Thermal spin torques in magnetic insulators

H. Yu,[1,2,*] S. D. Brechet,[2] P. Che,[1,2] F. A. Vetro,[2] M. Collet,[3] S. Tu,[1,2] Y. G. Zhang,[1] Y. Zhang,[1] T. Stueckler,[1] L. Wang,[1,4] H. Cui,[4] D. Wang,[4] C. Zhao,[4] P. Bortolotti,[3] A. Anane,[3] J-Ph. Ansermet,[2,†] and W. Zhao[1]

[1]*Fert Beijing Research Institute, School of Electrical and Information Engineering, BDBC, Beihang University, China*

[2]*Institute of Physics, Station 3, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne-EPFL, Switzerland*

[3]*Unité Mixte de Physique CNRS, Thales, Univ. Paris Sud, Université Paris-Saclay, 91767 Palaiseau, France*

[4]*Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100029, China*

(Received 15 December 2016; revised manuscript received 3 March 2017; published 23 March 2017)

The damping of spin waves transmitted through a two-port magnonic device implemented on a yttrium iron garnet thin film is shown to be proportional to the temperature gradient imposed on the device. The sign of the damping depends on the relative orientation of the magnetic field, the wave vector, and the temperature gradient. The observations are accounted for qualitatively and quantitatively by using an extension of the variational principle that leads to the Landau-Lifshitz equation. All parameters of the model can be obtained by independent measurements.

DOI: 10.1103/PhysRevB.95.104432

The discovery of giant magnetoresistance (GMR) revolutionized information storage technology [1,2] and the spin-transfer torque (STT), predicted two decades ago by Slonczewski [3] and Berger [4], may reshape once again the magnetic memory industry [5]. The concept of a heat-driven spin torque, or thermal spin-transfer torque (TST), has been suggested [6–8] and opened the world of spin caloritronics. Magnetic insulators are ideal for studying the fundamentals of spin caloritronics, because they are free of the effect of heat on charge transport. Here, we demonstrate that a spin torque can be induced in magnetic insulators by applying a thermal gradient. The effect is not linked to spin-dependent transport at interfaces since we observe a heat-driven contribution to damping of magnetization waves on a millimeter scale. We show that by adding to $\boldsymbol{M}(\boldsymbol{r})$ the bound magnetic current $(\boldsymbol{\nabla} \times \boldsymbol{M})$ as state variable, the variational principle that yields the Landau-Lifshitz equation predicts the presence of a thermal spin torque, from which we derive an expression for spin currents in insulators. Our experiments verify the key predictions of this model. Thermodynamics can predict a link between heat and magnetization, but cannot determine the strength of the effect [9].

Spin caloritronics studies the interplay of spin, charge, and heat transport [10]. As the spin dependence of the electrical conductivity proved to be important since it gives rise to GMR, the spin dependence of other transport parameters has been investigated, such as that of the Seebeck [11] and Peltier coefficients [12]. The combination of heat with spin and charge transport gained widespread attention owing to studies of the spin Seebeck effect [13,14]. The STT effect which uses a spin-polarized electrical current has shown promising applications, e.g., in magnetic memories (STT-MRAM). It was already established that heat flowing through a ferromagnetic metal can generate a diffusive spin current [15] which induces a spin torque when flowing through a magnetic nanostructure [6]. Experimentally, this effect was studied in Co/Cu/Co spin valve nanowires by observing the change in the switching field of magnetization due to a local thermal gradient [7]. It was later shown that heat couples to magnetization dynamics [16–18]. The effect of heat on magnetization was also found in magnetic tunnel junctions [19] and metallic spin valves [20]. Slonczewski predicted that a spin-transfer torque induced by thermal magnons could be more efficient than the usual electrically induced spin torques [8]. Combining TST and STT might further decrease the write-current magnitude of MRAMs [21].

A 20-nm-thick yittrium iron garnet (YIG) film was grown on gadolinium gallium garnet (GGG) substrate using pulsed laser deposition. Details of the growth condition and magnetic properties of the thin YIG layer can be found in Ref. [22].

Figure 1 shows the experimental principle of the measurement. Using inductively coupled plasma etching and photolithography, a YIG strip 100 μm wide and 4.8 mm long was prepared. The ends were designed with a 45° angle in order to avoid spin-wave reflection. Following the etching process, a 10-nm-thick copper or platinum bar was deposited on top of the YIG strip by electron-beam evaporation. This bar is connected to two large Au electrodes. These electrodes are designed for contact with a ground-signal-ground microprobe. The magnetic field is applied along the YIG strip, and spin waves are excited by one microprobe and detected by another. Alternatively, a microcoil [23] was used for excitation. Excitation and detection are 800 μm apart. The results were obtained with contacts made of Pt with a Ta seed layer. The resonance frequency could be tuned from 4 GHz up to 10 GHz. Lock-in detection with field modulation was used. The thermal gradient was generated by two Peltier elements and defined as $\nabla T = (T_B - T_A)/l$ with $l = 5$ mm being the distance between the Peltier elements. Using an infrared camera, we verified that the temperature changed linearly at the location of the sample.

As shown in Fig. 2, the linewidth changes linearly with temperature gradient. Furthermore, the slope changes sign when the field is reversed or when the propagation direction is reversed. For the latter case, we had to move the sample and this caused a change in the linewidth of 0.03 mT when the sample was at a uniform temperature. In Fig. 2, we translated all data points by this amount when the sample was flipped.

*haiming.yu@buaa.edu.cn

†jean-philippe.ansermet@epfl.ch

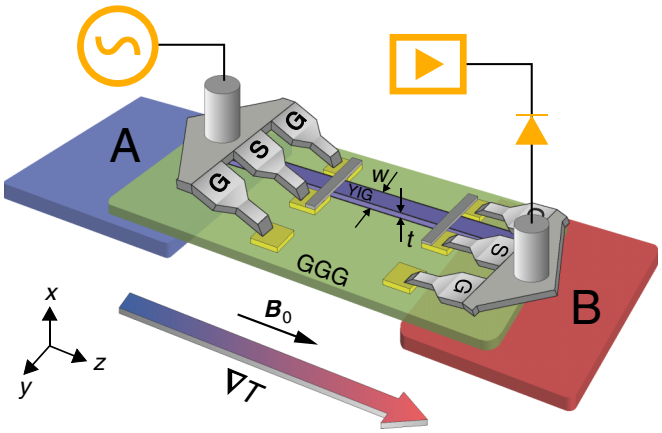

FIG. 1. Spin-wave propagation under a thermal gradient. 4.8-mm-long YIG strip fabricated on GGG substrate, width $w = 100\ \mu$m, thickness $t = 20$ nm, 10-nm-thick Cu contact connected to Au electrodes, microprobes for both excitation and detection, Peltier elements A and B heat sunk by copper blocks (not shown).

We can account for the observed effect of a temperature gradient on spin-wave transmission by a model based on an extension of the variation principle which yields the well-known Landau-Lifshitz-Gilbert (LLG) equation [24]. In the presence of an applied thermal gradient ∇T, the LLG equation

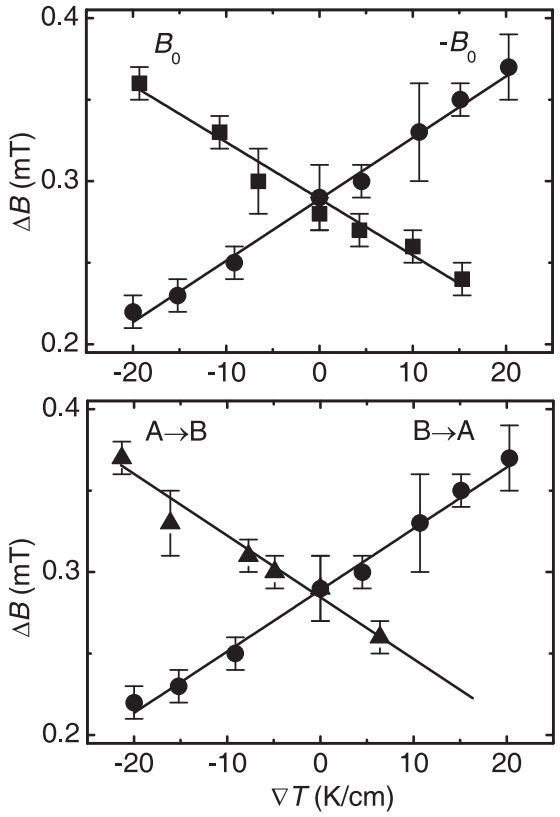

FIG. 2. Linewidth of the ferromagnetic resonance spectra at 4.2 GHz, as a function of temperature gradient. The slope changes sign upon flipping the field (top) or flipping the direction of propagation at fixed field orientation (bottom). $A \rightarrow B$ data are translated by 0.03 mT.

for the time evolution of the magnetization $\boldsymbol{M}$ contains a thermal spin torque term, i.e.,

$$\dot{\boldsymbol{M}} = \gamma\, \boldsymbol{M} \times \boldsymbol{B}_{\text{eff}} + \frac{\alpha}{M_S}\, \boldsymbol{M} \times \dot{\boldsymbol{M}} + \gamma \boldsymbol{\tau}_{\text{TST}}, \tag{1}$$

where $\gamma < 0$ is the gyromagnetic ratio, α is the magnetic damping parameter, and M_S is the saturation magnetization. The effective magnetic field $\boldsymbol{B}_{\text{eff}}$ is composed of the external field $\boldsymbol{B}_0$, the demagnetizing field $\boldsymbol{B}_{\text{dem}}$, the anisotropy field $\boldsymbol{B}_{\text{ani}}$, and the microwave excitation field $\boldsymbol{b}$ induced by the microwave antenna. The torque $\boldsymbol{\tau}_{\text{TST}}$ can be expressed as

$$\boldsymbol{\tau}_{\text{TST}} = \alpha_{\text{TST}} \frac{\omega}{|\gamma|} \frac{\boldsymbol{M}}{M_s^2} \times (\boldsymbol{M} \times \boldsymbol{m}_k), \tag{2}$$

where the effective thermal spin torque damping coefficient α_{TST} can be written as

$$\alpha_{\text{TST}} = -\frac{\omega_M}{\omega} \frac{k_T}{k}. \tag{3}$$

Here, ω corresponds to the microwave frequency and $\boldsymbol{m}_k$ is the out-of-equilibrium component of the magnetization for a mode of wave number k. In this work, we provide a quantitative expression for the thermal wave vector k_T with no adjustable parameter:

$$\boldsymbol{k}_T = \frac{\omega - \omega_0}{\omega_M} \left| \frac{1}{M_S} \frac{dM_S}{dT} \right| \nabla T, \tag{4}$$

where $\omega_0 = -\gamma B_0$ and $\omega_M = -\gamma M_S$. The lengthy derivations of the above equations are given in the Supplemental Material [25]. The effective damping parameter α_{eff} is the sum of the Gilbert damping parameter α and the thermal spin torque damping parameter α_{TST}. The observed spin-wave spectral linewidth is therefore given by [25]

$$\Delta B = \Delta B_0 + \frac{2}{\sqrt{3}} \alpha \left| \frac{\omega_K}{\gamma} \right| - \frac{2}{\sqrt{3}} \left| \frac{\omega_K - \omega_0}{\gamma} \right| \left| \frac{1}{M_S} \frac{dM_S}{dT} \right| \frac{1}{k} \nabla T, \tag{5}$$

where ω_K is the resonance frequency, given by the Kittel formula [26] and the first two terms are the usual ones [27].

Thus, our model predicts that the thermal spin torque changes sign under reversal of either the temperature gradient, the propagation direction, or the applied magnetic field (Fig. 2). Initially, we varied the applied thermal gradient and observed a linear change in the spin-wave spectral linewidth for one orientation of the field. This linear dependence is consistent with Eq. (5). Clearly, when the thermal gradient changes sign, the linewidth changes from a broadening to a narrowing with respect to its value in the isothermal condition. It must be noted that the temperature has hardly any influence on the linewidth [25]. The dependence of linewidth with thermal gradient changes sign when the magnetic field is reversed (Fig. 2, top). This can be understood as follows. If ω changes sign because B is reversed, then k must change sign also if we want propagation to be maintained in the same orientation [25]. Therefore, according to Eq. (5), the slope of the linewidth plotted vs temperature gradient must change sign when the magnetic field is reversed, as confirmed by Fig. 2 (top). Furthermore, if we swap the excitation and the detection, i.e., we reverse the spin-wave vector $\boldsymbol{k}$, then we observe that the thermal spin torque effect is also reversed, as shown in

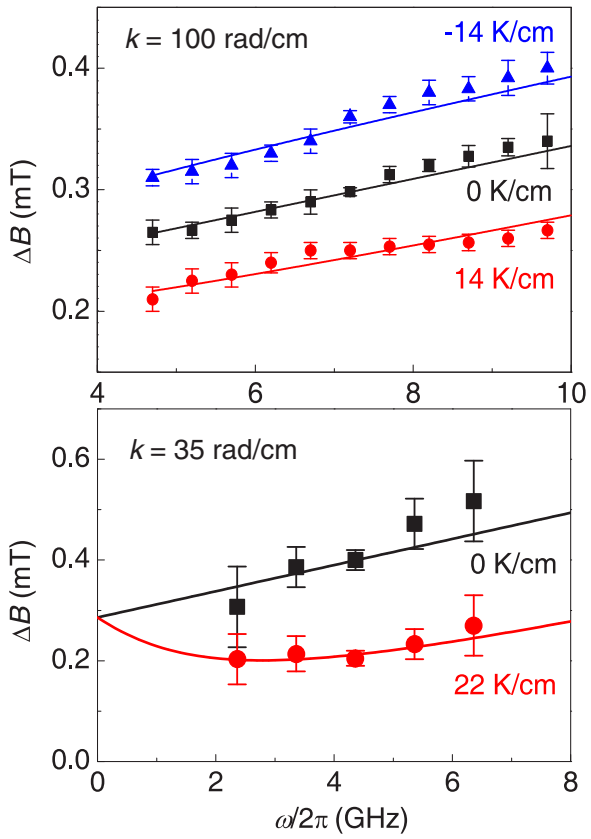

FIG. 3. Linewidth as a function of frequency at a set temperature gradient, using microprobe (top), or metal contacts (bottom) for excitation. Wave vector based on HFSS calculation. The applied temperature gradients are indicated in the figure. Top: black line yields $\alpha = 3.15 \times 10^{-4}$; red and blue lines using Eq. (5). Bottom: black line yields $\alpha = 6.30 \times 10^{-4}$; red line using Eq. (5). The error bars indicate the noise level.

Fig. 2 (bottom), which is consistent with the linewidth being proportional to $1/k$ [Eq. (5)].

We now investigate the frequency dependence of linewidth variation. The upper part of Fig. 3 shows the linewidth changes with frequencies from 4.7 GHz up to 9.7 GHz using a microprobe for excitation. We ran a high frequency electromagnetic field simulation (HFSS) taking into account the dimensions of the microprobe and acquired the field distribution at the injection area. We then used Fourier transformation to obtain the k space distribution [25]. Thus, we found that the most prominent excitation has a wave vector around 100 rad/cm, and that there are some higher order modes with much lower intensities. The lower part of Fig. 3 shows the frequency dependence of linewidth measured using the microcoil for excitation. According to the results from HFSS, we found that the dominant wave vector k of excitation is much smaller, namely, 35 rad/cm. The slope of the frequency dependence is proportional to the effective damping parameter. We can observe that the change of the slope is more significant for microcoil excitation than that for microprobe excitation. This can be understood from Eq. (3) where the thermal spin torque induced damping parameter is inversely proportional to the spin-wave wave vector. We can account for the data using the k values deduced from the HFSS calculation. We take the temperature dependence of the saturation magnetization to be $|\frac{1}{M_S}\frac{dM_S}{dT}| = 3.8 \times 10^{-3}$ K^{-1} based on Ref. [16] and confirmed by isothermal measurements of saturation magnetization [25]. In the lower part of Fig. 3, we fit the data based on Eq. (5), using the damping parameter $\alpha = 6.30 \times 10^{-4}$ deduced from the data taken without any thermal gradient. This smaller value could be due to the fact that when using the microcoil excitation, the detection was done using a Pt bar, whereas a Cu bar was used when taking data with the microprobe excitation. According to Ref. [18], the growth of Pt on YIG may introduce an increase of damping. In summary, the various data presented in Fig. 3 can be accounted for quantitatively with parameters that are all determined by independent measurements.

Finally, we note that the thermal spin torque [Eqs. (2) and (3)] can be expressed in terms of a spin current. To first order in the linear response, the thermal spin torque is given by [25]

$$\boldsymbol{\tau}_{\mathrm{TST}} = \boldsymbol{k}_{\mathrm{T}} \cdot \mathrm{j}_{\mathrm{s}}, \tag{6}$$

where the dot stands for the tensor contraction and the thermal spin current tensor j_{s} is defined by

$$\mathrm{j}_{\mathrm{s}} = -\mu_0 \boldsymbol{M}_{\mathrm{S}} \times \boldsymbol{\nabla}^{-1} \boldsymbol{m}_{\mathrm{k}}. \tag{7}$$

The spin current density tensor j_{s} has physical dimensions (J/m^2 in SI units) that correspond to the product of a spin density and a phase velocity. Expression (7) has the same geometry to first order as the spin-wave spin current tensor derived by Saitoh and Ando [28]. However, the physical origin of this spin current tensor is different since here, it is obtained specifically for the case of a spin current induced by a thermal gradient.

Very recently, self-oscillation based on spin-orbit torque was found in YIG/Pt pillars [29] and in permalloy/Pt nanowires [30]. By analogy, we may expect self-oscillation driven by a thermal spin torque as well.

In conclusion, we have prepared thin-film YIG microstrips and found that the linewidth of transmission spectra can be broadened or narrowed by applying a thermal gradient. These observations are accounted for by an effective damping that is due to a thermal spin torque. A comprehensive theoretical analysis provides an explicit expression for this torque, which is derived from an extension of the variational principle on which the Landau-Lifshitz equation is based. This study points to the possibility of damping control in magnonic devices using a local thermal gradient.

We wish to acknowledge the support by NSF China under Grants No. 11674020 and No. 11444005, for S.T. by the Sino-Swiss Science and Technology Cooperation SSSTC Grant No. EG 01-032015, for F.A.V and P.C. by the Polish-Swiss Research Program NANOSPIN PSRP-045/2010, for H.Y. by the Deutsche Forschungsgemeinschaft SPP 1538 (SpinCat) Grant No. AN762/1, and by the International Collaboration 111 Project B16001 from the Ministries of Education and Foreign Experts. The authors thank Vincent Cros for comments on the manuscript.

[1] M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas, Phys. Rev. Lett. **61**, 2472 (1988).

[2] G. Binasch, P. Grünberg, F. Saurenbach, and W. Zinn, Phys. Rev. B **39**, 4828(R) (1989).

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

[4] L. Berger, Phys. Rev. B **54**, 9353 (1996).

[5] A. D. Kent and D. C. Worledge, Nat. Nanotechnol. **10**, 187 (2015).

[6] M. Hatami, G. E. W. Bauer, Q. Zhang, and P. J. Kelly, Phys. Rev. Lett. **99**, 066603 (2007).

[7] H. Yu, S. Granville, D. P. Yu, and J.-P. Ansermet, Phys. Rev. Lett. **104**, 146601 (2010).

[8] J. C. Slonczewski, Phys. Rev. B **82**, 054403 (2010).

[9] S. D. Brechet, F. A. Vetro, E. Papa, S. E. Barnes, and J.-P. Ansermet, Phys. Rev. Lett. **111**, 087205 (2013).

[10] G. E. Bauer, E. Saitoh, and B. J. van Wees, Nat. Mater. **11**, 391 (2012).

[11] L. Piraux, A. Fert, P. Schroeder, R. Loloee, and P. Etienne, J. Magn. Magn. Mater. **110**, L247 (1992).

[12] J. Flipse, F. Bakker, A. Slachter, F. Dejene, and B. J. Van Wees, Nat. Nanotechnol. **7**, 166 (2012).

[13] K. Uchida, S. Takahashi, K. Harii, J. Ieda, W. Koshibae, K. Ando, S. Maekawa, and E. Saitoh, Nature (London) **455**, 778 (2008).

[14] C. Jaworski, R. Myers, E. Johnston-Halperin, and J. Heremans, Nature (London) **487**, 210 (2012).

[15] A. Slachter, F. L. Bakker, J.-P. Adam, and B. J. van Wees, Nat. Phys. **6**, 879 (2010).

[16] B. Obry, V. I. Vasyuchka, A. V. Chumak, A. A. Serga, and B. Hillebrands, Appl. Phys. Lett. **101**, 192406 (2012).

[17] G. da Silva, L. Vilela-Leao, S. Rezende, and A. Azevedo, Appl. Phys. Lett. **102**, 012401 (2013).

[18] L. Lu, Y. Sun, M. Jantz, and M. Wu, Phys. Rev. Lett. **108**, 257202 (2012).

[19] 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).

[20] G.-M. Choi, C.-H. Moon, B.-C. Min, K.-J. Lee, and D. G. Cahill, Nat. Phys. **11**, 576 (2015).

[21] N. Mojumder, D. Abraham, K. Roy, and D. Worledge, IEEE Trans. Magn. **48**, 2016 (2012).

[22] O. d'Allivy Kelly, A. Anane, R. Bernard, J. B. Youssef, C. Hahn, A. H. Molpeceres, C. Carrétéro, E. Jacquet, C. Deranlot, P. Bortolotti, R. Lebourgeois, J.-C. Mage, G. de Loubens, O. Klein, V. Cros, and A. Fert, Appl. Phys. Lett. **103**, 082408 (2013).

[23] E. Papa, S. Barnes, and J.-P. Ansermet, IEEE Trans. Magn. **49**, 1055 (2013).

[24] S. D. Brechet and J.-P. Ansermet, Europhys. Lett. **112**, 17006 (2015).

[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.95.104432, which includes Refs. [27,31], for additional details on theoretical derivation, sample fabrication, measurement technique, isothermal measurements, and HFSS simulations.

[26] C. Kittel, Phys. Rev. **73**, 155 (1948).

[27] A. G. Gurevich and G. A. Melkov, *Magnetization Oscillations and Waves* (CRC Press, Boca Raton, FL, 1996).

[28] D. C. Maekawa, S. O. Valenzuela, E. Saitoh, and T. Kimura, *Spin Current* (Oxford University Press, New York, 1996).

[29] M. Collet, X. de Milly, O. d'Allivy Kelly, V. V. Naletov, R. Bernard, P. Bortolotti, J. B. Youssef, V. E. Demidov, S. O. Demokritov, J. L. Prieto, M. Muñoz, V. Cros, A. Anane, G. de Loubens, and O. Klein, Nat. Commun. **7**, 10377 (2016).

[30] Z. Duan, A. Smith, L. Yang, B. Youngblood, J. Lindner, V. E. Demidov, S. O. Demokritov, and I. N. Krivorotov, Nat. Commun. **5**, 5616 (2014).

[31] D. C. Ralph and M. D. Stiles, J. Magn. Magn. Mater. **320**, 1190 (2008).

Supplementary Material:

Thermal spin torques in magnetic insulators

H. Yu,[1,2,*] S. D. Brechet,[2] P. Che,[1,2] F. A. Vetro,[2] M. Collet,[3] S. Tu,[1] Y. G. Zhang,[1] Y. Zhang,[1] T. Stueckler,[1] L. Wang,[1,4] H. Cui,[4] D. Wang,[4] C. Zhao,[4] P. Bortolotti,[3] A. Anane,[3] J-Ph. Ansermet,[2,†] and W. Zhao[1]

[1] *Fert Beijing Research Institute, School of Electronic and Information Engineering, Beihang University, Xueyuan Road 37, 100191 Beijing, China*

[2] *Institute of Physics, station 3, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne-EPFL, Switzerland*

[3] *Unité Mixte de Physique CNRS, Thales, Univ. Paris Sud, Université Paris-Saclay, 91767 Palaiseau, France*

[4] *Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100029, China*

(Dated: December 13, 2016)

* haiming.yu@buaa.edu.cn

† jean-philippe.ansermet@epfl.ch

Sample fabrication. A 20 nm-thick YIG film was grown on a gadolinium gallium garnet (GGG) substrate using pulsed laser deposition (PLD). Details of the growth condition and magnetic properties of the thin YIG layer can be found elsewhere [6]. Using inductively coupled plasma etching and photolithography, YIG strips of 100 μm wide and 4.8 mm long were prepared. The end was cut with a 45° angle in order to avoid spin-wave reflection. Afterwards, 10 nm-thick platinum (or copper) bars were deposited on top of the YIG strip by electron beam evaporation. The bars were connected to large Au electrodes which were designed for contact with a ground-signal-ground microprobe for microwave measurements.

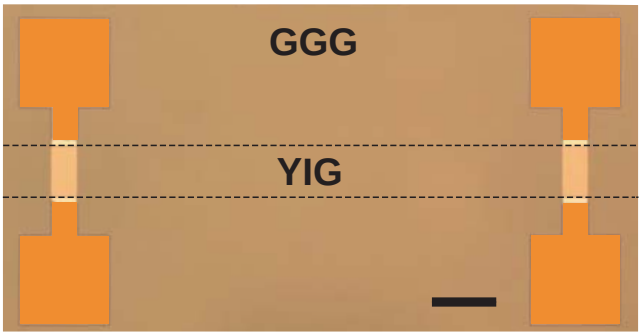

FIG. S1. Optical microscopy image of a typical device. The scale bar represents 100 μm. Dotted line: position of the transparent YIG strip.

Spin-wave transmission measurements. The magnetization waves were measured by field modulation and a Lock-in Amplifier. Magnetization waves were excited at gigahertz microwave frequencies and two picoprobes were used to inject microwave currents and to collect the microwave signals induced by propagating magnetization waves. The inductive microwave signals were rectified by a zero-bias Schottky diode and detected by a lock-in amplifier. The low frequency reference signal of the lock-in amplifier came from a function generator, which drove also two home-made modulation coils. The sample setup was placed in between the poles of an electromagnet held in place by a suction system. The YIG sample was glued on two Peltier elements on top of a copper disk. A gaussmeter probe was put at the center of the magnet, very close to the sample. Figure S2 shows the measurements of the transmission spectra between two microprobes. Using microprobes, one ensures local excitation and detection, with the propagation distance being 800 μm. Under a thermal gradient of 12 K/cm, either in positive or negative direction, the spectra linewidth are varied, as well as the amplitude.

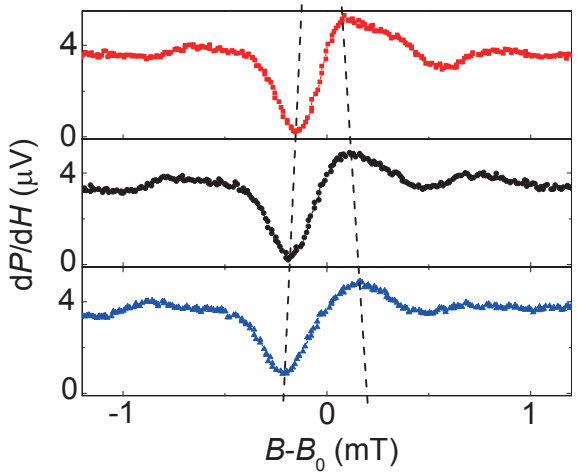

FIG. S2. Spin-wave spectra under different thermal gradients: using microprobe excitation, spectra obtained with $\nabla T = 0$ (black), 12 (red), -12 (blue) K/cm, centered at B_0, the resonance field for each case.

A secondary measurement technique is to excite spin waves in a contact-free manner. Fig. S3a shows different measured transmission spectra using a microcoil for excitation. The distance between excitation and detection is about 1800 μm. When a thermal gradient is applied, we observed that both the amplitude and the linewidth of the the spin-wave transmission spectra are dramatically modified, which exhibits consistent results with those obtained with the microprobe shown in the main text. In addition, the effect is strongly enhanced. A modest temperature gradient of 22 K/cm can enhance the amplitude of the measured spectrum by a factor of two.

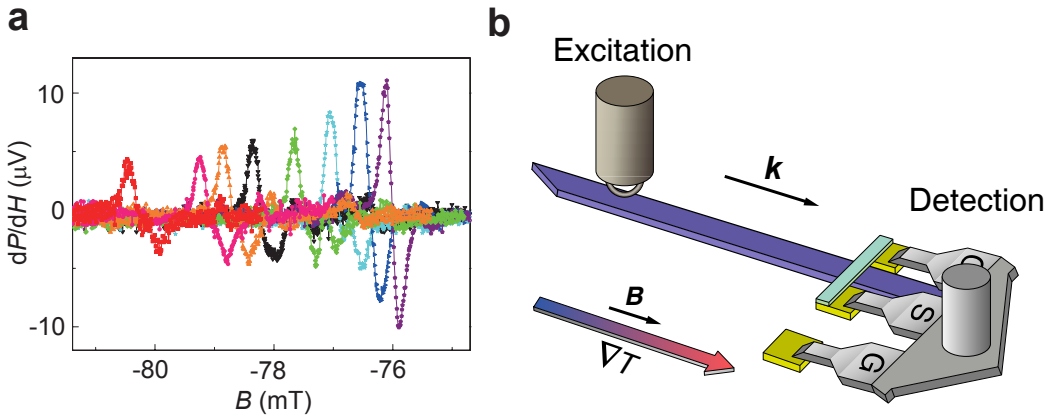

FIG. S3. **a**, Using microcoil excitation, from left to right, thermal gradients ∇T = +10 K/cm (red), +5 K/cm (pink), +2 K/cm (orange), 0 (black), -4 K/cm (green), -8 K/cm (cyan), -11 K/cm (blue), -22 K/cm (purple), respectively. **b**, Excitation using microcoil. Detection using microprobe connected to a 10 nm-thick Pt bar.

The generation and monitoring of thermal gradient. In-plane temperature gradients were created by two Peltier elements at the two ends of the sample. The upper sides of both Peltier elements could form a temperature gradient when they were charged by DC currents of opposite polarities. The sample was connected with the Peltier elements by thermal tape. The temperature at the lower sides was maintained by a copper disk which was well heat-sunk. In order to determine quantitatively the heat current flowing through the sample, two thermal sensors were installed on the upper surface of the Peltier elements near the sample. A thermal camera was used to confirm the existence of linear temperature gradient in the YIG strip.

Simulation using HFSS. We set up a model for the microcoil 1 mm above the YIG in the HFSS program. The simulation was set at 4.36 GHz. The amplitude and phase of the y component of the magnetic field H_y was extracted from the 1 mm sampling line along the YIG strip (x axis) by space steps of 0.01 mm. Using 100'000-point complex fast Fourier transform (FFT), the wavevector spectrum was then obtained, as shown in Fig. 4d. The simulation for microprobe is done in a similar way, taking into account of the detailed dimensions of the GSG picoprobe and the Cu or Pt bar.

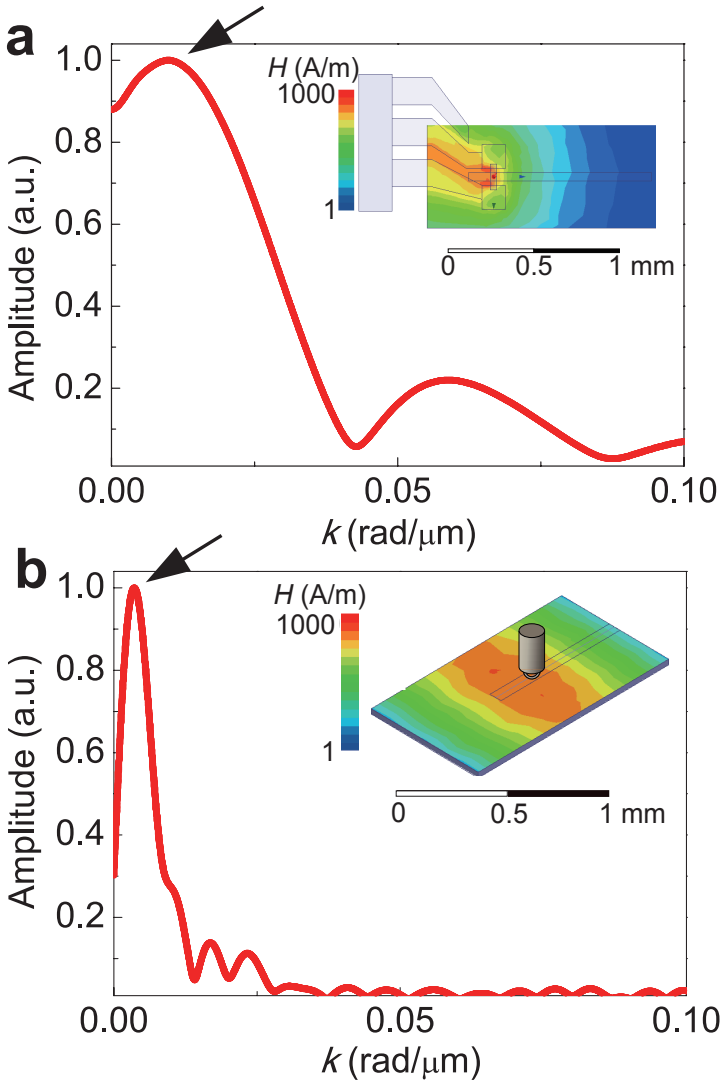

FIG. S4. **a**, Spin wave wavevector k distribution based on HFSS calculated for microprobe excitation. **b**, Spin wave wavevector k distribution based on the HFSS calculation for microcoil excitation.

I. THEORY OF THERMAL SPIN TORQUE

First, We determine the resonance frequency $\omega_{\rm K}$ which corresponds to the eigenfrequency of the magnetization dynamics in a stationnary state, in the absence of damping, i.e. $\alpha = 0$, and of a thermal gradient, i.e. $\boldsymbol{k}_{\rm T} = \boldsymbol{0}$. In a stationnary state, the time evolution of the excitation field $\boldsymbol{b}$ and of the magnetic response field $\boldsymbol{m}$ is given respectively by,

$$\dot{\boldsymbol{b}} = \boldsymbol{\omega}_{\rm K} \times \boldsymbol{b} \qquad \text{and} \qquad \dot{\boldsymbol{m}} = \boldsymbol{\omega}_{\rm K} \times \boldsymbol{m} \tag{S1}$$

where $\boldsymbol{\omega}_{\rm K} = \omega_{\rm K}\, \hat{\boldsymbol{z}}$ is the angular velocity at resonance. The time derivatives of the relations (S1) yield,

$$\ddot{\boldsymbol{b}} = -\,\omega_{\rm K}^2\, \boldsymbol{b} \qquad \text{and} \qquad \ddot{\boldsymbol{m}} = -\,\omega_{\rm K}^2\, \boldsymbol{m} \tag{S2}$$

In the absence of damping, i.e. $\alpha = 0$, and of a thermal gradient, i.e. $\boldsymbol{k}_{\rm T} = \boldsymbol{0}$, the linearised LLG equation reduces to,

$$\dot{\boldsymbol{m}} = \omega_0\, (\hat{\boldsymbol{z}} \times \boldsymbol{m}) - \frac{\omega_{\rm M}}{\mu_0}\, (\hat{\boldsymbol{z}} \times \boldsymbol{b}) + \omega_{\rm M}\, (\boldsymbol{m} \cdot \hat{\boldsymbol{x}})\, \hat{\boldsymbol{y}} \tag{S3}$$

The time derivative of relation (S3) is given by,

$$\ddot{\boldsymbol{m}} = \omega_0\, (\hat{\boldsymbol{z}} \times \dot{\boldsymbol{m}}) - \frac{\omega_{\rm M}}{\mu_0} \left(\hat{\boldsymbol{z}} \times \dot{\boldsymbol{b}}\right) + \omega_{\rm M}\, (\dot{\boldsymbol{m}} \cdot \hat{\boldsymbol{x}})\, \hat{\boldsymbol{y}} \tag{S4}$$

The substitution of (S3) into relation (S4) yields,

$$\ddot{\boldsymbol{m}} = -\,\omega_0\, (\omega_0 + \omega_{\rm M})\, \boldsymbol{m} + \frac{\omega_{\rm M}}{\mu_0} \left(\omega_0\, \boldsymbol{b} - \hat{\boldsymbol{z}} \times \dot{\boldsymbol{b}} + \omega_{\rm M}\, (\boldsymbol{b} \cdot \hat{\boldsymbol{y}})\, \hat{\boldsymbol{y}} \right) \tag{S5}$$

In order to satisfy the second equation (S2), relation (S5) reduces to,

$$\ddot{\boldsymbol{m}} = -\,\omega_0\, (\omega_0 + \omega_{\rm M})\, \boldsymbol{m} \tag{S6}$$

and thus the terms in the second brackets of relation (S5) vanish, i.e.

$$\omega_0\, \boldsymbol{b} - \hat{\boldsymbol{z}} \times \dot{\boldsymbol{b}} + \omega_{\rm M}\, (\boldsymbol{b} \cdot \hat{\boldsymbol{y}})\, \hat{\boldsymbol{y}} = \boldsymbol{0} \tag{S7}$$

which implies that,

$$\dot{\boldsymbol{b}} = -\,\omega_0\, (\hat{\boldsymbol{z}} \times \boldsymbol{b}) + \omega_{\rm M}\, (\boldsymbol{b} \cdot \hat{\boldsymbol{y}})\, \hat{\boldsymbol{x}} \tag{S8}$$

The time derivative of relation (S8) is given by,

$$\ddot{\boldsymbol{b}} = -\,\omega_0 \left(\hat{\boldsymbol{z}} \times \dot{\boldsymbol{b}}\right) + \omega_{\mathrm{M}} \left(\dot{\boldsymbol{b}} \cdot \hat{\boldsymbol{y}}\right) \hat{\boldsymbol{x}} \tag{S9}$$

The substitution of (S9) into relation (S8) yields,

$$\ddot{\boldsymbol{b}} = -\,\omega_0 \left(\omega_0 + \omega_{\mathrm{M}}\right) \boldsymbol{b} \tag{S10}$$

The comparision between the relations (S1) and the relations (S6) and (S10) respectively yields the Kittel resonance frequency [1],

$$\omega_{\mathrm{K}} = \pm\sqrt{\omega_0 \left(\omega_0 + \omega_{\mathrm{M}}\right)} \tag{S11}$$

where the sign is selected so that ω_{K} has sign as the angular frequencies ω_0 and ω_{M}.

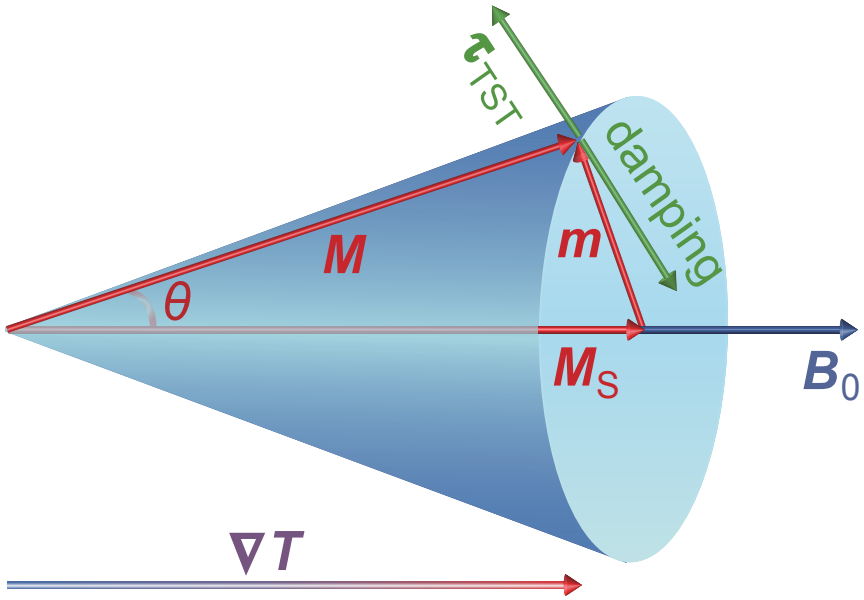

FIG. S5. Magnetization $\boldsymbol{M}$ precessing at an angle θ away from the equilibrium $\boldsymbol{M}_{\mathrm{S}}$ parallel to the applied field $\boldsymbol{B}_0$, thermal spin torque $\boldsymbol{\tau}_{\mathrm{TST}}$ orthogonal to $\boldsymbol{M}$.

Second, we analyze magnetization dynamics in the presence of a temperature gradient and derive a thermal spin torque and show that it can be expressed in terms of a spin current.

Taking into account an applied temperature gradient $\boldsymbol{\nabla} T$, the Landau-Lifshitz-Gilbert (LLG) equation for the time evolution of the magnetisation $\boldsymbol{M}$ contains an additional thermal spin torque (TST) term, i.e. [2],

$$\dot{\boldsymbol{M}} = \gamma\, \boldsymbol{M} \times \boldsymbol{B}_{\text{eff}} + \frac{\alpha}{M_{\text{S}}}\, \boldsymbol{M} \times \dot{\boldsymbol{M}} + \gamma\, \boldsymbol{M} \times \boldsymbol{B}_{\text{TST}} \tag{S12}$$

where $\boldsymbol{B}_{\text{eff}}$ is the effective magnetic field, $\boldsymbol{B}_{\text{TST}}$ is a magnetic field induced by the temperature gradient, $\gamma < 0$ is the gyromagnetic ratio, α is the magnetic damping parameter and M_{S} is the saturation magnetisation. The effective magnetic field $\boldsymbol{B}_{\text{eff}}$ is composed of the external field $\boldsymbol{B}_0$, the demagnetising field $\boldsymbol{B}_{\text{dem}}$, the anisotropy field $\boldsymbol{B}_{\text{ani}}$ and the microwave excitation field $\boldsymbol{b}$ induced by the microwave antenna. In order to focus on the significance of the thermal spin torque, we neglect the anisotropy field $\boldsymbol{B}_{\text{ani}}$. For the thin strip (Fig.1a in main text), the demagnetising field $\boldsymbol{B}_{\text{dem}}$ is given by, i.e.

$$\boldsymbol{B}_{\text{dem}} = -\,\mu_0\, (\boldsymbol{M} \cdot \hat{\boldsymbol{n}})\; \hat{\boldsymbol{n}} \tag{S13}$$

where $\hat{\boldsymbol{n}} = \hat{\boldsymbol{x}}$ is the unit vector orthogonal to the microstrip plane. The magnetisation $\boldsymbol{M}$ is the sum of the saturation magnetisation $\boldsymbol{M}_{\text{S}}$ and a magnetic response field $\boldsymbol{m}$ oscillating in a plane orthogonal to $\boldsymbol{M}_{\text{S}}$. In the linear response, the LLG equation (S12) is explicitly expressed as,

$$\dot{\boldsymbol{m}} = \gamma\, (\boldsymbol{m} \times \boldsymbol{B}_0 + \boldsymbol{M}_{\text{S}} \times \boldsymbol{b}) - \gamma\, \mu_0\, (\boldsymbol{m} \cdot \hat{\boldsymbol{x}})\, (\boldsymbol{M}_{\text{S}} \times \hat{\boldsymbol{x}}) + \frac{\alpha}{M_{\text{S}}}\, \boldsymbol{M}_{\text{S}} \times \dot{\boldsymbol{m}} + \gamma\, \boldsymbol{M}_{\text{S}} \times \boldsymbol{B}_{\text{TST}} \tag{S14}$$

The thermal spin torque is due to the thermal magnetic field $\boldsymbol{B}_{\text{TST}}$ obtained using a variational principle for the magnetisation [2],

$$\boldsymbol{B}_{\text{TST}} = -\,\mu_0\, \hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla}^{-1} \left(\frac{\partial}{\partial z}\, \left(\chi^{-1}\, \boldsymbol{m} \right) \right) \tag{S15}$$

where χ is the magnetic susceptibility of the YIG microstrip. To first-order, the heat-driven magnetic field $\boldsymbol{B}_{\text{TST}}$ is recast as,

$$\boldsymbol{B}_{\text{TST}} = \mu_0\, \frac{1}{\chi^2}\, \frac{\partial \chi}{\partial T}\, (\boldsymbol{\nabla} T) \cdot \boldsymbol{\nabla}^{-1}\, \boldsymbol{m} \tag{S16}$$

where the temperature gradient $\boldsymbol{\nabla} T$ is applied along the $\hat{\boldsymbol{z}}$-axis. According to the expression (S16), the heat-driven magnetic field $\boldsymbol{B}_{\mathrm{TST}}$ can also be written as,

$$\boldsymbol{B}_{\mathrm{TST}} = -\,\mu_0 \left(\boldsymbol{k}_{\mathrm{T}} \cdot \boldsymbol{\nabla}^{-1}\right) \boldsymbol{m} \tag{S17}$$

where the thermal wave vector $\boldsymbol{k}_{\mathrm{T}}$ is given by,

$$\boldsymbol{k}_{\mathrm{T}} = -\,\frac{1}{\chi^2}\,\frac{\partial \chi}{\partial T}\,\boldsymbol{\nabla} T \tag{S18}$$

According to the relations (S14) and (S17), the linearised LLG equation is expressed as,

$$\dot{\boldsymbol{m}} = \gamma\,(\boldsymbol{m} \times \boldsymbol{B}_0 + \boldsymbol{M}_{\mathrm{S}} \times \boldsymbol{b}) - \gamma\,\mu_0\,M_{\mathrm{S}}\,(\boldsymbol{m} \cdot \hat{\boldsymbol{x}})\;\hat{\boldsymbol{y}} + \frac{\alpha}{M_{\mathrm{S}}}\,\boldsymbol{M}_{\mathrm{S}} \times \dot{\boldsymbol{m}} - \gamma\,\mu_0\,\boldsymbol{M}_{\mathrm{S}} \times \left(\boldsymbol{k}_{\mathrm{T}} \cdot \boldsymbol{\nabla}^{-1}\right) \boldsymbol{m} \tag{S19}$$

The external magnetic field $\boldsymbol{B}_0$, the saturation magnetisation $\boldsymbol{M}_{\mathrm{S}}$ and the thermal wave vector $\boldsymbol{k}_{\mathrm{T}}$ are oriented along the $\hat{\boldsymbol{z}}$-axis, i.e.

$$\boldsymbol{B}_0 = B_0\,\hat{\boldsymbol{z}} \qquad \text{and} \qquad \boldsymbol{M}_{\mathrm{S}} = M_{\mathrm{S}}\,\hat{\boldsymbol{z}} \qquad \text{and} \qquad \boldsymbol{k}_{\mathrm{T}} = k_{\mathrm{T}}\,\hat{\boldsymbol{z}} \tag{S20}$$

and the magnetic excitation field $\boldsymbol{b}$ and the magnetic response field $\boldsymbol{m}$ are precessing in a plane orthogonal to $\boldsymbol{B}_0$, i.e.

$$\boldsymbol{b} = b_x\,\hat{\boldsymbol{x}} + b_y\,\hat{\boldsymbol{y}} \qquad \text{and} \qquad \boldsymbol{m} = m_x\,\hat{\boldsymbol{x}} + m_y\,\hat{\boldsymbol{y}} \tag{S21}$$

For convenience, we introduce the angular frequencies ω_0 and ω_{M} that are given respectively by,

$$\omega_0 = -\,\gamma\,B_0 > 0 \qquad \text{and} \qquad \omega_{\mathrm{M}} = -\,\gamma\,\mu_0\,M_{\mathrm{S}} > 0 \tag{S22}$$

since the gyromagnetic ratio of an electron is negative, i.e. $\gamma < 0$. Using the definitions (S22), the linearised LLG equation (S19) is recast as,

$$\dot{\boldsymbol{m}} = \omega_0\,(\hat{\boldsymbol{z}} \times \boldsymbol{m}) - \frac{\omega_{\mathrm{M}}}{\mu_0}\,(\hat{\boldsymbol{z}} \times \boldsymbol{b}) + \omega_{\mathrm{M}}\,(\boldsymbol{m} \cdot \hat{\boldsymbol{x}})\;\hat{\boldsymbol{y}} + \alpha\,(\hat{\boldsymbol{z}} \times \dot{\boldsymbol{m}}) + \omega_{\mathrm{M}}\,\hat{\boldsymbol{z}} \times \left(\boldsymbol{k}_{\mathrm{T}} \cdot \boldsymbol{\nabla}^{-1}\right) \boldsymbol{m} \tag{S23}$$

The propagation of the magnetisation waves occurs along the $\hat{\boldsymbol{z}}$-axis. In a stationary state, the time evolution of the magnetic response field is given by,

$$\dot{\boldsymbol{m}} = \omega\,(\hat{\boldsymbol{z}} \times \boldsymbol{m}) \tag{S24}$$

since for an electron the precession of the magnetisation is counterclockwise around the $\hat{\boldsymbol{z}}$-axis. The magnetisation waves propagating in the YIG microstrip are magnetostatic background volume modes characterised by the fact that the wave vector $\boldsymbol{k} = -k\,\hat{\boldsymbol{z}}$ is opposed to the propagation direction $\hat{\boldsymbol{z}}$, i.e.

$$\boldsymbol{k} \cdot \boldsymbol{z} = -kz \tag{S25}$$

Thus, the magnetic response field $\boldsymbol{m}$ can be expanded in plane waves of wave vector $\boldsymbol{k}$, i.e.

$$\boldsymbol{m} = \sum_k \boldsymbol{m}_k = \sum_k \Big(m_{kx}\,\cos\left(\omega\,t + k\,z - \phi\right)\,\hat{\boldsymbol{x}} + m_{ky}\,\sin\left(\omega\,t + k\,z - \phi\right)\,\hat{\boldsymbol{y}} \Big) \tag{S26}$$

Similarly, the magnetic excitation field $\boldsymbol{b}$ is expanded in plane waves, i.e.

$$\boldsymbol{b} = \sum_k \boldsymbol{b}_k = \sum_k \Big(b_{kx}\,\cos\left(\omega\,t + k\,z\right)\,\hat{\boldsymbol{x}} + b_{ky}\,\sin\left(\omega\,t + k\,z\right)\,\hat{\boldsymbol{y}} \Big) \tag{S27}$$

According to the relation (S26), the time derivative of the magnetic response field $\boldsymbol{m}$ is given by,

$$\dot{\boldsymbol{m}} = \sum_k \omega\,(\hat{\boldsymbol{z}} \times \boldsymbol{m}_k) \tag{S28}$$

In order to recast the last term of the linearised LLG equation (S19) in the form of a Gilbert term, we apply the operatorial identity $\boldsymbol{\nabla}^{-1} \cdot \boldsymbol{\nabla} = \mathbb{1}$ along the $\hat{\boldsymbol{z}}$-axis on the vector $\sum_k \frac{1}{k}\,(\hat{\boldsymbol{z}} \times \boldsymbol{m}_k)$, i.e.

$$\boldsymbol{k}_{\mathrm{T}} \cdot \boldsymbol{\nabla}^{-1} \left(\hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \left(\sum_k \frac{1}{k}\,(\hat{\boldsymbol{z}} \times \boldsymbol{m}_k) \right) \right) = \sum_k \frac{k_{\mathrm{T}}}{k}\,(\hat{\boldsymbol{z}} \times \boldsymbol{m}_k) \tag{S29}$$

where according to the relation (S26)

$$\begin{aligned}
\hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \left(\frac{1}{k}\,(\hat{\boldsymbol{z}} \times \boldsymbol{m}_k) \right) &= \hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \left(- m_{ky}\,\frac{\sin\left(\omega\,t + \,k\,z - \phi\right)}{k}\,\hat{\boldsymbol{x}} + m_{kx}\,\frac{\cos\left(\omega\,t + \,k\,z - \phi\right)}{k}\,\hat{\boldsymbol{y}} \right) \\
&= -\,m_{kx}\,\cos\left(\omega\,t + k\,z - \phi\right)\,\hat{\boldsymbol{x}} - \,m_{ky}\,\sin\left(\omega\,t + k\,z - \phi\right)\,\hat{\boldsymbol{y}} = -\,\boldsymbol{m}_k
\end{aligned}$$

This implies that the identity (S29) reduces to,

$$\left(\boldsymbol{k}_{\mathrm{T}} \cdot \boldsymbol{\nabla}^{-1}\right)\,\boldsymbol{m} = -\sum_k \frac{k_{\mathrm{T}}}{k}\,(\hat{\boldsymbol{z}} \times \boldsymbol{m}_k) \tag{S30}$$

We define the thermal spin torque damping parameter α_{TST} as,

$$\alpha_{\mathrm{TST}} = -\frac{\omega_{\mathrm{M}}}{\omega}\,\frac{k_{\mathrm{T}}}{k} \tag{S31}$$

and thus the effective damping parameter $\alpha_{\,\mathrm{eff}}$ is given by,

$$\alpha_{\,\mathrm{eff}} = \alpha + \alpha_{\mathrm{TST}} \tag{S32}$$

We now determine the magnetic susceptibility χ for the linear response at resonance. Using the relations (S28), (S30), (S31) and (S22), the linearised LLG equation (S23) is recast in reciprocal space as,

$$(\omega - \omega_0)\,(\hat{\boldsymbol{z}} \times \boldsymbol{m}_k) - \omega_{\mathrm{M}}\,(\boldsymbol{m}_k \cdot \hat{\boldsymbol{x}})\,\hat{\boldsymbol{y}} - \alpha_{\,\mathrm{eff}}\,\omega\left(\hat{\boldsymbol{z}} \times (\hat{\boldsymbol{z}} \times \boldsymbol{m}_k)\right) = -\frac{\omega_{\mathrm{M}}}{\mu_0}\,\hat{\boldsymbol{z}} \times \boldsymbol{b}_k \tag{S33}$$

which can be recast as,

$$(\omega - \omega_0)\,\boldsymbol{m}_k - \omega_{\mathrm{M}}\,(\boldsymbol{m}_k \cdot \hat{\boldsymbol{x}})\,\hat{\boldsymbol{x}} + \alpha_{\,\mathrm{eff}}\,\omega\,(\boldsymbol{m}_k \cdot \hat{\boldsymbol{y}})\,\hat{\boldsymbol{x}} - \alpha_{\,\mathrm{eff}}\,\omega\,(\boldsymbol{m}_k \cdot \hat{\boldsymbol{x}})\,\hat{\boldsymbol{y}} = -\frac{\omega_{\mathrm{M}}}{\mu_0}\,\boldsymbol{b}_k \tag{S34}$$

The relation (S34) is recast in matrix form as,

$$\begin{pmatrix} \omega - \omega_0 - \omega_{\mathrm{M}} & \alpha_{\,\mathrm{eff}}\,\omega \\ -\alpha_{\,\mathrm{eff}}\,\omega & \omega - \omega_0 \end{pmatrix} \begin{pmatrix} \boldsymbol{m}_k \cdot \hat{\boldsymbol{x}} \\ \boldsymbol{m}_k \cdot \hat{\boldsymbol{y}} \end{pmatrix} = -\frac{\omega_{\mathrm{M}}}{\mu_0} \begin{pmatrix} \boldsymbol{b}_k \cdot \hat{\boldsymbol{x}} \\ \boldsymbol{b}_k \cdot \hat{\boldsymbol{y}} \end{pmatrix} \tag{S35}$$

At resonance in the GHz range, the angular frequencies satisfy the following condition,

$$\alpha_{\,\mathrm{eff}} \ll \frac{\omega - \omega_0}{\omega} \quad \text{and} \quad \alpha_{\,\mathrm{eff}} \ll \frac{\omega - \omega_0 - \omega_{\mathrm{M}}}{\omega} \tag{S36}$$

which implies that $\phi \ll 1$. Thus, taking into account the conditions (S36), the magnetic constitutive relation (S35) is recast as,

$$\boldsymbol{b}_k = \mu_0\,\boldsymbol{\chi}^{-1} \cdot \boldsymbol{m}_k \tag{S37}$$

where the inverse of the magnetic susceptibility tensor is expressed as,

$$\boldsymbol{\chi}^{-1} = \begin{pmatrix} \chi^{-1} + 1 & 0 \\ 0 & \chi^{-1} \end{pmatrix} \tag{S38}$$

and the magnetic susceptibility scalar χ yields,

$$\chi = -\frac{\omega_{\mathrm{M}}}{\omega - \omega_0} \tag{S39}$$

The temperature derivative of the inverse magnetic susceptibility tensor (S38) is given by,

$$\frac{\partial \boldsymbol{\chi}^{-1}}{\partial T} = \frac{\partial \chi^{-1}}{\partial T}\,\mathbb{1} = -\frac{1}{\chi^2}\,\frac{\partial \chi}{\partial T}\,\mathbb{1} \tag{S40}$$

Taking into account the relation (S39), the temperature derivative of the magnetic susceptibility χ yields,

$$\frac{d\chi}{dT} = -\frac{\omega_{\mathrm{M}}}{\omega - \omega_0}\,\left|\,\frac{1}{M_{\mathrm{S}}}\,\frac{dM_{\mathrm{S}}}{dT}\,\right| \tag{S41}$$

Using the expressions (S11), and the relations (S39) and (S41), the thermal wave vector (S18) is recast at resonance as,

$$\boldsymbol{k}_{\mathrm{T}} = \frac{\omega_{\mathrm{K}} - \omega_0}{\omega_{\mathrm{M}}}\,\left|\,\frac{1}{M_{\mathrm{S}}}\,\frac{dM_{\mathrm{S}}}{dT}\,\right|\,\boldsymbol{\nabla}\,T \tag{S42}$$

The predicted effect is inversely proportional to the wave numbers k (S31). In our analysis, we shall only consider the dominant mode. Using the expression (S42) for the thermal wave vector $\boldsymbol{k}_{\mathrm{T}}$ and the expression $\boldsymbol{\nabla}\,T = \nabla_z\,T\,\hat{\boldsymbol{z}}$ for the thermal gradient, the expression (S31) for the thermal spin torque damping parameter α_{TST} becomes,

$$\alpha_{\mathrm{TST}} = -\left(1 - \frac{\omega_0}{\omega_{\mathrm{K}}}\right)\,\left|\,\frac{1}{M_{\mathrm{S}}}\,\frac{dM_{\mathrm{S}}}{dT}\,\right|\,\frac{1}{k}\,\nabla_z\,T \tag{S43}$$

The thermal spin torque damping parameter α_{TST} changes sign under reversals of either the temprature gradient, the propagation direction or the applied magnetic field. According to relation (S32), the effective damping parameter $\alpha_{\,\mathrm{eff}}$ is the sum of the Gilbert damping parameter α and the heat driven spin torque damping parameter α_{TST}. The inhomogeneous line width is frequency independent and the homogeneous line width is proportional to the damping parameter, [3], i.e.

$$\Delta B = \Delta B_0 + \frac{2}{\sqrt{3}}\,\alpha_{\,\mathrm{eff}}\,\left|\,\frac{\omega}{\gamma}\,\right| \tag{S44}$$

At resonance, i.e. $\omega = \omega_{\rm K}$, using the relation (S43), the expression (S44) for the homogeneous line width is recast explicitly as,

$$\Delta B = \Delta B_0 + \frac{2}{\sqrt{3}}\,\alpha\,\left|\,\frac{\omega_{\rm K}}{\gamma}\,\right| - \frac{2}{\sqrt{3}}\,\left|\,\frac{\omega_{\rm K}-\omega_0}{\gamma}\,\right|\,\left|\,\frac{1}{M_{\rm S}}\,\frac{dM_{\rm S}}{dT}\,\right|\,\frac{1}{k}\,\nabla_z T \tag{S45}$$

We now determine an explicit expression of the thermal spin torque $\boldsymbol{\tau}_{\rm TST}$. To first-order, the thermal spin torque is given by,

$$\boldsymbol{\tau}_{\rm TST} = \boldsymbol{M}_{\rm S} \times \boldsymbol{B}_{\rm TST} \tag{S46}$$

Using the expression (S17) of the heat-driven magnetic field $\boldsymbol{B}_{\rm TST}$ and the definition (S22) of the angular frequency $\omega_{\rm M}$, the thermal spin torque (S46) becomes,

$$\boldsymbol{\tau}_{\rm TST} = \frac{\omega_{\rm M}}{\gamma}\,\hat{\boldsymbol{z}} \times \left(\boldsymbol{k}_{\rm T}\cdot\boldsymbol{\nabla}^{-1}\right)\boldsymbol{m} \tag{S47}$$

At resonance, i.e. $\omega = \omega_{\rm K}$, the expression (S31) of the heat-driven damping coefficent $\alpha_{\rm TST}$ yields,

$$\alpha_{\rm TST} = -\,\frac{\omega_{\rm M}}{\omega_{\rm K}}\,\frac{k_{\rm T}}{k} \tag{S48}$$

For the dominant k-mode, using the vectorial identity (S30) and the expression (S48) of $\alpha_{\rm TST}$, the expression (S46) of the thermal spin torque reduces to,

$$\boldsymbol{\tau}_{\rm TST} = -\,\alpha_{\rm TST}\,\frac{\omega_{\rm K}}{|\,\gamma\,|}\,\boldsymbol{m}_k \tag{S49}$$

since $\gamma < 0$. The thermal spin torque (S49) can be recast as,

$$\boldsymbol{\tau}_{\rm TST} = \alpha_{\rm TST}\,\frac{\omega_{\rm K}}{|\,\gamma\,|\,M_{\rm S}^2}\,\boldsymbol{M}_{\rm S} \times (\boldsymbol{M}_{\rm S} \times \boldsymbol{m}_k) \tag{S50}$$

In the non-linear response where the precession cone angle $\theta = m_k/M_{\rm S}$ is large, the torque is orthogonal to the magnetisation $\boldsymbol{M}$, i.e.

$$\boldsymbol{\tau}_{\rm TST} = \alpha_{\rm TST}\,\frac{\omega_{\rm K}}{|\,\gamma\,|\,M_{\rm S}^2}\,\boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{m}_k) \tag{S51}$$

The thermal spin torque (S51) is parallel (i.e. $\alpha_{\rm TST} > 0$) or anti-parallel (i.e. $\alpha_{\rm TST} < 0$) to the magnetic damping torque corresponding to the Gilbert damping term.

We now express the thermal spin torque $\boldsymbol{\tau}_{\mathrm{TST}}$ in terms of a spin current density. Taking into account the expression (S22) of ω_{M} and the expression (S48) of α_{TST}, the thermal spin torque (S51) can be recast as,

$$\boldsymbol{\tau}_{\mathrm{TST}} = -\frac{k_{\mathrm{T}}}{M_{\mathrm{S}}^2}\,\boldsymbol{M}\times(\boldsymbol{M}\times\boldsymbol{j}_{\mathrm{s}}) \tag{S52}$$

where the spin current density vector $\boldsymbol{j}_{\mathrm{s}}$ is given by,

$$\boldsymbol{j}_{\mathrm{s}} = \frac{\mu_0\,M_{\mathrm{S}}}{k}\,\boldsymbol{m}_k \tag{S53}$$

The expression (S52) of the thermal spin torque (TST) has the same geometry as the spin-transfer torque (STT) in a metallic ferromagnet [5]. The thermal spin current density vector $\boldsymbol{j}_{\mathrm{s}}$ in the thermal spin torque (TST) plays an analogous role to the charge current density in the spin transfer torque (STT). To first-order in the linear response, the thermal spin torque (S52) reduces to,

$$\boldsymbol{\tau}_{\mathrm{TST}} = k_{\mathrm{T}}\,\boldsymbol{j}_{\mathrm{s}} \tag{S54}$$

Alternatively, using the expression (S17) for the heat-driven magnetic field $\boldsymbol{B}_{\mathrm{TST}}$, the thermal spin torque (S46) is expressed as,

$$\boldsymbol{\tau}_{\mathrm{TST}} = \boldsymbol{k}_{\mathrm{T}}\cdot\mathrm{j}_{\mathrm{s}} \tag{S55}$$

where the dot stands for the tensor contraction. The i-component of the thermal spin torque is given explicitly by,

$$\left(\tau_{\mathrm{TST}}\right)_i = \sum_{j=1}^{3}\left(k_{\mathrm{T}}\right)^j\left(j_{\mathrm{s}}\right)_{ij} = -\sum_{j,k,\ell=1}^{3}\mu_0\,\varepsilon_{\,ik\ell}\left(M_{\mathrm{S}}\right)^k\left(k_{\mathrm{T}}\right)^j\partial_j^{-1}\left(m_k\right)^\ell$$

where $\varepsilon_{\,ik\ell}$ is the totally antisymmetric Levi-Civita tensor. Thus, the heat-driven spin current tensor j_{s} yields,

$$\mathrm{j}_{\mathrm{s}} = -\,\mu_0\,\boldsymbol{M}_{\mathrm{S}}\times\boldsymbol{\nabla}^{-1}\,\boldsymbol{m}_k \tag{S56}$$

Note that the spin current density vector $\boldsymbol{j}_{\mathrm{s}}$ and the spin current density tensor j_{s} have the same physical dimensions, i.e. $\mathrm{J/m^2}$ in SI units, that correspond to the product of the spin density and the phase velocity of the heat-driven spin waves.

II. FIELD AT RESONANCE AS A FUNCTION OF TEMPERATURE

Isothermal measurements are conducted when both Peltier elements are set at the same temperature. We observe a systematic change of the value of the field at resonance with respect to the temperature T, which is attributed to the temperature dependence of the saturation magnetisation. (Fig.S1a). We observe that $\frac{dB_0}{dT} = 0.18\,\mathrm{mT/K}$ for which we deduce $\frac{1}{M_\mathrm{S}}\frac{dM_\mathrm{S}}{dT} = -3.8 \times 10^{-4}$ using the Kittel formula. The absolute value of saturation magnetization M_s was accurately measured by D. Kelly et al [6]. In a second

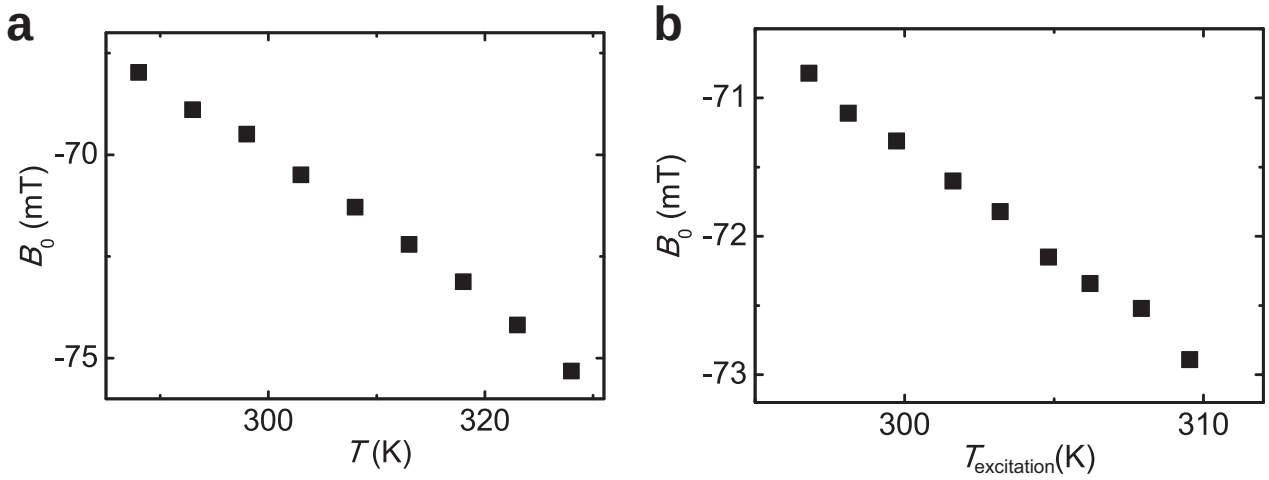

FIG. S6. **Field at resonance as a function temperature. a**, Isothermal measurement of field at resonance. **b**, Field at resonance as a function of injector temperature. Detector maintained at 303 K. Microwave frequency: 4.2 GHz.

experiment, we vary only the excitation temperature, and keep the detection temperature relatively constant. We find again a clear temperature dependence for the field B_0 at resonance(Fig.S1b). We extract the slope $\frac{dB_0}{dT} = 0.16\,\mathrm{mT/K}$ from Fig.S1b, which is fairly similar to that of isothermal measurements (Fig.S1a). This indicates that our excitation and detection are local and what matters for the field at resonance is the temperature at the excitation.

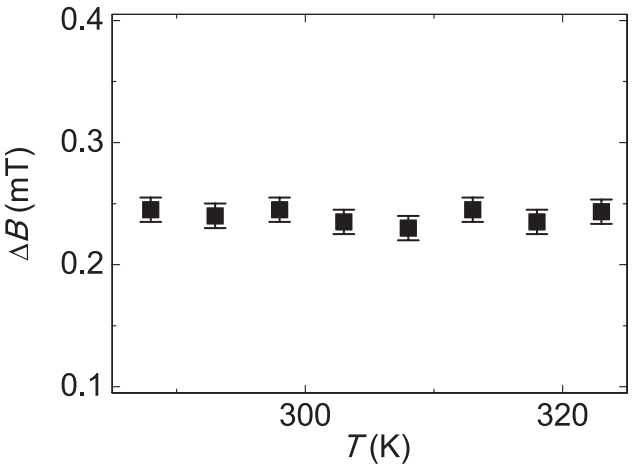

FIG. S7. Isothermal measurement of line width at 4.35 GHz as a function of temperature.

[1] C. Kittel, *Phys. Rev.* **73**, 155 (1948).

[2] Brechet, S. D. & Ansermet J.-P. Variational principle for magnetisation dynamics in a temperature gradient. *EPL* **112**, 17006 (2015).

[3] Gurevich, A. G. & Melkov, G. A. *Magnetization oscillations and waves* (CRC press, 1996).

[4] Obry, B., Vasyuchka, V. I., Chumak, A. V., Serga, A. A. & Hillebrands, B. Spin-wave propagation and transformation in a thermal gradient. *Appl. Phys. Lett.* **101**, 192406 (2012).

[5] Ralph, D. C. & Stiles, M. D. Spin transfer torques. *J. Magn. Magn. Mater.* **320**, 1190 (2008).

[6] O. d. Kelly *et al.*, Inverse spin hall effect in nanometer-thick yttrium iron garnet/pt system. *Appl. Phys. Lett.* **103**, 082408 (2013).