Spin-torque resonance due to diffusive dynamics at the surface of a topological insulator

R. J. Sokolewicz,[1] I. A. Ado,[1] M. I. Katsnelson,[1,2] P. M. Ostrovsky,[3,4] and M. Titov[1,5]

[1]*Radboud University, Institute for Molecules and Materials, NL-6525 AJ Nijmegen, The Netherlands*
[2]*Theoretical Physics and Applied Mathematics Department, Ural Federal University, Mira Str. 19, 620002 Ekaterinburg, Russia*
[3]*Max Planck Institute for Solid State Research, Heisenbergstrasse 1, 70569 Stuttgart, Germany*
[4]*L. D. Landau Institute for Theoretical Physics RAS, 119334 Moscow, Russia*
[5]*ITMO University, Saint Petersburg 197101, Russia*

We investigate spin-orbit torques on magnetization in an insulating ferromagnetic layer that is brought into close proximity to a topological insulator (TI). In addition to the well-known fieldlike spin-orbit torque, we identify an anisotropic anti-damping-like spin-orbit torque that originates in a diffusive motion of conduction electrons. This diffusive torque is vanishing in the limit of zero momentum (i.e., for a spatially homogeneous electric field or current), but it may, nevertheless, have a strong impact on spin-torque resonance at finite frequency provided the external field is neither parallel nor perpendicular to the TI surface. The required electric-field configuration can be created by a grated top gate.

I. INTRODUCTION

It is widely known that spin-orbit interaction provides an efficient way to couple electronic and magnetic degrees of freedom. It is, therefore, no wonder that the largest torque on magnetization, which is also referred to as the spin-orbit torque, emerges in magnetic systems with strong spin-orbit interaction [1,2], as has long been anticipated [3].

The spin-orbit coupling may be enhanced by confinement potentials in effectively two-dimensional systems consisting of conducting and magnetic layers. The in-plane current may efficiently drive domain walls or switch magnetic orientation in such structures with the help of spin-orbit torque [4–7], which is present even for uniform magnetization, or with the help of spin-transfer torque, which requires the presence of a magnetization gradient (due to, e.g., domain wall) [8–11].

Topological insulators (TIs) [12–15] may be thought of as materials with an ultimate spin-orbit coupling. Indeed, the effective Hamiltonian of conduction electrons at the TI surface contains essentially nothing but a spin-orbit interaction term that provides a perfect spin-momentum locking. Thus, the magnetization dynamics in a thin ferromagnetic (FM) film in proximity to a TI surface is expected to be strongly affected by electric currents and/or electric fields [16]. There seems to be, indeed, substantial experimental evidence that the efficiency of domain switching in TI/FM heterostructures is dramatically enhanced as compared to that in metals [17–22].

Nowadays the symmetry of spin-orbit torques is routinely inferred from the ferromagnetic resonance measurements in which an alternating microwave-frequency current (with frequencies 7–12 GHz) is applied within the sample plane [17,23–26].

II. TWO-DIMENSIONAL DIRAC FERMIONS WITH SD-INTERACTION ON THE TI/FM INTERFACE

In this work, we identify an anti-damping-like torque originating in a diffusive motion of conduction electrons at the TI surface. Such a torque originates in a nonlocal diffusive response of z component of the conduction electron spin density to the in-plane electric field. The nonlocality of the response is determined by the so-called diffusion pole in analogy to the density-density response of a disordered system. It is, however, important that the diffusive response of the spin-density in the TI is always present in the perpendicular-to-the-plane component of the spin density, irrespective of the magnetization direction in the FM. In nontopological FM/metal systems, such a diffusive response is present only in the spin-density component that is directed along the local magnetization of the FM. Thus, the diffusive antidamping spin-orbit torque, which we describe below, is specific for the TI/FM interfaces. Similarly, we identify a strong anisotropy of the Gilbert damping in the TI/FM system due to a combination of electron elastic scattering on nonmagnetic impurities and a spin-momentum locking in the TI.

Diffusive antidamping spin-orbit torque, which we are going to study, can be related to a response of conduction electron spin density to an electric field at a finite, but small, frequency and momentum. Such a field can be created, e.g., by applying an ac gate voltage to a grated top-gate as shown in Fig. 1. The presence of the diffusive spin-orbit torque can be detected by rather unusual spin-orbit-torque resonances in the TI/FM structures that we also investigate in this work.

Microscopic theory of current-induced magnetization dynamics in TI/FM heterostructures has been limited up to now to some particular cases: (i) the specific direction of
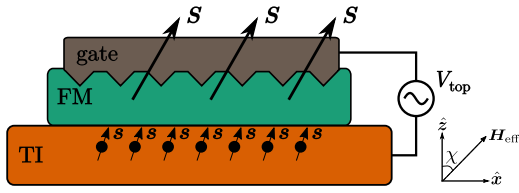
214444-1

FIG. 1. Proposed experimental setup. Nonhomogeneous in-plane electric-field components are created by an ac top-gate voltage V_{top} that induces a strong diffusive spin-orbit torque (4) of the damping-like symmetry. An effective magnetic field H is directed at the angle χ with respect to $\hat{z}$.

magnetization and (ii) the limit of vanishing exchange interaction between FM angular momenta and the spins of conduction electrons. In particular, an analytic estimate of spin-transfer and spin-orbit torques in a TI/FM bilayer was given in Ref. [27] for magnetization perpendicular to the TI surface. An attempt to generalize these results to an arbitrary magnetization direction was undertaken more recently in Ref. [28]. The nonlocal transport on the surface of the TI was first discussed in Ref. [29]. The results of that work were later applied to TI/FM systems [30,31] in a perturbative approach with respect to a weak sd-type exchange. However, the nonlocal behavior of nonequilibrium out-of-plane spin polarization in TI/FM systems, which gives rise to diffusive spin-orbit torques, was overlooked in all these publications.

To describe magnetization dynamics at a TI/FM interface, we employ an effective two-dimensional Dirac model for conduction electrons,

$$\mathcal{H} = v \left[(\boldsymbol{p} - e\boldsymbol{A}) \times \boldsymbol{\sigma} \right]_z - \Delta_{sd} \boldsymbol{m} \cdot \boldsymbol{\sigma} + V(\boldsymbol{r}), \qquad (1)$$

where $\boldsymbol{A}$ stands for the vector potential, $e = -|e|$ is the electron charge, z is the direction perpendicular to the TI surface, v is the effective velocity of Dirac electrons, and $V(\boldsymbol{r})$ is a disorder potential that models the main relaxation mechanism of conduction electrons. The energy $\Delta_{sd} = J_{sd}S$ is characterizing the local exchange interaction $\mathcal{H}_{ex} = -J_{sd} \sum_n \boldsymbol{S}_n \cdot c_n^\dagger \boldsymbol{\sigma} c_n$ between localized classical magnetic moments $\boldsymbol{S}_n$ on an FM lattice (with conserved absolute value $S = |\boldsymbol{S}_n|$ per unit cell area $\mathcal{A}$) and the electron spin density [represented by the vector operator $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ on the TI surface] [32]. Here σ_α stands for Pauli matrices and J_{sd} quantifies the sd-type exchange-interaction strength.

A classical equation of motion for the unit magnetization vector $\boldsymbol{m} = \boldsymbol{S}/S$ is determined by the sd-like exchange interaction $\mathcal{H}_{ex}$ as

$$\partial \boldsymbol{m}/\partial t = \gamma \boldsymbol{H} \times \boldsymbol{m} + \boldsymbol{T}, \quad \boldsymbol{T} = (J_{sd}\mathcal{A}/\hbar) \boldsymbol{m} \times \boldsymbol{s}, \quad (2)$$

where $\hbar = h/2\pi$ is the Planck constant and γ is a gyromagnetic ratio for the FM spin. The effective field $\boldsymbol{H}$ represents the combined contribution of the external magnetic field and the field produced by neighboring magnetic moments in the FM (e.g., due to direct exchange), while the term $\boldsymbol{T}$ represents the effect of the conduction electron spin density $\boldsymbol{s}(\boldsymbol{r}, t) = \langle c_n^\dagger \boldsymbol{\sigma} c_n \rangle$ on the TI surface.

To quantify the leading contributions to $\boldsymbol{T}$, we microscopically compute (i) a linear response of $\boldsymbol{s}$ to the in-plane electric field $\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}_{\boldsymbol{q},\omega} \exp(-i\omega t + i\boldsymbol{q} \cdot \boldsymbol{r})$, and (ii) a

linear response of $\boldsymbol{s}$ to the time derivative $\partial \boldsymbol{m}/\partial t$. The former response defines the spin-orbit torque, while the latter defines the Gilbert damping.

Before we proceed with the analysis, we shall note that the velocity operator $\boldsymbol{v} = v (\boldsymbol{\sigma} \times \hat{z})$ in the model of Eq. (1) is directly related to the spin operator $\boldsymbol{\sigma}$. As the result, the response of the in-plane spin density $\boldsymbol{s}_\parallel = (s_x, s_y)$ to electric field $\boldsymbol{E} = -\partial \boldsymbol{A}/\partial t$ is defined by the conductivity tensor [28,33]. This also means that the nonequilibrium contribution to $\boldsymbol{s}_\parallel$ from the electric current density $\boldsymbol{J}$ is given by $\boldsymbol{s}_\parallel = (\hat{z} \times \boldsymbol{J})/ev$ for any frequency and momentum irrespective of the type of scattering for conduction electrons and even beyond the linear response.

Thus, the response of $\boldsymbol{s}_\parallel$ defines an exceptionally universal fieldlike spin-orbit torque

$$\boldsymbol{T}_{FL}^{SOT} = (J_{sd}\mathcal{A}/\hbar ev) \boldsymbol{m} \times (\hat{z} \times \boldsymbol{J}), \qquad (3)$$

which acts in the same way as an in-plane external magnetic field applied perpendicular to the charge current.

Apart from the universal response of $\boldsymbol{s}_\parallel$, there might also exist a nonequilibrium spin polarization s_z perpendicular to the TI surface. This component plays no role in Eq. (2) for $\boldsymbol{m} = \pm \hat{z}$ due to the vector product involved. Also, the s_z component is vanishing by symmetry for $\boldsymbol{m} = \boldsymbol{m}_\parallel$, where we decompose $\boldsymbol{m} = \boldsymbol{m}_\parallel + \boldsymbol{m}_\perp$ to in-plane and perpendicular-to-the-plane components.

We find, however, that for a general direction of $\boldsymbol{m}$, the spin density s_z may be strongly affected by the in-plane electric field at a small but finite frequency and a small but finite wave vector. In the leading approximation, the result can be cast in the following form:

$$\boldsymbol{T}_{diff}^{SOT} = \eta \, \boldsymbol{m} \times \boldsymbol{m}_\perp \frac{iD\boldsymbol{q} \cdot \boldsymbol{E}}{i\omega - Dq^2}, \quad \eta = \frac{eJ_{sd}^2 \mathcal{A}S}{2\pi \hbar^3 v^2}, \qquad (4)$$

where D is a diffusion coefficient for conduction electrons at the TI surface, and $\boldsymbol{E} = \boldsymbol{E}_{\boldsymbol{q},\omega} \exp(-i\omega t + i\boldsymbol{q} \cdot \boldsymbol{r})$. Note that the diffusive torque is nonlinear with respect to $\boldsymbol{m}$ and, from the point of view of the time-reversal symmetry, is analogous to antidamping torque. The denominator $i\omega - Dq^2$ in Eq. (4) reflects diffusive (Brownian) motion of conduction electrons that defines the time-delayed diffusive torque on magnetization, $\boldsymbol{T}_{diff}^{SOT}$.

It is interesting to note that the torque of Eq. (4) has an antidamping symmetry (when expressed through electric current rather than electric field). Moreover, the torque formally diverges as $1/q$ in the dc limit $\omega = 0$. This singularity is well known in the theory of disordered systems [34,35] and originates in the diffusive (Brownian) motion of conduction electrons in a disorder potential. The dc limit singularity in Eq. (4) is, in fact, regularized by the dephasing length of conduction electrons on the surface of the TI. The length is strongly temperature- and material-dependent, and at low temperatures it can reach hundreds of microns. Thus, the result of Eq. (4) also predicts large antidamping spin-orbit torque in the dc limit that originates in a mechanism that is specific for the TI interface.

III. DERIVATION OF DIFFUSIVE SPIN-ORBIT TORQUE AND GILBERT DAMPING FROM LINEAR-RESPONSE THEORY

To derive the result of Eq. (4) and the expressions for Gilbert damping, we shall adopt a particular relaxation model for both spin and orbital angular momenta of conduction electrons. For the model of Eq. (1), those are provided by scattering on a disorder potential. We choose the latter to be the white-noise Gaussian disorder potential that is fully characterized by a single dimensionless parameter $\alpha \ll 1$,

$$\langle V(\boldsymbol{r}) \rangle = 0, \quad \langle V(\boldsymbol{r})V(\boldsymbol{r}') \rangle = 2\pi\alpha\,(\hbar v)^2\,\delta(\boldsymbol{r}-\boldsymbol{r}'), \quad (5)$$

where angular brackets denote the averaging over the ensemble of disordered systems.

Since both the vector potential $\boldsymbol{A}$ and the magnetization $\boldsymbol{m}$ couple to spin operators in Eq. (1), the linear response of $\boldsymbol{s}$ to $\boldsymbol{E} = -\partial\boldsymbol{A}/\partial t$ and $\partial\boldsymbol{m}/\partial t$ is defined in the frequency-momentum domain as

$$\boldsymbol{s} = (v^2 h)^{-1}\hat{K}(\boldsymbol{q},\omega)[ev\,(\boldsymbol{E}\times\hat{z}) - i\omega\Delta_{\rm sd}\,\boldsymbol{m}]. \quad (6)$$

Here, the dimensionless nine-component tensor $\hat{K}(\boldsymbol{q},\omega)$ is given by the Kubo formula

$$\hat{K}_{\alpha\beta}(\boldsymbol{q},\omega) = v^2\int\frac{d^2\boldsymbol{p}}{(2\pi)^2}\,\mathrm{Tr}\,\big\langle\sigma_\alpha G^R_{\boldsymbol{p}+\hbar\boldsymbol{q},\varepsilon+\hbar\omega}\sigma_\beta G^A_{\boldsymbol{p},\varepsilon}\big\rangle, \quad (7)$$

where the notation $G^{R(A)}_{\boldsymbol{p},\varepsilon}$ stands for the retarded (advanced) Green's function for the Hamiltonian of Eq. (1), the angular brackets denote the averaging over disorder realizations, while the energy ε refers to the Fermi energy (zero-temperature limit is assumed).

The tensor $\hat{K}$ can be represented by the matrix

$$\hat{K} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & Q_y \\ \sigma_{yx} & \sigma_{yy} & -Q_x \\ Q_y & -Q_x & \zeta \end{pmatrix}, \quad (8)$$

of which $\sigma_{\alpha\beta}$ are the components of the two-dimensional conductivity tensor at the TI surface (all conductivities are expressed in units of e^2/h), the vector $\boldsymbol{Q}$ defines the diffusive spin-orbit torque of Eq. (4) (its contribution to Gilbert damping is negligible), while ζ determines the response of s_z to $\partial m_z/\partial t$. The components of $\hat{K}$ correspond to different responses at different limits. When discussing the response to an electric field $\boldsymbol{E}_{\boldsymbol{q}\omega}$, we are primarily interested in the limit $\omega \ll Dq^2$, whereas the response to the time derivative of magnetization $\boldsymbol{m}$ is defined by the limit $q \to 0$.

In the linear-response theory of Eq. (6), one needs to compute the tensor in Eq. (7) for a constant direction $\boldsymbol{m}$ and for $\boldsymbol{A} = 0$. In usual systems (conducting ferromagnets), the response of $\boldsymbol{s}$ in the direction of $\boldsymbol{m}$ is always diffusive. This response, however, plays no role in the torque since $\boldsymbol{T} \propto \boldsymbol{m} \times \boldsymbol{s}$. The situation at the TI surface is, however, special. Here, the in-plane components of magnetization m_x, m_y play no role in Eq. (1), since those are simply equivalent to a constant in-plane vector potential for conduction electrons, and therefore they can be excluded by a gauge transform (shift of the Dirac cone). Consequently, all observable quantities in the model (including all components of the tensor $\hat{K}$) may only depend on the field $\Delta_z = \Delta_{\rm sd}m_z$. As a result, the diffusive response
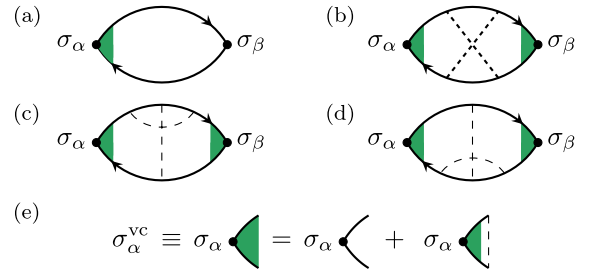

FIG. 2. Diagrams considered in the calculation of $\hat{K}$: (a) non-crossing diagram, (b) X diagram, and (c),(d) Ψ diagrams. Green areas indicate the ladder summation (e) for the vertex correction in the noncrossing approximation [36].

occurs exclusively in the s_z component of spin polarization and can easily enter the expression for the torque.

The conductivity tensor in the model of Eqs. (1) and (5) has been analyzed in detail in Ref. [36] in the limit $\omega = q = 0$ (and for $\alpha \ll 1$) with the result $\sigma_{xx} = \sigma_{yy} = \sigma_0$ and $\sigma_{xy} = -\sigma_{yx} = \sigma_{\rm H}$, where

$$\sigma_0 = \frac{\varepsilon^2 - \Delta_z^2}{\pi\alpha\,(\varepsilon^2 + 3\Delta_z^2)}, \quad \sigma_{\rm H} = \frac{8\varepsilon\Delta_z^3}{(\varepsilon^2 + 3\Delta_z^2)^2}. \quad (9)$$

Since the anomalous Hall conductivity $\sigma_{\rm H} \propto \alpha\,\sigma_0$ is subleading with respect to σ_0, it has to be computed beyond the Born approximation (see Refs. [36–38]).

Here we generalize the analysis to calculate the tensor $\hat{K}$ for finite ω and $\boldsymbol{q}$ assuming $\alpha \ll 1$, $\omega\tau_{\rm tr} \ll 1$, and $\omega \propto Dq^2$, where $D = \hbar v^2\sigma_0/\varepsilon$ is the diffusion coefficient and $\tau_{\rm tr} = \hbar\varepsilon\sigma_0/(\varepsilon^2 + \Delta_z^2)$ is the transport scattering time for the problem. In real samples, $\tau_{\rm tr} = 0.01$–1 ps [39–42].

A. Disorder averaging: Born approximation and vertex corrections

The main building block of our analysis is the averaged Green's function in the first Born approximation

$$G^R_{\boldsymbol{p},\varepsilon} = \frac{\varepsilon^R + v(\boldsymbol{p}\times\boldsymbol{\sigma})_z - \Delta_z^R\sigma_z}{(\varepsilon^R)^2 - v^2p^2 - (\Delta_z^R)^2}, \quad (10)$$

where the complex parameters $\varepsilon^R = \varepsilon(1 + i\pi\alpha/2)$ and $\Delta_z^R = \Delta_z(1 - i\pi\alpha/2)$ are found from the corresponding self-energy

$$\Sigma^R(\varepsilon) = 2\pi\alpha\,v^2\int\frac{d^2\boldsymbol{p}}{(2\pi)^2}G^R_{\boldsymbol{p},\varepsilon}, \quad (11)$$

which gives rise to $\mathrm{Im}\,\Sigma^R = -\pi\alpha(\varepsilon - \Delta_z\sigma_z)/2$ (strictly speaking, the RG analysis [36] has to be applied). In the Green's function of Eq. (10) we shift the momentum $\boldsymbol{p}$ such that there is no direct dependence on the in-plane magnetization components m_x and m_y.

The next step in disorder-averaging requires the computation of vertex corrections. This means we need to replace the spin operator σ_α with a vertex-corrected spin operator $\sigma_\alpha^{\rm vc}$ in the ladder approximation as depicted in Fig. 2(e). The crossed diagrams in Figs. 2(b)–2(d) give a contribution to the components of $\hat{K}$ of the order $O(\alpha^0)$. The only components that are modified to this order are those corresponding to the

Hall conductivity (i.e., σ_{xy} and σ_{yx}). Details of this calculation can be found in Ref. [36].

The dressing of σ_α with a single disorder line is denoted by $\sigma_\alpha^{1 \times \mathrm{dr}}$ and is conveniently represented in matrix form by introducing a matrix $\hat{M}$ with 16 components $M_{\alpha\beta}$ for $\alpha, \beta = 0, x, y, z$ (with $\sigma_0 = 1$),

$$\sigma_\alpha^{1 \times \mathrm{dr}} = 2\pi\alpha\, v^2 \int \frac{d^2\boldsymbol{p}}{(2\pi)^2} G_{\varepsilon+\omega, \boldsymbol{p}+\boldsymbol{q}}^{\mathrm{A}} \sigma_\alpha G_{\boldsymbol{p}}^{R} = \pi\alpha M_{\alpha\beta}\sigma_\beta, \quad (12)$$

where the summation of the repeating index $\beta = 0, x, y, z$ is assumed. Full expressions of the components of $\hat{M}$ up to second order in ω and q are given by Eqs. (B1a)–(B1f).

In our calculation, the terms of the order of $\alpha \ln p_{\mathrm{cutoff}}/\varepsilon$ (where p_{cutoff} is the ultraviolet momentum cutoff) are disregarded with respect to 1. This approximation is legitimate since we assume that all model parameters ϵ, Δ_{sd}, and α are first renormalized such that $p_{\mathrm{cutoff}} \approx \varepsilon$.

It is easy then to see that the vertex-corrected spin operator is readily obtained from the geometric series of powers of $\pi\alpha\hat{M}$,

$$\sigma_\alpha^{\mathrm{vc}} = \sigma_\alpha + \pi\alpha\hat{M}_{\alpha\beta}\sigma_\beta + (\pi\alpha)^2(\hat{M}^2)_{\alpha\beta}\sigma_\beta + \cdots$$
$$= [1 - \pi\alpha\hat{M}]_{\alpha\beta}^{-1}\sigma_\beta. \quad (13)$$

Thus, in the noncrossing approximation [illustrated in Fig. 2(a)], one simply finds $\hat{K} = \hat{M}[1 - \pi\alpha\hat{M}]^{-1}$.

Dressed spin-spin correlators are defined by the components $\hat{K}_{\alpha\beta}$ with $\alpha, \beta = x, y, z$. The vector $\boldsymbol{q}$ selects a particular direction in space that makes the conductivity tensor anisotropic. By choosing the x direction along the $\boldsymbol{q}$ vector, we find the conductivity components $\sigma_{xx} = \sigma_0$, $\sigma_{xy} = -\sigma_{yx} = \sigma_{\mathrm{H}}$, and $\sigma_{yy} = i\omega\sigma_0/(i\omega - Dq^2)$, where we have kept only the leading terms in the limits $\alpha \ll 1$, $\omega\tau_{\mathrm{tr}} \ll 1$ [more general expressions are given by Eqs. (B3a)–(B3d)]. We can see that the σ_{yy} component also acquires a diffusion pole. One needs to go beyond the noncrossing approximation in the computation of anomalous Hall conductivity [36–38].

B. Diffusive spin-orbit torque and Gilbert damping

Clearly, the components $\sigma_{\alpha\beta}$ define the fieldlike contribution $\boldsymbol{T}_{\mathrm{FL}}^{\mathrm{SOT}}$, which was already discussed above. It is interesting to note that the conductivity is isotropic, $\sigma_{xx} = \sigma_{yy} = \sigma_0$, only if the limit $q = 0$ is taken before the limit $\omega = 0$. If the limit $\omega = 0$ is taken first, the conductivity remains anisotropic with respect to the direction of $\boldsymbol{q}$ even for $q = 0$.

The vector $\boldsymbol{Q} = (Q_x, Q_y)$ quantifies both the response of s_z to an electric field or to $\partial\boldsymbol{m}_\parallel/\partial t$ as well as the response of $\boldsymbol{s}_\parallel$ to $\partial m_z/\partial t$. From Eq. (7) we find

$$\boldsymbol{Q}(\omega, \boldsymbol{q}) = \frac{\Delta_z}{\hbar v} \frac{iD\boldsymbol{q}}{i\omega - Dq^2}[1 + O(\omega\tau_{\mathrm{tr}})], \quad (14)$$

where we again assumed $\omega\tau_{\mathrm{tr}} \ll 1$. The result of Eq. (14) then corresponds to an additional diffusive spin-orbit torque of the form Eq. (4).

Finally, the response of s_z to $\partial m_z/\partial t$ is defined by

$$\zeta = \frac{\Delta_z^2}{i\hbar\varepsilon\omega}[1 + O(\omega^2\tau_{\mathrm{tr}}^2)], \quad (15)$$

where the limit $q = 0$ is taken. Thus, we find from Eq. (6) that there exists no response of s_z to $\partial m_z/\partial t$. Instead, the

quantity ζ defines the additional spin polarization in the z direction, $\delta s_z = -\Delta_{\mathrm{sd}}^3 m_z^3/(2\pi\hbar^2 v^2\varepsilon)$, which we ignore below. Equations (14) and (15) including subleading terms in $\omega\tau_{\mathrm{tr}}$ are presented in Eq. (B4).

We also note that $\boldsymbol{Q}(q = 0) = 0$, hence there is no term in s_z that is proportional to $\partial\boldsymbol{m}/\partial t$. This reflects the highly anisotropic nature of the Gilbert damping in the model of Eq. (1).

The remaining parts of the Gilbert damping can be cast in the following form:

$$\boldsymbol{T}^{\mathrm{GD}} = \frac{J_{\mathrm{sd}}^2 \mathcal{A}S}{\pi\hbar^2 v^2} \boldsymbol{m} \times \left(\sigma_0 \frac{\partial\boldsymbol{m}_\parallel}{\partial t} + \frac{\sigma_{\mathrm{H}}}{m_z} \frac{\partial\boldsymbol{m}_\parallel}{\partial t} \times \boldsymbol{m}_\perp\right), \quad (16)$$

where the coefficients σ_0 and σ_{H}/m_z from Eq. (9) depend on m_z^2, which is yet another source of the Gilbert damping anisotropy. We note that even though Eq. (16) does not contain a term proportional to $\partial m_z/\partial t$, the existing in-plane Gilbert damping is sufficient to relax the magnetization along the $\hat{z}$ direction.

Despite the strongly anisotropic nature of the diffusive torque (the torque is vanishing for purely in-plane or purely perpendicular-to-the-plane magnetization), its strength for a generic direction of magnetization may be quite large. For example, for $\boldsymbol{m}$ directed approximately at 45° to the TI surface, the ratio of amplitudes of diffusive and fieldlike torques is readily estimated as

$$\frac{T_{\mathrm{diff}}^{\mathrm{SOT}}}{T_{\mathrm{FL}}^{\mathrm{SOT}}} = \frac{\Delta_{\mathrm{sd}}}{\hbar qv} \frac{1}{\sigma_0}, \quad (17)$$

where we used the condition $\omega \ll Dq^2$. Let us assume that a top gate in Fig. 1 induces an ac in-plane electric field with the characteristic period $2\pi q^{-1} \approx 1\ \mu\mathrm{m}$ and a typical FM resonance frequency, $\omega \approx 7$–12 GHz. Then, for realistic materials one can estimate $Dq^2 \approx 100$ GHz, hence $\omega \ll Dq^2$ indeed. For a typical velocity $v = 10^6$ m/s one finds $\hbar qv \approx 4$ meV. Thus, the ratio $\Delta_{\mathrm{sd}}/\hbar qv$ in Eq. (17) may reach three orders of magnitude, while the value of σ_0 is typically 10. This estimate suggests that, for a generic direction of $\boldsymbol{m}$, the magnitude of diffusive torque can become three orders of magnitude larger than that of the fieldlike spin-orbit torque.

The diffusive torque at the TI surface can be most directly probed by the corresponding spin-torque resonance. In this case, one can disregard the effect of the fieldlike torque, so that Eq. (2) is simplified to

$$\frac{\partial\boldsymbol{m}}{\partial t} = \gamma \boldsymbol{H} \times \boldsymbol{m} + f(\boldsymbol{r}, t)\boldsymbol{m} \times \boldsymbol{m}_\perp + \alpha_{\mathrm{G}}\boldsymbol{m} \times \frac{\partial\boldsymbol{m}_\parallel}{\partial t}, \quad (18)$$

where $\alpha_{\mathrm{G}} = J_{\mathrm{sd}}^2 \mathcal{A}S\sigma_0/\pi(\hbar v)^2$ is the Gilbert damping amplitude (which is a constant for $\varepsilon \gg \Delta_{\mathrm{sd}}$), while the terms containing σ_{H} are omitted. The function

$$f(\boldsymbol{r}, t) = \eta \int d^2\boldsymbol{r}' \int_{-\infty}^{t} dt' \frac{e^{-(\boldsymbol{r}-\boldsymbol{r}')^2/4D(t-t')}}{4\pi(t-t')} \boldsymbol{\nabla} \cdot \boldsymbol{E}(\boldsymbol{r}', t')$$

defines the strength of the diffusive spin-orbit torque (4) in real space and time.
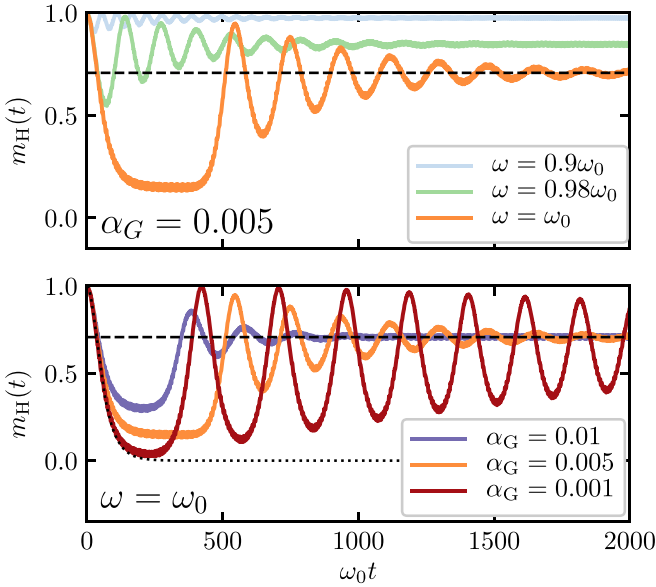
FIG. 3. The projection $m_H(t)$ as simulated from Eq. (18) for $f_0 = 0.1\,\omega_0$. The top panel illustrates the behavior at different frequencies for $\alpha_G = 0.005$. The lower panel illustrates the resonant behavior at different values of α_G. The dashed horizontal line corresponds to $m_H = 1/\sqrt{2}$. Dots indicate the asymptotic solution for $\alpha_G = 0$ as given by Eq. (19).

IV. RESONANT MAGNETIZATION DYNAMICS

Resonant magnetization dynamics defined by Eq. (18) is illustrated in Fig. 3 for $\boldsymbol{H}$ directed at the angle $\chi = \pi/4$ with respect to $\hat{z}$ and for frequencies that are close to the resonant frequency $\omega_0 = \gamma H$. The time evolution of magnetization projection $m_H = \boldsymbol{m} \cdot \boldsymbol{H}/H$ is induced by the diffusive torque with $f(t) = f_0 \cos \omega t$ (magnetization at different $\boldsymbol{r}$ is simply different by a phase).

Resonant dynamics at $\omega = \omega_0$ in Eq. (18) consists of a precession of $\boldsymbol{m}$ around the vector $\boldsymbol{H}$ such that the azimuth (precession) angle is changing linearly with time, $\phi(t) = \omega_0 t - \pi/2$ (for $f_0 \ll \omega_0$ and $\alpha_G \ll 1$). In addition, the projection m_H oscillates between 1 and 0 on much larger time scales. Such oscillations are damped by a finite α_G to the limiting value $m_H = 1/\sqrt{2}$.

In the limit of vanishing Gilbert damping, $\alpha_G = 0$, one simply finds the result

$$m_H(t) = \left\{ \cosh \left[\tfrac{1}{4} f_0 t \sin(2\chi) \right] \right\}^{-1}, \tag{19}$$

which clearly illustrates the absence of the effect for both perpendicular-to-the-plane ($\chi = 0$) and in-plane ($\chi = \pi/2$) magnetization. The qualitative behavior at the resonance ($\omega = \omega_0$) is illustrated in the lower panel of Fig. 3 for different values of α_G.

V. CONCLUSIONS

In conclusion, we consider magnetization dynamics in a model TI/FM system at a finite frequency ω and $\boldsymbol{q}$ vector. We identify diffusive antidamping spin-orbit torque that is specific to the TI/FM system. Such a torque is absent in usual (nontopological) FM/metal systems, where the diffu-

sive response of conduction electron spin density is always aligned with the magnetization direction of the FM. In contrast, the electrons at the TI surface give rise to a singular diffusive response of the conduction electron spin density in the direction perpendicular to the TI surface, irrespective of the FM magnetization direction. Such a response leads to strong nonadiabatic antidamping spin-orbit torque that has a diffusive nature. This response is specific for a system with an ultimate spin-momentum locking and gives rise to abnormal antidamping diffusive torque that can be detected by performing spin-torque resonance measurements. We also show that, in realistic conditions, the anti-damping-like diffusive torque may become orders of magnitude larger than the usual fieldlike spin-orbit torque. We investigate the peculiar magnetization dynamics induced by the diffusive torque at the frequency of the ferromagnet resonance. Our theory also predicts the ultimate anisotropy of the Gilbert damping in the TI/FM system. In contrast to the phenomenological approaches [43,44], our microscopic theory is formulated in terms of very few effective parameters. Our results are complementary to previous phenomenological studies of Dirac ferromagnets [45–67].

APPENDIX A: KUBO FORMULA

The linear-response formula used in the main text can be obtained in a Keldysh framework. We start by introducing the Green function $\mathcal{G}$ in rotated Keldysh space (see, e.g., Ref. [68]),

$$\mathcal{G} = \begin{pmatrix} G^R & G^K \\ 0 & G^A \end{pmatrix}, \tag{A1}$$

where R, A, and K denote retarded, advanced, and Keldysh Green functions, respectively. In this notation, a perturbation to a classical field $V(\boldsymbol{x}, t)$ is given by

$$\delta\mathcal{G}(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2) = \int d\boldsymbol{x}_3 \int dt_3 \, \mathcal{G}^{(0)}(\boldsymbol{x}_1, t_1; \boldsymbol{x}_3, t_3) \hat{V}(\boldsymbol{x}_3, t_3)$$
$$\times \mathcal{G}^{(0)}(\boldsymbol{x}_3, t_3; \boldsymbol{x}_2, t_2) + O(V^2) \tag{A2}$$

with $\mathcal{G}^{(0)}$ equilibrium Green functions. The Wigner transform of a function $F(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2)$ is given by

$$F(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2) = \int \frac{d^2\boldsymbol{p}}{(2\pi\hbar)^2} \int \frac{d\varepsilon}{2\pi\hbar} e^{-i\varepsilon(t_1 - t_2)/\hbar} e^{i\boldsymbol{p}\cdot(\boldsymbol{x}_1 - \boldsymbol{x}_2)/\hbar}$$
$$\times F(\varepsilon, \boldsymbol{p}, R, T) \tag{A3}$$

with energy ε, momentum $\boldsymbol{p}$, time $T = \frac{t_1 + t_2}{2}$, and position $\boldsymbol{R} = \frac{\boldsymbol{x}_1 + \boldsymbol{x}_2}{2}$. In equilibrium, the Green functions $\mathcal{G}^{(0)}$ do not depend on $\boldsymbol{R}$ and T, so that the momentum-frequency representation of Eq. (A2) becomes $\delta\mathcal{G}(\varepsilon, \omega, \boldsymbol{p}, \boldsymbol{q}) = \mathcal{G}^{(0)}_{\varepsilon_+, \boldsymbol{p}_+} V_{\omega, \boldsymbol{q}} \mathcal{G}^{(0)}_{\varepsilon_-, \boldsymbol{q}_-}$, with

subscripts $\varepsilon_\pm = \varepsilon \pm \hbar\omega/2$ and $\boldsymbol{p}_\pm = \boldsymbol{p} \pm \hbar\boldsymbol{q}/2$, and $V_{\omega,\boldsymbol{q}}$ is the Fourier transform of $V(\boldsymbol{R}, T)$.

The spin density $\boldsymbol{s}_{\omega,\boldsymbol{q}}$ is given by

$$\boldsymbol{s}_{\omega,\boldsymbol{q}} = i\hbar \int \frac{d\varepsilon}{2\pi\hbar} \int \frac{d^2\boldsymbol{p}}{(2\pi\hbar)^2} \, \mathrm{Tr}[\delta G^<(\varepsilon, \omega, \boldsymbol{p}, \boldsymbol{q}, T)\boldsymbol{\sigma}], \quad \text{(A4)}$$

where

$$\delta G^<(\varepsilon, \omega, \boldsymbol{p}, \boldsymbol{q}) = 1/2[\delta G^K(\varepsilon, \omega, \boldsymbol{p}, \boldsymbol{q}) - \delta G^R(\varepsilon, \omega, \boldsymbol{p}, \boldsymbol{q}) \\ + \delta G^A(\varepsilon, \omega, \boldsymbol{p}, \boldsymbol{q})]. \quad \text{(A5)}$$

In equilibrium we have the fluctuation-dissipation theorem $G^K_{\varepsilon_\pm, \boldsymbol{p}_\pm} = (1 - 2f_{\varepsilon_\pm})(G^R_{\varepsilon_\pm, \boldsymbol{p}_\pm} - G^A_{\varepsilon_\pm, \boldsymbol{p}_\pm})$ with $f_{\varepsilon_\pm}$ the Fermi distribution, so that the spin density now becomes

$$\boldsymbol{s}_{\omega,\boldsymbol{q}} = i\hbar \int \frac{d\varepsilon}{2\pi\hbar} \int \frac{d^2\boldsymbol{p}}{(2\pi\hbar)^2} \\ \times \mathrm{Tr}\Big\langle -(f_{\varepsilon_+} - f_{\varepsilon_-})\boldsymbol{\sigma} G^R_{\varepsilon_+, \boldsymbol{p}_+} V_{\omega,\boldsymbol{q}} G^A_{\varepsilon_-, \boldsymbol{p}_-} \\ - f_{\varepsilon_+}\boldsymbol{\sigma} G^R_{\varepsilon_+, \boldsymbol{p}_+} V_{\omega,\boldsymbol{q}} G^R_{\varepsilon_-, \boldsymbol{p}_-} + f_{\varepsilon_-}\boldsymbol{\sigma} G^A_{\varepsilon_+, \boldsymbol{p}_+} V_{\omega,\boldsymbol{q}} G^A_{\varepsilon_-, \boldsymbol{p}_-} \Big\rangle, \quad \text{(A6)}$$

where the angular brackets stand for impurity averaging. The latter amounts to the replacement of the Green's functions with the corresponding impurity-averaged Greens functions (in the Born approximation) and to the replacement of one of the spin operators with the corresponding vertex-corrected operator (in the noncrossing approximation). The corrections beyond the noncrossing approximation are important for those tensor components that lack a leading-order contribution [36]. To keep our notations more compact, we ignore here the fact that the Green's functions before disorder averaging lack translational invariance, i.e., they depend on both Wigner coordinates: momentum and coordinate.

In the limit of small frequency, i.e., $\hbar\omega \ll \varepsilon$, we obtain $s_\alpha = s_\alpha^{\mathrm{I}} + s_\alpha^{\mathrm{II}}$,

$$s_\alpha^{\mathrm{I}} = \frac{i\omega}{2\hbar} \int \frac{d\varepsilon}{2\pi} \int \frac{d^2\boldsymbol{p}}{(2\pi)^2} \left(-\frac{\partial f}{\partial\varepsilon} \right) \mathrm{Tr} \big\langle 2\sigma_\alpha G^R_{\varepsilon_+, \boldsymbol{p}_+} V_{\omega,\boldsymbol{q}} G^A_{\varepsilon_-, \boldsymbol{p}_-} \\ - \sigma_\alpha G^A_{\varepsilon_+, \boldsymbol{p}_+} V_{\omega,\boldsymbol{q}} G^A_{\varepsilon_-, \boldsymbol{p}_-} - \sigma_\alpha G^R_{\varepsilon_+, \boldsymbol{p}_+} V_{\omega,\boldsymbol{q}} G^R_{\varepsilon_-, \boldsymbol{p}_-} \big\rangle, \quad \text{(A7)}$$

$$s_\alpha^{\mathrm{II}} = \frac{i}{\hbar} \int \frac{d\varepsilon}{2\pi} \int \frac{d^2\boldsymbol{p}}{(2\pi)^2} f_\varepsilon \, \mathrm{Tr} \big\langle \sigma_\alpha G^A_{\varepsilon_+, \boldsymbol{p}_+} V_{\omega,\boldsymbol{q}} G^A_{\varepsilon_-, \boldsymbol{p}_-} \\ - \sigma_\alpha G^R_{\varepsilon_+, \boldsymbol{p}_+} V_{\omega,\boldsymbol{q}} G^R_{\varepsilon_-, \boldsymbol{p}_-} \big\rangle, \quad \text{(A8)}$$

where s^{I} and s^{II} are the Kubo and Streda contributions, respectively. The Streda contribution is subleading in powers of weak disorder strength $\alpha \ll 1$ as long as the Fermi energy lies outside the gap. Similarly, the AA and RR bubbles in the expression of s_α^{I} are subleading and may be neglected. Furthermore, we work in the zero-temperature limit.

The linear response to the electric field and the time derivative of magnetization corresponds to $V_{\boldsymbol{q},\omega} = -\hat{\boldsymbol{j}} \cdot \boldsymbol{A} - \Delta_{\mathrm{sd}}\boldsymbol{m} \cdot \boldsymbol{\sigma}$, so that we obtain

$$\boldsymbol{s}_{\boldsymbol{q},\omega} = \frac{1}{v^2 h}\hat{K}(\boldsymbol{q}, \omega)[ev(\boldsymbol{E}_{\boldsymbol{q},\omega} \times \hat{\boldsymbol{z}}) - i\omega\Delta_{\mathrm{sd}}\boldsymbol{m}_\omega], \quad \text{(A9)}$$

where the components of the tensor $\hat{K}$ are given by

$$\hat{K}_{\alpha\beta}(\boldsymbol{q}, \omega) = v^2 \int \frac{d^2\boldsymbol{p}}{(2\pi)^2} \, \mathrm{Tr} \big\langle \sigma_\alpha G^R_{\boldsymbol{p}+\hbar q, \varepsilon+\hbar\omega} \sigma_\beta G^A_{\boldsymbol{p},\varepsilon} \big\rangle. \quad \text{(A10)}$$

Equations (A9) and (A10) correspond to Eqs. (6) and (7) of the main text. Here we used the expression for the current operator $\hat{\boldsymbol{j}} = v_f(\boldsymbol{\sigma} \times \hat{\boldsymbol{z}})$ and electric field $\boldsymbol{E}_{\boldsymbol{q},\omega} = i\omega\boldsymbol{A}_{\boldsymbol{q},\omega}$.

APPENDIX B: CALCULATION OF THE SPIN-SPIN CORRELATOR

We shall compute the matrix $\hat{M}$ to second order in powers of ω and q. The result is represented as

$$M = M_0 + M_\omega + M_{\omega^2} + M_{q\omega} + M_{q^2}, \quad \text{(B1a)}$$

$$M_0 = \frac{1}{\pi\alpha(\varepsilon^2 + \Delta_z^2)} \begin{pmatrix} \varepsilon^2 & 0 & 0 & -\varepsilon\Delta_z \\ 0 & (\varepsilon^2 - \Delta_z^2)/2 & \pi\alpha\varepsilon\Delta_z & 0 \\ 0 & -\pi\alpha\varepsilon\Delta_z & (\varepsilon^2 - \Delta_z^2)/2 & 0 \\ -\varepsilon\Delta_z & 0 & 0 & \Delta_z^2 \end{pmatrix}, \quad \text{(B1b)}$$

$$M_\omega = \frac{i\omega\varepsilon}{\left[\pi\alpha(\varepsilon^2 + \Delta_z^2)\right]^2} \begin{pmatrix} \varepsilon^2 & 0 & 0 & -\varepsilon\Delta_z \\ 0 & (\varepsilon^2 - \Delta_z^2)/2 & \pi\alpha(\varepsilon^2 - \Delta_z^2)\Delta_z/2\varepsilon & 0 \\ 0 & -\pi\alpha(\varepsilon^2 - \Delta_z^2)\Delta_z/2\varepsilon & (\varepsilon^2 - \Delta_z^2)/2 & 0 \\ -\varepsilon\Delta_z & 0 & 0 & \Delta_z^2 \end{pmatrix}, \quad \text{(B1c)}$$

$$M_{\omega^2} = \frac{(i\omega\varepsilon)^2}{\left[\pi\alpha(\varepsilon^2 + \Delta_z^2)\right]^3} \begin{pmatrix} \varepsilon^2 & 0 & 0 & -\varepsilon\Delta_z \\ 0 & (\varepsilon^2 - \Delta_z^2)/2 & \pi\alpha(\varepsilon^2 - \Delta_z^2)\Delta_z/2\varepsilon & 0 \\ 0 & -\pi\alpha(\varepsilon^2 - \Delta_z^2)\Delta_z/2\varepsilon & (\varepsilon^2 - \Delta_z^2)/2 & 0 \\ -\varepsilon\Delta_z & 0 & 0 & \Delta_z^2 \end{pmatrix}, \quad \text{(B1d)}$$

$$M_{q\omega} = \frac{v(\varepsilon^2 - \Delta_z^2)}{\left[\pi\alpha(\varepsilon^2 + \Delta_z^2)\right]^2} \left(\frac{-i}{2} + \frac{\varepsilon\omega}{\left[\pi\alpha(\varepsilon^2 + \Delta_z^2)\right]} \right) \begin{pmatrix} 0 & \varepsilon q_x & \varepsilon q_y & 0 \\ \varepsilon q_x & 0 & 0 & -\Delta_z q_x \\ \varepsilon q_y & 0 & 0 & -\Delta_z q_y \\ 0 & -\Delta_z q_x & -\Delta_z q_y & 0 \end{pmatrix}, \quad \text{(B1e)}$$

$$M_{q^2} = \frac{v^2(\varepsilon^2 - \Delta_z^2)}{2[\pi\alpha(\varepsilon^2 + \Delta_z^2)]^3} \begin{pmatrix} \varepsilon^2 q^2 & 0 & 0 & -\varepsilon\Delta_z q^2 \\ 0 & -(\varepsilon^2 - \Delta_z^2)(3q_x^2 - q_y^2)/4 & -(\varepsilon^2 - \Delta_z^2)q_x q_y/2 & 0 \\ 0 & -(\varepsilon^2 - \Delta_z^2)q_x q_y/2 & -(\varepsilon^2 - \Delta_z^2)(3q_y^2 - q_x^2)/4 & 0 \\ -\varepsilon\Delta_z q^2 & 0 & 0 & \Delta_z^2 q^2 \end{pmatrix}, \quad \text{(B1f)}$$

from which the components of $\hat{K}$ are obtained. Complete expressions for the components are cumbersome, therefore we proceed by first analyzing their denominator, which is proportional to $\det[1 - \pi\alpha M]$,

$$\det[1 - \pi\alpha M] = -\frac{\varepsilon(\varepsilon^2 + 3\Delta_z^2)^2}{4\pi\alpha(\varepsilon^2 + \Delta_z^2)^3}\left[i\omega\left(1 - i\omega\tau_{\text{tr}}\frac{\varepsilon^2 - 5\Delta_z^2}{\varepsilon^2 - \Delta_z^2} + O[(\omega\tau_{\text{tr}})^2]\right) - Dq^2\left(1 + i\omega\tau_{\text{tr}}\frac{13\Delta_z^4 + 10\Delta_z^2\varepsilon^2 + \varepsilon^4}{(\varepsilon^2 - \Delta_z^2)(\varepsilon^2 + \Delta_z^2)}\right.\right.$$

$$\left.\left. - (i\omega\tau_{\text{tr}})^2\frac{(\varepsilon^2 + 3\Delta^2)(\varepsilon^4 - 14\varepsilon^2\Delta_z - 35\Delta_z^4)}{(\varepsilon^2 - \Delta_z^2)(\varepsilon^2 + \Delta_z^2)} + O[(\omega\tau_{\text{tr}})^3]\right) + O[(Dq^2)^2\tau_{\text{tr}}]\right]. \quad \text{(B2)}$$

By restricting ourselves to perturbations that vary slowly in time compared to the transport time τ_{tr}, and smoothly in space compared to the diffusion length $L_D = \sqrt{D\tau_{\text{tr}}}$, i.e., $\omega\tau_{\text{tr}}, Dq^2\tau_{\text{tr}} \ll 1$, we are able to extract the diffusion pole $(i\omega - Dq^2)^{-1}$.

The components of the conductivity tensor $\hat{\sigma}$ at finite ω and $\boldsymbol{q}$ are given by

$$\sigma_{xx} = \sigma_0 + \frac{Dq^2}{i\omega - Dq^2}\left[\frac{q_y^2}{q^2}\sigma_0 - i\omega\tau_{\text{tr}}\left(\frac{2}{\pi\alpha}\frac{\varepsilon^2 + 2\Delta_z^2}{\varepsilon^2 + \Delta_z^2} + \frac{3}{\pi\alpha}\frac{q_x^2 - q_y^2}{2q^2}\right)\right], \quad \text{(B3a)}$$

$$\sigma_{yy} = \sigma_0 + \frac{Dq^2}{i\omega - Dq^2}\left[\frac{q_x^2}{q^2}\sigma_0 - i\omega\tau_{\text{tr}}\left(\frac{2}{\pi\alpha}\frac{\varepsilon^2 + 2\Delta_z^2}{\varepsilon^2 + \Delta_z^2} - \frac{3}{\pi\alpha}\frac{q_x^2 - q_y^2}{2q^2}\right)\right], \quad \text{(B3b)}$$

$$\sigma_{xy} = \sigma_{\text{H}} + \frac{Dq^2}{i\omega - Dq^2}\left(-\frac{q_x q_y}{q^2}\sigma_0 - i\omega\tau_{\text{tr}}\frac{3}{\pi\alpha}\frac{q_x q_y}{q^2}\right), \quad \text{(B3c)}$$

$$\sigma_{yx} = -\sigma_{\text{H}} + \frac{Dq^2}{i\omega - Dq^2}\left(-\frac{q_x q_y}{q^2}\sigma_0 - i\omega\tau_{\text{tr}}\frac{3}{\pi\alpha}\frac{q_x q_y}{q^2}\right), \quad \text{(B3d)}$$

where σ_0 and σ_{H} are given in Eq. (9) of the main text. The remaining components of $\hat{K}$ are given by

$$Q = \frac{\Delta_z}{v}\frac{iD\boldsymbol{q}}{i\omega - Dq^2}\left(1 + i\omega\tau_{\text{tr}}\frac{(\varepsilon^2 + 7\Delta_z^2)}{\varepsilon^2 + \Delta_z^2}\right), \quad \text{(B4a)}$$

$$\zeta = \frac{\Delta_z^2}{\varepsilon}\frac{1 - i\omega\tau_{\text{tr}}(\varepsilon^2 - 5\Delta_z^2)/(\varepsilon^2 - \Delta_z^2)}{i\omega - Dq^2 + \omega^2\tau_{\text{tr}}(\varepsilon^2 - 5\Delta_z^2)/(\varepsilon^2 - \Delta_z^2)}, \quad \text{(B4b)}$$

where the ω^2 term was included in the denominator of ζ because of its importance when taking the limit $q \to 0$. The leading contributions to Eq. (B4a) in the limit $\omega\tau_{\text{tr}} \ll 1$ together with Eq. (B4b) in the limit $q \to 0$ correspond to Eqs. (8), (9), and (15) of the main text.

It is convenient to rotate the coordinate system such that the new $\hat{x}$ axis lies along $\boldsymbol{q}$. Let us introduce a rotation matrix U to transform the tensor $\hat{K}$,

$$U = \begin{pmatrix} q_x/q & -q_y/q & 0 \\ q_y/q & q_x/q & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \tilde{K} = U^\top \hat{K} U, \quad \text{(B5)}$$

so that the new components of Eqs. (B3) become

$$\tilde{\sigma}_{xx} = \sigma_0 - \frac{Dq^2}{i\omega - Dq^2}i\omega\tau_{\text{tr}}\frac{7\varepsilon^2 + 11\Delta_z^2}{2\pi\alpha(\varepsilon^2 + \Delta_z^2)}, \quad \text{(B6a)}$$

$$\tilde{\sigma}_{yy} = \sigma_0 + \frac{Dq^2}{i\omega - Dq^2}\left(\sigma_0 - i\omega\tau_{\text{tr}}\frac{\varepsilon^2 + 5\Delta_z^2}{2\pi\alpha(\varepsilon^2 + \Delta_z^2)}\right), \quad \text{(B6b)}$$

$$\tilde{\sigma}_{yx} = -\tilde{\sigma}_{xy} = \sigma_{\text{H}}, \quad \text{(B6c)}$$

and the rotated tensor, $\tilde{K}$, is conveniently written as

$$\tilde{K} = \begin{pmatrix} \tilde{\sigma}_{xx} & \sigma_{\text{H}} & 0 \\ -\sigma_{\text{H}} & \tilde{\sigma}_{yy} & Q \\ 0 & Q & \zeta \end{pmatrix}. \quad \text{(B7)}$$

APPENDIX C: LIMITING BEHAVIOR OF $\boldsymbol{m}(t)$

To illustrate the behavior of $\boldsymbol{m}(t)$, we consider $f = f_0 \cos(\omega t)$ at a particular point $\boldsymbol{r}$. It is also convenient to let the field $\boldsymbol{H}_{\text{eff}}$ lie in the $\hat{x}$-$\hat{z}$ plane and rotate the coordinate system such that $\boldsymbol{H}_{\text{eff}}$ lies along the new z direction. This is achieved by introducing the rotation matrix $\hat{R}$,

$$\hat{R} = \begin{pmatrix} \cos\chi & 0 & -\sin\chi \\ 0 & 1 & 0 \\ \sin\chi & 0 & \cos\chi \end{pmatrix}, \quad \text{(C1)}$$

where χ is the angle between $\hat{\boldsymbol{z}}$ and $\boldsymbol{H}_{\text{eff}}$. Furthermore, introducing the frequency $\omega_0 = |\gamma \boldsymbol{H}_{\text{eff}}|$ and the unit vector $\bar{\boldsymbol{h}} = (-\sin\chi, 0, \cos\chi)^\top$, we can write the equation of motion in the rotated coordinate frame as

$$\partial_t \boldsymbol{m} = -\omega_0 \, \boldsymbol{m} \times \hat{\boldsymbol{z}} + f(\boldsymbol{r}, t) \, (\boldsymbol{m} \cdot \bar{\boldsymbol{h}}) \, [\boldsymbol{m} \times \bar{\boldsymbol{h}}]$$
$$+ \alpha_{\text{G}} \, [\boldsymbol{m} \times (\partial_t \boldsymbol{m})] - \alpha_{\text{G}} \, (\partial_t \boldsymbol{m}) \cdot \hat{\boldsymbol{z}} \, [\boldsymbol{m} \times \hat{\boldsymbol{z}}], \quad \text{(C2)}$$

where the vector $\hat{\boldsymbol{z}}$ is defined now as the unit vector along $\boldsymbol{H}_{\text{eff}}$, hence the magnetization projection $m_{\text{H}} = \boldsymbol{m} \cdot \boldsymbol{h}$ is simply given by m_z.

In the regime of $\alpha_{\text{G}} \ll f_0 \lll \omega_0$ we can find the asymptotic behavior of m_{H} at sufficiently small times. To do so, it is convenient to represent $\boldsymbol{m}$ in spherical coordinates: $\boldsymbol{m} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^\top$, where θ is the polar angle between $\boldsymbol{m}$ and $\hat{\boldsymbol{z}}$, and ϕ is the azimuth. In the limit $\alpha_{\text{G}} \to 0$, we find the equations of motion on θ and ϕ:

$$\partial_t \theta = \sin\chi \sin\phi \, f(\boldsymbol{r}, t)(\sin\chi \sin\theta \cos\phi - \cos\chi \cos\theta),$$
$$\text{(C3)}$$
$$\partial_t \phi = \omega_0 + f(\boldsymbol{r}, t)\cos\theta \Big[\cos^2\chi \cos^2\phi - \sin^2\phi$$
$$- \tfrac{1}{2}\sin\chi \, (\cot\theta - \sin\theta) \Big]. \quad \text{(C4)}$$

We take $f(\boldsymbol{r}, t) = f_0 \cos\omega t$ and assume that $f_0 \ll \omega_0$, so that we find $\phi = \omega_0 t - \phi_0$. It is convenient to choose $\phi_0 = \pi/2$ so that

$$\partial_t \theta = -f_0 \sin\chi \cos^2\omega_0 t \, (\sin\theta \sin\chi \sin\omega_0 t - \cos\theta \cos\chi).$$
$$\text{(C5)}$$

Because we assumed that $f_0 \ll \omega_0$, the dynamics of ϕ is much faster than the dynamics of θ. Therefore, we average Eq. (C5) over ϕ and obtain

$$\partial_t \theta = \frac{f_0}{4} \cos\theta \sin 2\chi. \quad \text{(C6)}$$

This equation is readily solved by means of the substitution $\cos\theta = 1/\cosh x$, $\sin\theta = -\tanh x$. Using the initial condition $\theta(0) = 0$, one finds

$$\cos\theta(t) = \frac{1}{\cosh\left(\tfrac{1}{4}f_0 t \sin 2\chi\right)}, \quad \text{(C7)}$$

which gives the result of Eq. (19) of the main text.

[1] I. M. Miron, G. Gaudin, S. Auffret, B. Rodmacq, A. Schuhl, S. Pizzini, J. Vogel, and P. Gambardella, Nat. Mater. **9**, 230 (2010).

[2] P. M. Haney, H.-W. Lee, K.-J. Lee, A. Manchon, and M. D. Stiles, Phys. Rev. B **87**, 174411 (2013).

[3] M. I. Dyakonov and V. I. Perel, Phys. Lett. A **35**, 459 (1971).

[4] D. Awschalom and N. Samarth, Physics **2**, 50 (2009).

[5] A. Manchon and S. Zhang, Phys. Rev. B **78**, 212405 (2008).

[6] I. Garate and A. H. MacDonald, Phys. Rev. B **80**, 134403 (2009).

[7] A. Manchon and S. Zhang, Phys. Rev. B **79**, 094422 (2009).

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

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

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

[11] M. D. Stiles and A. Zangwill, Phys. Rev. B **66**, 014407 (2002).

[12] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).

[13] J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306(R) (2007).

[14] R. Roy, Phys. Rev. B **79**, 195322 (2009).

[15] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **452**, 970 (2008).

[16] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Nat. Phys. **4**, 273 (2008).

[17] A. R. Mellnik, J. S. Lee, A. Richardella, J. L. Grab, P. J. Mintun, M. H. Fischer, A. Vaezi, A. Manchon, E.-A. Kim, N. Samarth, and D. C. Ralph, Nature (London) **511**, 449 (2014).

[18] Y. Wang, P. Deorani, K. Banerjee, N. Koirala, M. Brahlek, S. Oh, and H. Yang, Phys. Rev. Lett. **114**, 257202 (2015).

[19] Y. Fan, P. Upadhyaya, X. Kou, M. Lang, S. Takei, Z. Wang, J. Tang, L. He, L.-T. Chang, M. Montazeri, G. Yu, W. Jiang, T. Nie, R. N. Schwartz, Y. Tserkovnyak, and K. L. Wang, Nat. Mater. **13**, 699 (2014).

[20] Y. Fan, X. Kou, P. Upadhyaya, Q. Shao, L. Pan, M. Lang, X. Che, J. Tang, M. Montazeri, K. Murata, L.-T. Chang, M. Akyol, G. Yu, T. Nie, K. L. Wong, J. Liu, Y. Wang, Y. Tserkovnyak, and K. L. Wang, Nat. Nanotechnol. **11**, 352 (2016).

[21] K. Yasuda, A. Tsukazaki, R. Yoshimi, K. Kondou, K. S. Takahashi, Y. Otani, M. Kawasaki, and Y. Tokura, Phys. Rev. Lett. **119**, 137204 (2017).

[22] S. Cha, M. Noh, J. Kim, J. Son, H. Bae, D. Lee, H. Kim, J. Lee, H.-S. Shin, S. Sim, S. Yang, S. Lee, W. Shim, C.-H. Lee, M.-H. Jo, J. S. Kim, D. Kim, and H. Choi, Nat. Nanotechnol. **13**, 910 (2018).

[23] L. Liu, T. Moriyama, D. C. Ralph, and R. A. Buhrman, Phys. Rev. Lett. **106**, 036601 (2011).

[24] M. Akyol, G. Yu, J. G. Alzate, P. Upadhyaya, X. Li, K. L. Wong, A. Ekicibil, P. Khalili Amiri, and K. L. Wang, Appl. Phys. Lett. **106**, 162409 (2015).

[25] D. MacNeill, G. M. Stiehl, M. H. D. Guimaraes, R. A. Buhrman, J. Park, and D. C. Ralph, Nat. Phys. **13**, 300 (2016).

[26] M. H. D. Guimaraes, G. M. Stiehl, D. MacNeill, N. D. Reynolds, and D. C. Ralph, Nano Lett. **18**, 1311 (2018).

[27] A. Sakai and H. Kohno, Phys. Rev. B **89**, 165307 (2014).

[28] P. B. Ndiaye, C. A. Akosa, M. H. Fischer, A. Vaezi, E.-A. Kim, and A. Manchon, Phys. Rev. B **96**, 014408 (2017).

[29] A. A. Burkov and D. G. Hawthorn, Phys. Rev. Lett. **105**, 066802 (2010).

[30] K. Taguchi, K. Shintani, and Y. Tanaka, Phys. Rev. B **92**, 035425 (2015).

[31] K. Shintani, K. Taguchi, Y. Tanaka, and Y. Kawaguchi, Phys. Rev. B **93**, 195415 (2016).

[32] S. V. Vonsovsky, *Magnetism* (Wiley, New York, 1974).

[33] S. Ghosh and A. Manchon, Phys. Rev. B **97**, 134402 (2018).

[34] S. Maleev and B. Toperverg, Zh. Eksp. Teor. Fiz. **69**, 1440 (1976) [Sov. J. Exp. Theor. Phys. **42**, 734 (1976)].

[35] B. Altshuler and A. G. Aronov, *Electron-electron Interaction in Disordered Systems* (North-Holland, Amsterdam, 1985), Vol. 1, p. 155.

[36] I. A. Ado, I. A. Dmitriev, P. M. Ostrovsky, and M. Titov, Europhys. Lett. **111**, 37004 (2015).

[37] I. A. Ado, I. A. Dmitriev, P. M. Ostrovsky, and M. Titov, Phys. Rev. Lett. **117**, 046601 (2016).

[38] I. A. Ado, I. A. Dmitriev, P. M. Ostrovsky, and M. Titov, Phys. Rev. B **96**, 235148 (2017).

[39] D. Kong, J. J. Cha, K. Lai, H. Peng, J. G. Analytis, S. Meister, Y. Chen, H.-J. Zhang, I. R. Fisher, Z.-X. Shen, and Y. Cui, ACS Nano **5**, 4698 (2011).

[40] V. S. Kamboj, A. Singh, T. Ferrus, H. E. Beere, L. B. Duffy, T. Hesjedal, C. H. W. Barnes, and D. A. Ritchie, ACS Photon. **4**, 2711 (2017).

[41] S.-M. Huang, S.-J. Huang, C. Hsu, P. V. Wadekar, Y.-J. Yan, S.-H. Yu, and M. Chou, Sci. Rep. **7**, 5133 (2017).

[42] F. X. Xiang, X. L. Wang, and S. X. Dou, arXiv:1401.6732.

[43] E. van der Bijl and R. A. Duine, Phys. Rev. B **86**, 094406 (2012).

[44] K. M. D. Hals and A. Brataas, Phys. Rev. B **88**, 085423 (2013).

[45] Y. Tserkovnyak and C. H. Wong, Phys. Rev. B **79**, 014402 (2009).

[46] F. Mahfouzi, N. Nagaosa, and B. K. Nikolić, Phys. Rev. Lett. **109**, 166602 (2012).

[47] Y. Ferreiros, F. J. Buijnsters, and M. I. Katsnelson, Phys. Rev. B **92**, 085416 (2015).

[48] M. H. Fischer, A. Vaezi, A. Manchon, and E.-A. Kim, Phys. Rev. B **93**, 125303 (2016).

[49] T. Yokoyama, J. Zang, and N. Nagaosa, Phys. Rev. B **81**, 241410 (2010).

[50] T. Yokoyama, Phys. Rev. B **84**, 113407 (2011).

[51] Z. B. Siu, H. C. Son, M. b. A. Jalil, and S. G. Tan, J. Appl. Phys. **122**, 073903 (2017).

[52] F. Mahfouzi, B. K. Nikolić, and N. Kioussis, Phys. Rev. B **93**, 115419 (2016).

[53] M. Soleimani, S. Jalili, F. Mahfouzi, and N. Kioussis, Europhys. Lett. **117**, 37001 (2017).

[54] D. Kurebayashi and K. Nomura, arXiv:1702.04918.

[55] J. Chen, Y. Peng, and J. Zhou, J. Magn. Magn. Mater. **432**, 554 (2017).

[56] M. Rodriguez-Vega, G. Schwiete, J. Sinova, and E. Rossi, Phys. Rev. B **96**, 235419 (2017).

[57] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Phys. Rev. B **78**, 195424 (2008).

[58] I. Garate and M. Franz, Phys. Rev. Lett. **104**, 146802 (2010).

[59] K. Nomura and N. Nagaosa, Phys. Rev. B **82**, 161401(R) (2010).

[60] Y. Tserkovnyak and D. Loss, Phys. Rev. Lett. **108**, 187201 (2012).

[61] J. Linder, Phys. Rev. B **90**, 041412(R) (2014).

[62] Y. Tserkovnyak, D. A. Pesin, and D. Loss, Phys. Rev. B **91**, 041121(R) (2015).

[63] H. T. Ueda, A. Takeuchi, G. Tatara, and T. Yokoyama, Phys. Rev. B **85**, 115110 (2012).

[64] X. Liu and J. Sinova, Phys. Rev. Lett. **111**, 166801 (2013).

[65] P.-H. Chang, T. Markussen, S. Smidstrup, K. Stokbro, and B. K. Nikolić, Phys. Rev. B **92**, 201406(R) (2015).

[66] J. Fujimoto and H. Kohno, Phys. Rev. B **90**, 214418 (2014).

[67] N. Okuma and M. Ogata, Phys. Rev. B **93**, 140205(R) (2016).

[68] J. Rammer and H. Smith, Rev. Mod. Phys. **58**, 323 (1986).