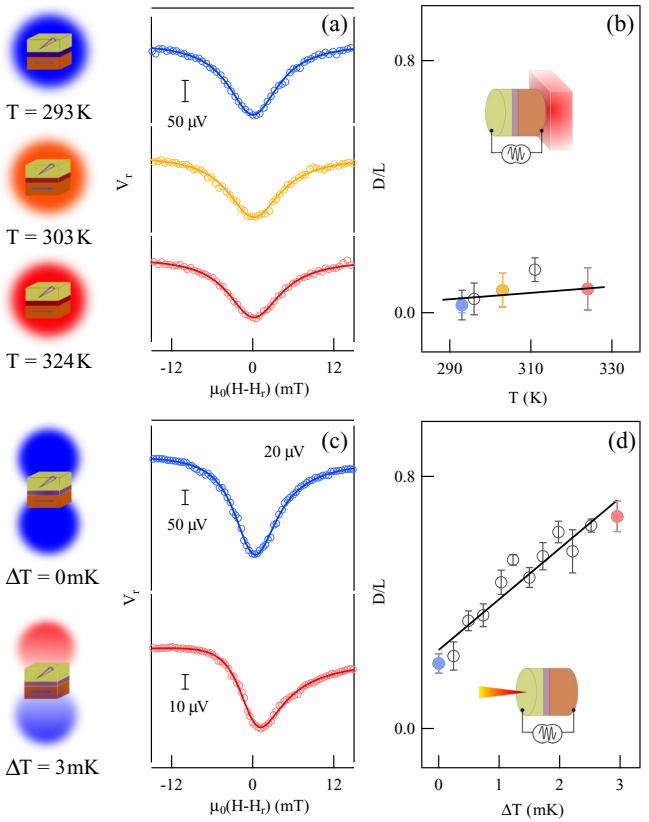

FIG. 7. In the left column, the sketches show that the samples were set into various temperatures or have a temperature difference from the top to the bottom side. (a) The FMR spectra for three different temperatures. (b) D/L of the FMR spectra at various temperatures. The black line is a guide to the eye and the blue, yellow, and red solid dots are D/L at the three temperatures corresponding to the three temperatures in (a). (c) The FMR spectra at different temperature difference ΔT. (d) D/L at various ΔT. The black line is a guide to the eye and the blue and solid red dots are the values of D/L at ΔT of 0 and 3 mK corresponding to the cases in (c).

$d\tau_{t\perp}/dT$ are not so clear yet, and other possible mechanism(s) might be considered.

We would like to point out that if ΔT induces a direct change on $d\tau_{b\perp}/dI$ or $d\tau_{b\parallel}/dI$ our experimental results would have another explanation. However, based on the Seebeck effect, ΔT drives a charge current on the order of nA which is ignorable from the perspective of changing the torques on both components compared to the microwave current ($\sim 10\,\mu$A).

3. *Exclusion of temperature rise*

To exclude the effects made by temperature rise, the temperature dependence of the electrically detected FMR was systematically studied.

The FMR spectra were measured on sample B by heating the sample with a Peltier device as shown in Fig. 3(b), where the bottom side of the sample was well attached to a Peltier device by thermally conductive adhesive. The temperature of the Peltier device could be controlled by a dc current and detected by a thermal couple. At each temperature, the sample was allowed to be heated adequately by waiting for 20 min. Figure 7(a) shows the FMR under three different temperatures. θ was 52° and a MW current with the amplitude of $I_{\rm RF} = 86\,\mu$A was applied. The line shape does not change significantly by increasing temperature. Figure 7(b) shows that D/L is almost independent of temperature T.

For comparison, the laser heating spectra of sample A performed at $\theta = 91°$ were replotted in Fig. 7(c) and D/L as a function of ΔT was plotted in Fig. 7(d). By comparing Figs. 7(b) and 7(d), one can see that D/L is very sensitive to ΔT compared to T. A ΔT as small as 3 mK makes a change up to 0.47 in D/L, while a temperature increase up to 40 K does not make a big change in D/L. This comparison gives the conclusion that the effects on FMR line shapes by laser heating are dominated by temperature difference.

Joule heating was also considered. In the laser heating case, we are comparing the FMR line shape before and after laser heating and the effects of Joule heating in those two cases were the same since the magnitude of the microwave current was fixed. Thus, Joule heating was not involved when we studied the difference of FMR line shape before and after laser heating.

V. CONCLUSION

In conclusion, our work presents evidence of the existence of TSTT in MTJs. We established a temperature difference ΔT across an MTJ by heating one side of the MTJ using a laser beam. This ΔT generated a TSTT on the magnetization of the free FM layer. The FMR line-shape change indicated by the ratio of dispersive and Lorentz components of the FMR spectra (D/L) was studied systematically. D/L is enhanced with larger ΔT due to the enhancement of TSTT. The increasing of D/L under ΔT is dependent on the polarity of the angle of the magnetization between two FM layers. A positive angle corresponds to a positive increase while a negative angle corresponds to a negative increase. The angular dependence of D/L under TSTT differs from the angular dependence under STT driven by a dc bias, indicating that the angular dependences of TSTT and dc biased STT are different. Our work demonstrates that electrically detected FMR can be used as a sensitive tool for the measurement of TSTT in MTJs, making our work of general interest to the spintronics community.

ACKNOWLEDGMENTS

We thank Gerrit Bauer, Ke Xia, and Michael Harder for helpful discussions. This work was funded by Natural Sciences and Engineering Research Council, CFI, and URGP grants (C.-M.H.). Z.Z. was supported by the China Scholarship Council.

APPENDIX: THE DEDUCTION OF RECTIFICATION VOLTAGE V_r UNDER STT

In this Appendix, we will derive the FMR and rectification voltage in MTJs under STT. Generally, the current sent into the sample has two components and can be represented as

$$I(t) = I + \delta I(t) = I + I_{\rm RF}\cos\omega t. \qquad \text{(A1)}$$

Here I is a dc component while $I_{\mathrm{RF}}\cos\omega t$ is an ac component which is a microwave of amplitude I_{RF}. The magnetization of the free layer $\mathbf{m}$ is processing around the $\hat{z}$ direction; thus by ignoring the slight change in the z component we can write $\mathbf{m}$ as

$$\mathbf{m} = (m'_x e^{i\omega t}, m'_y e^{i\omega t}, m'_z), \tag{A2}$$

where m'_x, m'_y, and m'_z are the three components of magnetization in the orthogonal coordinate as shown in Fig. 1. Normalizing the three components of $\mathbf{m}$ by dividing by m'_z taking $m_z = M_s$, where M_s is the saturated magnetization of the free layer, the unit vector of $\mathbf{m}$ can be represented as

$$\hat{\mathbf{m}} = (m_x e^{i\omega t}, m_y e^{i\omega t}, 1), \tag{A3}$$

where $m_x = m'_x/m_z$ and where $m_y = m'_y/m_z$. Similarly, since the magnetization of the fixed layer is in the y-z plane, the unit vector of the fixed layer's magnetization $\hat{\mathbf{M}}$ can be represented as

$$\hat{\mathbf{M}} = (0, \sin\theta, \cos\theta), \tag{A4}$$

where θ is the angle made by $\hat{\mathbf{m}}$ and $\hat{\mathbf{M}}$. Precession of $\hat{\mathbf{m}}$ changes θ, the direction cosine of which can be written as

$$\begin{aligned}\cos\theta(t) &= \hat{\mathbf{m}}\cdot\hat{\mathbf{M}} = \mathrm{Re}(m_x e^{i\omega t}, m_y e^{i\omega t}, 1)(0, \sin\theta_0, \cos\theta_0)\\ &= \cos\theta_0 + \sin\theta_0 \mathrm{Re}(m_y e^{i\omega t}),\end{aligned} \tag{A5}$$

where θ_0 is the intersection angle between $\hat{\mathbf{M}}$ and the $\hat{z}$ axis (assuming that $\hat{\mathbf{M}}$ does not change with the microwave field). By comparing with a Taylor expansion of $\cos\theta(t)$ at θ_0,

$$\cos\theta(t) = \cos\theta_0 - \sin\theta_0\delta\theta,$$

we have

$$\delta\theta = -\mathrm{Re}(m_y e^{i\omega t}). \tag{A6}$$

The voltage we measured is a function of $I(t)$ and θ. For a small microwave current, we can expand the dynamic voltage as [38]

$$\begin{aligned}V = V(I_0,\theta_0) &+ \frac{\partial V}{\partial I}\delta I + \frac{\partial V}{\partial\theta}\delta\theta + \frac{1}{2}\frac{\partial^2 V}{\partial I^2}(\delta I)^2\\ &+ \frac{\partial^2 V}{\partial I\partial\theta}\delta I\delta\theta + \frac{1}{2}\frac{\partial^2 V}{\partial\theta^2}(\delta\theta)^2,\end{aligned} \tag{A7}$$

where $V(I_0,\theta_0)$ is a dc background. The rectification voltage V_r is the time average of the terms related to δI and $\delta\theta$ over one period, since

$$\begin{aligned}\langle\delta I\rangle &= 0,\\ \langle\delta\theta\rangle &= 0,\\ \langle(\delta I)^2\rangle &= I_{\mathrm{RF}}^2\langle\cos^2\omega t\rangle = I_{\mathrm{RF}}^2/2,\\ \langle\delta I\delta\theta\rangle &= -I_{\mathrm{RF}}\langle\cos\omega t\,\mathrm{Re}(m_y e^{i\omega t})\rangle = -I_{\mathrm{RF}}\mathrm{Re}(m_y)/2,\\ \langle(\delta\theta)^2\rangle &= \langle(\mathrm{Re}(m_y e^{i\omega t}))^2\rangle = |m_y|^2.\end{aligned}$$

V_r becomes

$$V_r = \frac{1}{4}\frac{\partial^2 V}{\partial I^2}I_{\mathrm{RF}}^2 - \frac{1}{2}I_{\mathrm{RF}}\mathrm{Re}(m_y)\frac{\partial^2 V}{\partial I\partial\theta} + \frac{1}{4}\frac{\partial^2 V}{\partial\theta^2}|m_y|^2. \tag{A8}$$

To get the value of V_r, m_y needs to be calculated.

The LLG equation for a vector magnetization $\hat{\mathbf{m}}$ is [42]

$$\begin{aligned}\frac{d\hat{\mathbf{m}}}{dt} &= -\gamma\hat{\mathbf{m}}\times\mathbf{H} + \alpha\hat{\mathbf{m}}\times\frac{d\hat{\mathbf{m}}}{dt} - \gamma\frac{\tau_{b\parallel}(I,\theta)}{M_s}\\ &\times\frac{\hat{\mathbf{m}}\times(\hat{\mathbf{M}}\times\hat{\mathbf{m}})}{\sin\theta} - \gamma\frac{\tau_{b\perp}(I,\theta)}{M_s}\frac{\hat{\mathbf{M}}\times\hat{\mathbf{m}}}{\sin\theta},\end{aligned} \tag{A9}$$

where $\tau_{b\parallel}(I,\theta)$ is in-plane STT and $\tau_{b\perp}(I,\theta)$ is out-of-plane STT. (In this equation, the positive current I is assumed to flow from the free layer to the fixed layer.) There is a disturbance δm applied on $\hat{\mathbf{m}}$ by a temperature difference, ΔT, across the MTJ, and $\hat{\mathbf{m}}$ under ΔT can be written as $\hat{\mathbf{m}} \approx (m_x e^{i\omega t} + \delta m_x)\hat{\mathbf{x}} + (m_y e^{i\omega t} + \delta m_y)\hat{\mathbf{y}} + (1+\delta m_z)\hat{\mathbf{z}}$, where δm_x, δm_y, and δm_z are the three components of δm in the x, y, and z direction. Therefore

$$\hat{\mathbf{M}}\times\hat{\mathbf{m}} = \sin\theta\hat{\mathbf{x}} + \delta m_x\cos\theta\hat{\mathbf{y}} - \delta m_x\sin\theta\hat{\mathbf{z}} + O(e^{i\omega t}). \tag{A10}$$

Since $\tau_{b\perp}$ is dominated along the x direction, we have

$$\boldsymbol{\tau}_{b\perp} \approx \tau_{b\perp}\hat{\mathbf{x}} + \tau_{b\perp}\cdot\cot\theta\cdot\delta m_x\cdot\hat{\mathbf{y}}, \tag{A11}$$

where $\delta m_i \ll m_i$ ($i = x, y$ or z) and $\cot\theta\cdot\delta m_x \ll 1$. Similarly, we have

$$\hat{\mathbf{m}}\times(\hat{\mathbf{M}}\times\hat{\mathbf{m}}) = \sin\theta\hat{\mathbf{y}} - \delta m_x\cos\theta\hat{\mathbf{x}} + \delta m_x\sin\theta\hat{\mathbf{z}} + O(e^{i\omega t}). \tag{A12}$$

To a good approximation

$$\boldsymbol{\tau}_{b\parallel} \approx \tau_{b\parallel}\hat{\mathbf{y}} - \tau_{b\parallel}\cdot\cot\theta\cdot\delta m_x\cdot\hat{\mathbf{x}}. \tag{A13}$$

Therefore, Eq. (A9) turns out to be

$$\begin{aligned}\frac{d\hat{\mathbf{m}}}{dt} &= -\gamma\hat{\mathbf{m}}\times\mathbf{H} + \alpha\hat{\mathbf{m}}\times\frac{d\hat{\mathbf{m}}}{dt}\\ &- \gamma\frac{[\tau_{b\parallel}(I,\theta) - \tau_{b\perp}(I,\theta)\cot\theta\delta m_x]}{M_s}\hat{\mathbf{y}}\\ &- \gamma\frac{[\tau_{b\perp}(I,\theta) + \tau_{b\parallel}(I,\theta)\cot\theta\delta m_x]}{M_s}\hat{\mathbf{x}}.\end{aligned} \tag{A14}$$

The microwave field is $\hat{\mathbf{h}} = (h_x\hat{\mathbf{x}} + h_y\hat{\mathbf{y}} + h_z\hat{\mathbf{z}})e^{i\omega t}$. Considering the demagnetization effect, the total magnetic field $\hat{\mathbf{H}}$ can be represented as

$$\begin{aligned}\mathbf{H} &= \left[h_x e^{i\omega t} - N_x^0 M_s m_x(t)\right]\hat{\mathbf{x}} + \left[h_y e^{i\omega t} - N_y^0 M_s m_y(t)\right]\hat{\mathbf{y}}\\ &+ (h_z e^{i\omega t} + H\hat{\mathbf{z}}),\end{aligned} \tag{A15}$$

where N_x^0 and N_y^0 are the demagnetization factor in the x and y direction, respectively. Then, each term in Eq. (A14) can be written separately as

$$-\gamma\hat{\mathbf{m}}\times\hat{\mathbf{H}} \approx \gamma\left[\left(N_x^0 M_s + H\right)m_x\hat{\mathbf{y}} - \left(N_y^0 M_s + H\right)m_y\hat{\mathbf{x}}\right]e^{i\omega t},$$

$$\alpha\hat{\mathbf{m}}\times\frac{d\hat{\mathbf{m}}}{dt} \approx i\alpha\omega(m_x\hat{\mathbf{y}} - m_y\hat{\mathbf{x}})e^{i\omega t},$$

$$\begin{aligned}\tau_b(I,\theta) &= \tau_b^0 + \frac{d\tau_b}{dI}\delta I + \frac{d\tau_b}{d\theta}\delta\theta\\ &= \tau_b^0 + \frac{d\tau_b}{dI}I_{\mathrm{RF}}e^{i\omega t} - \frac{d\tau_b}{d\theta}m_y e^{i\omega t},\end{aligned}$$

where I_{RF} is the magnitude of the microwave current. Combining individual terms and ignoring constant terms gives us

$$i\omega m_x = -m_y\gamma\left(N_y^0 M_s + H\right) - im_y\alpha\omega - \frac{\gamma}{M_s}\left(\frac{d\tau_{b\perp}}{dI}I_{\mathrm{RF}} - \frac{d\tau_{b\parallel}}{dI}\cot\theta\cdot\delta m_x I_{\mathrm{RF}} - \frac{d\tau_b}{d\theta}m_y\right), \tag{A16}$$

$$i\omega m_y = m_x\gamma\left(N_x^0 M_s + H\right) - im_x\alpha\omega - \frac{\gamma}{M_s}\left(\frac{d\tau_{b\parallel}}{dI}I_{\mathrm{RF}} + \frac{d\tau_{b\perp}}{dI}\cot\theta\cdot\delta m_x I_{\mathrm{RF}} - \frac{d\tau_b}{d\theta}m_y\right), \tag{A17}$$

which leads to

$$\begin{pmatrix} i\omega & \left(\gamma N_y^0 M_s + \gamma H + i\alpha\omega\right) - \frac{\gamma}{M_s}\frac{d\tau_{b\perp}}{d\theta} \\ -\left(\gamma N_x^0 M_s + \gamma H + i\alpha\omega\right) & i\omega - \frac{\gamma}{M_s}\frac{d\tau_{b\parallel}}{d\theta} \end{pmatrix}\begin{pmatrix} m_x \\ m_y \end{pmatrix} = \begin{pmatrix} -\frac{\gamma}{M_s}I_{\mathrm{RF}}\left(\frac{d\tau_{b\perp}}{dI} - \frac{d\tau_{b\parallel}}{dI}\cot\theta\cdot\delta m_x\right) \\ -\frac{\gamma}{M_s}I_{\mathrm{RF}}\left(\frac{d\tau_{b\perp}}{dI} + \frac{d\tau_{b\parallel}}{dI}\cot\theta\cdot\delta m_x\right) \end{pmatrix}. \tag{A18}$$

Let $A = -(\gamma N_x^0 M_s + \gamma H + i\alpha\omega)$, $B = (\gamma N_y^0 M_s + \gamma H + i\alpha\omega) - \frac{\gamma}{M_s}\frac{d\tau_{b\perp}}{d\theta}$, and $C = i\omega - \frac{\gamma}{M_s}\frac{d\tau_{b\parallel}}{d\theta}$; then we have

$$\begin{pmatrix} m_x \\ m_y \end{pmatrix} = -\frac{\gamma I_{\mathrm{RF}}}{M_s}\frac{1}{i\omega C - AB}\begin{pmatrix} C & -B \\ -A & i\omega \end{pmatrix}\begin{pmatrix} -\frac{d\tau_{b\perp}}{dI} - \frac{d\tau_{b\parallel}}{dI}\cot\theta\cdot\delta m_x \\ \frac{d\tau_{b\perp}}{dI} + \frac{d\tau_{b\parallel}}{dI}\cot\theta\cdot\delta m_x \end{pmatrix}. \tag{A19}$$

(1) When there is no STT and no damping, we have

$$\begin{pmatrix} i\omega & \left(\gamma N_y^0 M_s + \gamma H\right) \\ -\left(\gamma N_x^0 M_s + \gamma H\right) & i\omega \end{pmatrix}\begin{pmatrix} m_x \\ m_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A20}$$

In order to have a nonzero solution of m_x and m_y, the determinant of the equation coefficient matrix should be zero:

$$\begin{vmatrix} i\omega & \left(\gamma N_y^0 M_s + \gamma H\right) \\ -\left(\gamma N_x^0 M_s + \gamma H\right) & i\omega \end{vmatrix} = 0. \tag{A21}$$

This gives us the relation between resonant frequency ω and resonant field H for FMR in this situation:

$$\omega^2 = \left(\gamma N_x^0 M_s + \gamma H\right)\left(\gamma N_y^0 M_s + \gamma H\right) \approx \gamma^2 H_r(H_r + M_0) \tag{A22}$$

with $M_0 \equiv (N_x^0 + N_y^0)M_s/\mathrm{Vol} \approx N_x^0 M_s/\mathrm{Vol}$ and $H_r = \frac{1}{2}(-M_0 + \sqrt{M_0^2 + 4\omega^2/\gamma^2})$ the resonant position of applied magnetic field when there is no STT and no damping [43].

(2) Assuming $N_y^0 \ll N_x^0$ (for a flat disk in the y-z plane, $N_x^0 = 4\pi$ and $N_y^0 = N_z^0 = 0$) and $\alpha \ll 1$, the denominator in Eq. (A19), $i\omega C - AB$, can be written as

$$i\omega C - AB = 2\gamma^2 H_r(H - H_r' + i\Delta H), \tag{A23}$$

where

$$H_r' = H_r + \frac{1}{2M_s}\frac{d\tau_{b\perp}}{d\theta}, \tag{A24}$$

and

$$\Delta H = \sqrt{1 + M_0/H_r}\left[\alpha(H_r + M_0/2) - \frac{1}{2M_s}\frac{d\tau_{b\parallel}}{d\theta} - \frac{\alpha}{2M_s}\frac{d\tau_{b\perp}}{d\theta}\right]. \tag{A25}$$

Then, from Eq. (A19), we can get

$$m_y = -\frac{I_{\mathrm{RF}}}{2M_s}\frac{\sqrt{1 + M_0/H_r}}{(H - H_r' + i\Delta H)}\left[\sqrt{1 + M_0/H_r}\left(\frac{d\tau_{b\perp}}{dI} - \frac{d\tau_{b\parallel}}{dI}\cot\theta\cdot\delta m_x\right) + i\left(\frac{d\tau_{b\parallel}}{dI} + \frac{d\tau_{b\perp}}{dI}\cot\theta\cdot\delta m_x\right)\right]. \tag{A26}$$

Therefore, the voltage V_r can be expressed as

$$\begin{aligned} V_r = \frac{1}{4}\frac{\partial^2 V}{\partial I^2}I_{\mathrm{RF}}^2 + \frac{I_{\mathrm{RF}}^2}{4M_s}\frac{\partial^2 V}{\partial\theta\partial I}\sqrt{1 + \frac{M_0}{H_r}}\frac{1}{\Delta H}\cdot\left[D(H)\sqrt{1 + \frac{M_0}{H_r}}\left(\frac{d\tau_{b\perp}}{dI} - \frac{d\tau_{b\parallel}}{dI}\cot\theta\cdot\delta m_x\right)\right. \\ \left. - L(H)\left(\frac{d\tau_{b\parallel}}{dI} + \frac{d\tau_{b\perp}}{dI}\cot\theta\cdot\delta m_x\right)\right]. \end{aligned} \tag{A27}$$

When there is only TSTT without dc biased STT, the amplitudes of dispersive and Lorentz components D and L are

$$D = \frac{I_{\mathrm{RF}}^2}{4M_s}\frac{\partial^2 V}{\partial\theta\partial I}\left[1+\frac{M_0}{H_r}\right]\frac{1}{\Delta H}\left(\frac{d\tau_{b\perp}}{dI}-\frac{d\tau_{b\parallel}}{dI}\cot\theta\cdot\delta m_x\right) \tag{A28}$$

and

$$L = -\frac{I_{\mathrm{RF}}^2}{4M_s}\frac{\partial^2 V}{\partial\theta\partial I}\sqrt{1+\frac{M_0}{H_r}}\frac{1}{\Delta H}\frac{d\tau_{b\parallel}}{dI}. \tag{A29}$$

Our measurements were performed under open circuits conditions, i.e., no dc charge currents, therefore $\frac{d\tau_{b\perp}}{dI}$ is very small considering the symmetry of $\tau_{b\perp}$ at small voltage bias [38]. The ratio D/L can be written as

$$D/L \approx \sqrt{1+\frac{M_0}{H_r}}\cot\theta\cdot\delta m_x. \tag{A30}$$

As we discussed in the main text, δm_x is a small disturbance on $\hat{\mathbf{m}}$ under a temperature difference ΔT across the sample and can be expressed as [31]

$$\begin{aligned}\delta m_x &= \sum_{\beta=\perp,\parallel}\frac{dm_x}{d\tau_{t,\beta}}\frac{d\tau_{t,\beta}}{dT}\Delta T \\ &= \frac{1}{M_s}\sum_{\beta=\perp,\parallel}\chi_{x\beta}\frac{\partial\tau_{t,\beta}}{\partial T}\Delta T\end{aligned} \tag{A31}$$

where $\chi_{\alpha\beta}$ is the magnetic susceptibility tensor of the free layer. Finally, D/L can be written as

$$D/L = \frac{1}{M_s}\sqrt{1+\frac{M_0}{H_r}}\cdot\cot\theta\cdot\sum_{\beta=\perp,\parallel}\chi_{x\beta}\frac{\partial\tau_{t,\beta}}{\partial T}\Delta T. \tag{A32}$$

For the dc biased case, from Eq. (A27), the ratio D/L is

$$D/L = \sqrt{1+\frac{M_0}{H_r}}\frac{(d\tau_{b\perp}/dI)|_{I_0=V/R}}{(d\tau_{b\parallel}/dI)_{I_0=V/R}}, \tag{A33}$$

where V is the voltage bias and R is the resistance of the MTJ.

[1] D. Ralph and M. Stiles, J. Magn. Magn. Mater. **320**, 1190 (2008).
[2] L. Berger, J. Appl. Phys. **49**, 2156 (1978).
[3] J. Slonczewski, J. Magn. Magn. Mater. **159**, L1 (1996).
[4] L. Berger, Phys. Rev. B **54**, 9353 (1996).
[5] E. B. Myers, D. C. Ralph, J. A. Katine, R. N. Louie, and R. A. Buhrman, Science **285**, 867 (1999).
[6] A. Brataas, A. D. Kent, and H. Ohno, Nat. Mater. **11**, 372 (2012).
[7] J. A. Katine, F. J. Albert, R. A. Buhrman, E. B. Myers, and D. C. Ralph, Phys. Rev. Lett. **84**, 3149 (2000).
[8] S. I. Kiselev, J. C. Sankey, I. N. Krivorotov, N. C. Emley, R. J. Schoelkopf, R. A. Buhrman, and D. C. Ralph, Nature (London) **425**, 380 (2003).
[9] S. Yuasa, A. Fukushima, T. Nagahama, K. Ando, and Y. Suzuki, Jpn. J. Appl. Phys. **43**, L588 (2004).
[10] S. Yuasa, T. Nagahama, A. Fukushima, Y. Suzuki, and K. Ando, Nat. Mater. **3**, 868 (2004).
[11] S. S. P. Parkin, C. Kaiser, A. Panchula, P. M. Rice, B. Hughes, M. Samant, and S.-H. Yang, Nat. Mater. **3**, 862 (2004).
[12] D. D. Djayaprawira, K. Tsunekawa, M. Nagai, H. Maehara, S. Yamagata, N. Watanabe, S. Yuasa, Y. Suzuki, and K. Ando, Appl. Phys. Lett. **86**, 092502 (2005).
[13] S. Yuasa and D. D. Djayaprawira, J. Phys. D **40**, R337 (2007).
[14] 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, in *IEEE International Electron Devices Meeting, 2005, IEDM Technical Digest* (IEEE, Piscataway, NJ, 2005), p. 459.
[15] S. Ikeda, J. Hayakawa, Y. M. Lee, F. Matsukura, Y. Ohno, T. Hanyu, and H. Ohno, IEEE Trans. Electron Devices **54**, 991 (2007).
[16] G. D. Fuchs, J. A. Katine, S. I. Kiselev, D. Mauri, K. S. Wooley, D. C. Ralph, and R. A. Buhrman, Phys. Rev. Lett. **96**, 186603 (2006).
[17] S.-C. Oh, S.-Y. Park, A. Manchon, M. Chshiev, J.-H. Han, H.-W. Lee, J.-E. Lee, K.-T. Nam, Y. Jo, Y.-C. Kong, B. Dieny, and K.-J. Lee, Nature Physics **5**, 898 (2009).
[18] Y. Huai, F. Albert, P. Nguyen, M. Pakala, and T. Valet, Appl. Phys. Lett. **84**, 3118 (2004).
[19] G. D. Fuchs, N. C. Emley, I. N. Krivorotov, P. M. Braganca, E. M. Ryan, S. I. Kiselev, J. C. Sankey, D. C. Ralph, R. A. Buhrman, and J. A. Katine, Appl. Phys. Lett. **85**, 1205 (2004).
[20] V. S. Pribiag, I. N. Krivorotov, G. D. Fuchs, P. M. Braganca, O. Ozatay, J. C. Sankey, D. C. Ralph, and R. A. Buhrman, Nature Physics **3**, 498 (2007).
[21] K. Ando, S. Fujita, J. Ito, S. Yuasa, Y. Suzuki, Y. Nakatani, T. Miyazaki, and H. Yoda, J. Appl. Phys. **115**, 172607 (2014).
[22] M. Hatami, G. E. W. Bauer, Q. Zhang, and P. J. Kelly, Phys. Rev. Lett. **99**, 066603 (2007).
[23] M. Hatami, G. E. W. Bauer, Q. Zhang, and P. J. Kelly, Phys. Rev. B **79**, 174426 (2009).
[24] G. E. W. Bauer, S. Bretzel, A. Brataas, and Y. Tserkovnyak, Phys. Rev. B **81**, 024427 (2010).
[25] G. E. W. Bauer, E. Saitoh, and B. J. van Wees, Nat. Mater. **11**, 391 (2012).
[26] C. Heiliger, C. Franz, and M. Czerner, J. Appl. Phys. **115**, 172614 (2014).
[27] X. Jia, K. Xia, and G. E. W. Bauer, Phys. Rev. Lett. **107**, 176603 (2011).
[28] X.-T. Jia and K. Xia, Frontiers of Physics **9**, 768 (2014).
[29] J. C. Leutenantsmeyer, M. Walter, V. Zbarsky, M. Münzenberg, R. Gareev, K. Rott, A. Thomas, G. Reiss, P. Peretzki, H. Schuhmann *et al.*, Spin **03**, 1350002 (2013).
[30] H. Yu, S. Granville, D. P. Yu, and J.-P. Ansermet, Phys. Rev. Lett. **104**, 146601 (2010).
[31] L. Fitoussi, F. A. Vetro, C. Caspers, L. Gravier, H. Yu, and J.-P. Ansermet, Appl. Phys. Lett. **106**, 162401 (2015).

[32] A. Pushp, T. Phung, C. Rettner, B. P. Hughes, S.-H. Yang, and S. S. P. Parkin, Proc. Natl. Acad. Sci. USA **112**, 6585 (2015).
[33] Z. H. Zhang, Y. S. Gui, L. Fu, X. L. Fan, J. W. Cao, D. S. Xue, P. P. Freitas, D. Houssameddine, S. Hemour, K. Wu, and C.-M. Hu, Phys. Rev. Lett. **109**, 037206 (2012).
[34] M. Walter, J. Walowski, V. Zbarsky, M. Münzenberg, M. Schäfers, D. Ebke, G. Reiss, A. Thomas, P. Peretzki, M. Seibt *et al.*, Nat. Mater. **10**, 742 (2011).
[35] M. V. Costache, S. M. Watts, M. Sladkov, C. H. Van Der Wal, and B. J. Van Wees, Appl. Phys. Lett. **89**, 232115 (2006).
[36] Y. S. Gui, N. Mecking, X. Zhou, G. Williams, and C.-M. Hu, Phys. Rev. Lett. **98**, 107602 (2007).
[37] Y. Gui, L. Bai, and C. Hu, Science China: Physics, Mechanics, and Astronomy **56**, 124 (2013).
[38] J. C. Sankey, Y.-T. Cui, J. Z. Sun, J. C. Slonczewski, R. A. Buhrman, and D. C. Ralph, Nature Physics **4**, 67 (2008).
[39] C. Wang, Y.-T. Cui, J. Z. Sun, J. A. Katine, R. A. Buhrman, and D. C. Ralph, Phys. Rev. B **79**, 224416 (2009).
[40] M. Julliere, Phys. Lett. A **54**, 225 (1975).
[41] I. Theodonis, N. Kioussis, A. Kalitsov, M. Chshiev, and W. H. Butler, Phys. Rev. Lett. **97**, 237205 (2006).
[42] A. A. Kovalev, G. E. W. Bauer, and A. Brataas, Phys. Rev. B **75**, 014430 (2007).
[43] N. Mecking, Y. S. Gui, and C. M. Hu, Phys. Rev. B **76**, 224430 (2007).