

Spin Aharonov-Bohm effect and topological spin transistor

Joseph Maciejko,^{1,2} Eun-Ah Kim,³ and Xiao-Liang Qi^{1,2}¹*Department of Physics, Stanford University, Stanford, California 94305, USA*²*Stanford Institute for Materials and Energy Sciences, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA*³*Department of Physics, Cornell University, Ithaca, New York 14853, USA*

(Received 25 September 2009; revised manuscript received 19 August 2010; published 4 November 2010)

Ever since its discovery, the electron spin has only been measured or manipulated through the application of an electromagnetic force acting on the associated magnetic moment. In this work, we propose a spin Aharonov-Bohm effect in which the electron spin is controlled by a magnetic flux while no electromagnetic field is acting on the electron. Such a nonlocal spin manipulation is realized in an Aharonov-Bohm ring made from the recently discovered quantum spin Hall insulator, by taking advantage of the defining property of the quantum spin Hall edge states: the one-to-one correspondence between spin polarization and direction of propagation. The proposed setup can be used to realize a new spintronics device, the topological spin transistor, in which the spin rotation is completely controlled by a magnetic flux of $hc/2e$, independently of the details of the sample.

DOI: [10.1103/PhysRevB.82.195409](https://doi.org/10.1103/PhysRevB.82.195409)

PACS number(s): 72.25.Dc, 73.43.-f, 73.23.Ad

I. INTRODUCTION

The spin of the electron is one of the most fundamental quantum-mechanical degrees of freedom in nature. Historically, the discovery of the electron spin helped to lay the foundation of relativistic quantum mechanics. In recent years, the electron spin has been proposed as a possible alternate state variable for the next generation of computers, which led to extensive efforts toward achieving control and manipulation of the electron spin, a field known as spintronics.¹ Despite the great variety of currently used or theoretically proposed means of manipulating the electron spin, a feature common to all of them is that they all make use of the *classical* electromagnetic force or torque acting *locally* on the magnetic moment associated with the spin.

On the other hand, it is known that due to the Aharonov-Bohm (AB) effect,² electrons in a ring can be affected in a purely *quantum-mechanical* and *nonlocal* way by the flux enclosed by the ring even though no magnetic field—hence no classical force—is acting on them. This effect could be termed “charge AB effect,” as it relies only on the electron carrying an electric charge. In systems with spin-orbit coupling or magnetic fields, a spin-dependent phase factor can be obtained and leads to modifications to the AB effect.³ However, these effects usually involve classical forces acting on the spin such as electromagnetic fields, and the pure gauge potential leading to the charge AB effect does not directly couple to spin. This observation leads naturally to the question of whether it is possible to observe a “spin AB effect” which would enable one to manipulate the electron spin in a purely nonlocal and quantum-mechanical way, without any classical force or torque acting locally on the spin magnetic moment.

In this work, we show that the spin AB effect is indeed possible by making use of the edge states of the recently discovered quantum spin Hall (QSH) insulators. In recent years, the QSH insulator state has been proposed in several different materials.^{4–9} In particular, this topologically nontrivial state of matter has been recently predicted⁷ and realized experimentally^{10–12} in HgTe quantum wells (QWs). The

QSH insulator is invariant under time reversal (TR), has a charge excitation gap in the bulk, but has topologically protected gapless edge states that lie inside the bulk insulating gap. These edge states have a distinct helical property: two states with opposite spin polarization counterpropagate at a given edge.^{13–15} The edge states come in Kramers doublets, and TR symmetry ensures the crossing of their energy levels at TR invariant points in the Brillouin zone. Because of this level crossing, the spectrum of a QSH insulator cannot be adiabatically deformed into that of a topologically trivial insulator without closing the bulk gap. The helicity of the QSH edge states is the decisive property which allows the spin AB effect to exist: the perfect correlation between spin orientation and direction of propagation allows the transmutation of a usual charge AB effect into a spin AB effect, as will be explained in detail below.

The mechanism we propose to realize the spin AB effect is illustrated in Fig. 1. Consider a two-terminal device consisting of a bounded QSH insulator region pierced by a hole which is threaded by a magnetic flux ϕ . If the edge electrons propagating clockwise have their spin pointing out-of-plane along z (spin up $|\uparrow\rangle$) due to TR symmetry the electrons propagating counterclockwise must have opposite spin along $-z$ (spin down $|\downarrow\rangle$). If we inject electrons spin polarized along the x direction $|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$ from a ferromagnetic (FM) lead on the left, the electron beam will be split coherently upon entering the QSH region at the left junction into a $|\uparrow\rangle$ beam propagating along the top edge and a $|\downarrow\rangle$ beam propagating along the bottom edge. When the electron beams are recombined on the right side of the ring, the electrons along top and bottom edges will acquire a phase difference of $\varphi = 2\pi\phi/\phi_0$ due to the AB effect, where $\phi_0 = hc/e$ is the flux quantum. Consequently, the output state is given by $\frac{1}{\sqrt{2}}(|\uparrow\rangle + e^{-i\varphi}|\downarrow\rangle)$, such that the electron spin is rotated by an angle φ in the xy plane. The magnetic flux being confined to the hole in the device (Fig. 1), the electromagnetic fields are zero in the region where the electrons propagate, and the spin is rotated by a purely quantum-mechanical Berry phase effect. In particular, for collinear FM leads ($\theta=0$ in Fig. 1), one expects the conductance to be maximal for $\phi=0(\text{mod } \phi_0)$

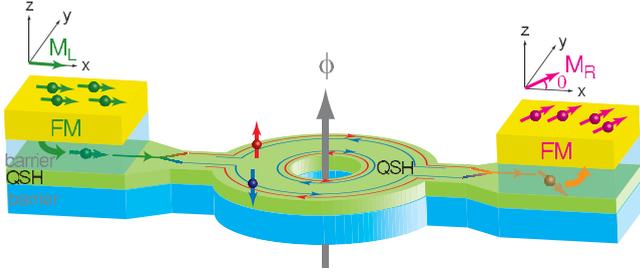


FIG. 1. (Color online) Schematic picture of the spin AB effect. A ring of QSH insulator threaded by a magnetic flux ϕ is connected to two magnetic leads. Spin-polarized electrons injected from the left lead enter the QSH region as a superposition of spin-up and spin-down states. The spin-up (down) state can only propagate along the top (bottom) edge of the QSH ring, and the two spin states thus acquire an AB phase difference proportional to ϕ . Consequently, upon exiting the QSH region the two edge states recombine into a state with spin rotated with respect to the injected direction. The magnetization direction of the right lead generally differs from that of the left lead by an angle θ . The two-terminal conductance $G=G(\phi, \theta)$ of the device depends on the relative angle between the spin polarization of the outgoing state and that of the right lead.

and minimal for $\phi = \frac{1}{2}\phi_0 \pmod{\phi_0}$, thus realizing a “topological” spin transistor [Fig. 3(c)]. This effect is topological in the sense that the spin is always rotated by one cycle for each period of flux ϕ_0 , regardless of the details of the device, such as the size of the system or the shape of the ring.

II. PHENOMENOLOGICAL SCATTERING MATRIX ANALYSIS

Before considering any microscopic model of transport in a QSH system, generic features of two-terminal transport in the device of Fig. 1 that depend only on symmetry considerations can be extracted from a simple phenomenological scattering matrix or S -matrix analysis.¹⁶ The left and right junctions are each described by a scattering matrix S_L and S_R , respectively [e.g., Fig. 2(a) for the left junction]. Considering the left junction first, S_L consists of four submatrices t_L, t'_L, r_L, r'_L which correspond, respectively, to transmission from left to right, transmission from right to left, reflection from the left, and reflection from the right. One can define similar submatrices for S_R . We wish to obtain an effective S -matrix S [see Eq. (A5)] for the whole device, by combining the S matrices of the junctions together with the S matrix for the central QSH region. Inside the QSH region, the AB effect is described by the matrix $\Phi \equiv e^{-i\varphi\sigma_z/2}$, where $\sigma_x, \sigma_y, \sigma_z$ are the three Pauli matrices. In addition to the geometric phase φ , the edge electrons also acquire a dynamical phase $\lambda = 2k_F\ell$ identical for both spin polarizations, where ℓ is the distance traveled by the edge electrons from left to right junction and k_F is the edge-state Fermi wave vector. Details of the analysis are presented in Appendix A; here we discuss only the main results. We obtain the effective 2×2 device scattering matrix S ,

$$S(\phi, \theta) = [1 - e^{i\lambda}\Phi r'_L(0)\Phi r_R(\theta)]^{-1} e^{i\lambda/2}\Phi, \tag{1}$$

where the junction reflection matrices $r'_L(\theta_L)$ and $r_R(\theta_R)$ depend on the angles θ_L, θ_R of the magnetization $\mathbf{M}_{L,R}$ in the

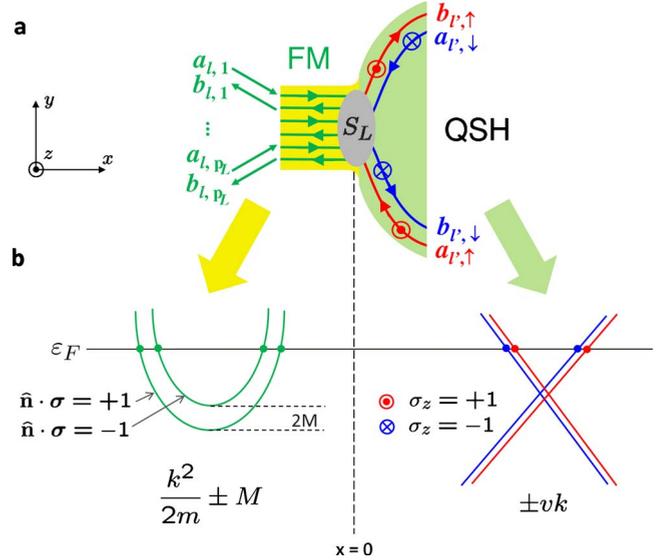


FIG. 2. (Color online) Illustration of the minimal model describing a FM/QSH junction. (a) Schematic picture of the junction between the left FM lead and the QSH insulator. Incoming channels $a_{l,1}, \dots, a_{l,p_L}$ from the left lead scatter at the junction into transmitted QSH edge channels $b_{l,\uparrow}, b_{l,\downarrow}$ and reflected lead channels $b_{l,1}, \dots, b_{l,p_L}$. This scattering process is described by a scattering matrix S_L . (b) Minimal model description of the junction. The FM lead is described by 1D parabolic bands with a spin splitting $2M$ while the QSH edge states are linearly dispersing and TR invariant, with opposite spin states counterpropagating.

left and right leads. For simplicity we consider $\theta_L=0$ and define $\theta \equiv \theta_R$ (Fig. 1).

The two-terminal conductance G of the device can be written as

$$G = \frac{e^2}{h} \text{tr } \rho_R S \rho_L S^\dagger, \tag{2}$$

using Eq. (A5) of Appendix A. Here ρ_L, ρ_R are 2×2 effective spin-density matrices for the FM leads, and have the form

$$\rho_\alpha(\theta_\alpha) = \frac{1}{2} T_\alpha(\theta_\alpha) [1 + \mathbf{P}_\alpha(\theta_\alpha) \cdot \boldsymbol{\sigma}] \tag{3}$$

with $\alpha=L,R$, where $T_\alpha = \text{tr } \rho_\alpha$ is the transmission coefficient of the junction and \mathbf{P}_α is a polarization vector. For simplicity, we can assume the device to have a π -rotation symmetry, which together with TR symmetry restricts the generic form of the reflection matrices r'_L and r_R in Eq. (1) to be

$$r'_L(\theta) = \begin{pmatrix} \alpha_\theta & \beta_\theta \\ \gamma_\theta & \alpha_{\theta+\pi} \end{pmatrix}, \quad r_R(\theta) = \begin{pmatrix} \alpha_{\theta+\pi} & \beta_\theta \\ \gamma_\theta & \alpha_\theta \end{pmatrix}. \tag{4}$$

Physically, α_θ is a nonspin-flip reflection amplitude whereas $\beta_\theta, \gamma_\theta$ are spin-flip reflection amplitudes, with β_θ corresponding to a $|\downarrow\rangle \rightarrow |\uparrow\rangle$ reflection and γ_θ to a $|\uparrow\rangle \rightarrow |\downarrow\rangle$ reflection. These amplitudes are generally different due to the breaking of TR symmetry at the junctions by the nearby FM leads.

III. MINIMAL MODEL DESCRIPTION

These expressions being so far very general, to make further progress it is useful to consider a simple continuum Hamiltonian model for the FM/QSH junctions in which the reflection matrices r'_L and r_R can be calculated explicitly. This model satisfies the symmetries invoked earlier and will be seen to be a good description of the realistic HgTe system in spite of its simplicity. We model the FM leads as one-dimensional (1D) spin- $\frac{1}{2}$ fermions with a term which explicitly breaks the $SU(2)$ spin-rotation symmetry,¹⁷

$$H_{\text{FM}} = \int dx \Psi^\dagger \left[-\frac{1}{2m} \frac{\partial^2}{\partial x^2} - \mathbf{M}(\theta) \cdot \boldsymbol{\sigma} \right] \Psi,$$

where $\mathbf{M}(\theta) = M\hat{\mathbf{n}}$, with $\hat{\mathbf{n}} = \hat{\mathbf{x}} \cos \theta + \hat{\mathbf{y}} \sin \theta$, is an in-plane magnetization vector and Ψ is a two-component spinor $\Psi \equiv (\psi_\uparrow \ \psi_\downarrow)^T$. In the absence of AB flux, the QSH edge liquid consists of 1D massless helical fermions.^{14,15} When the spins of the edge states are polarized along the z direction, the Hamiltonian is given by

$$H_{\text{QSH}} = -iv \sum_{\alpha=t,b} \eta_\alpha \int dx (\psi_{\alpha\uparrow}^\dagger \partial_x \psi_{\alpha\uparrow} - \psi_{\alpha\downarrow}^\dagger \partial_x \psi_{\alpha\downarrow}),$$

where v is the edge-state velocity and $\alpha=t, b$ refers to the top and bottom edge, respectively, with $\eta_t=1$ and $\eta_b=-1$.

In this simple model, the junction is described as a sharp interface between the FM region and the QSH region, from which the reflection matrix r'_L in Eq. (1) and the spin density matrix ρ_L in Eq. (2) can be obtained. The calculation yields the reflection matrices precisely in the form of Eq. (4) with $\alpha_\theta = a$ and $\beta_\theta = \gamma_\theta^* = be^{-i\theta}$. In the limit of small spin splitting $M/\varepsilon_F \ll 1$, where ε_F is the Fermi energy in the leads [Fig. 2(b)], the reflection amplitudes a and b are given by

$$a \simeq \frac{v - v_F}{v + v_F}, \quad b \simeq \frac{M}{2\varepsilon_F} \frac{v_F^3}{v(v + v_F)^2}, \quad (5)$$

where $v_F = \sqrt{2\varepsilon_F/m}$ is the Fermi velocity in the FM leads. The off-diagonal spin-flip reflection amplitude b is proportional to the magnetization M and along with its accompanying scattering phase shift $e^{\pm i\theta}$ is an explicit signature of TR symmetry breaking at the junction. The diagonal nonspin-flip reflection amplitude a does not break TR symmetry and is the same as would be obtained in the scattering from a nonmagnetic metal with $\mathbf{M}=0$. The lead spin-density matrices ρ_L, ρ_R can also be calculated explicitly and are found to follow the form of Eq. (3) as expected from the general S -matrix analysis. In the limit $M/\varepsilon_F \ll 1$, we obtain $T_L = T_R = 8v v_F / (v + v_F)^2$ and

$$\mathbf{P}_L(\theta) = \mathbf{P}_R(\theta) \equiv \mathbf{P}(\theta) = -\frac{\mathbf{M}(\theta)}{4\varepsilon_F} \frac{v_F^2}{v(v + v_F)}, \quad (6)$$

i.e., the spin-polarization vector is directly proportional to the magnetization \mathbf{M} .

From the results obtained above, we can readily evaluate the conductance G , which has the following expression in the limit $M/\varepsilon_F, P \equiv |\mathbf{P}(\theta)| \ll 1$ and $\lambda=0$:

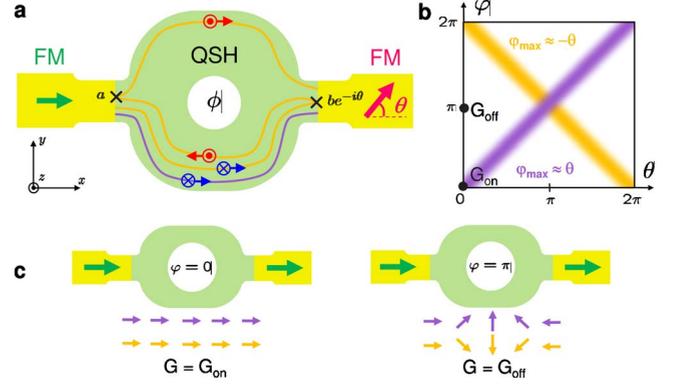


FIG. 3. (Color online) Phenomenological analysis of the two-terminal conductance (top view of Fig. 1). (a) The two leading contributions to the spin AB rotation. The darker path stands for the process with no spin flips, which leads to a spin rotation of $\varphi \equiv 2\pi\phi/\phi_0$. The lighter path stands for the process with spin-dependent reflections, which leads to a spin rotation of $-\varphi$. (b) Schematic intensity map of the two-terminal conductance $G(\varphi, \theta)$. The conductance reaches its maximum along the lines $\varphi = \theta$ (darker line) and $\varphi = -\theta$ (lighter line), which are contributed by the darker and lighter paths in panel (a), respectively. (c) The on and off states of the topological spin transistor are defined for $\theta=0$ by $\varphi=0$ and $\varphi=\pi$, respectively, as also indicated in panel (c).

$$G(\varphi, \theta; \lambda = 0) = \frac{e^2}{h} \frac{T_L T_R / 2}{1 - 2a^2 \cos \varphi + a^4} \left[1 + \frac{\cos(\theta - \varphi) + (1 - t^2)^2 \cos(\theta + \varphi) + C(\varphi, \theta)}{1 - 2a^2 \cos \varphi + a^4} \times P^2 + \mathcal{O}(P^4) \right], \quad (7)$$

where $t=1-a$ and $C(\varphi, \theta) \equiv \gamma \cos \varphi + \delta \cos \theta$ with γ, δ some constants depending only on a . The effect of a finite λ will be addressed in the next section, where we study numerically a more realistic model of the QSH state in HgTe QWs. Physically, a and t can be interpreted as reflection and transmission coefficients for the S_z spin current. The generic behavior of Eq. (7) is illustrated in Fig. 3. The term $C(\varphi, \theta)$ is an uninteresting background term which manifests no correlation between AB phase φ and rotation angle of the electron spin θ . The term $\propto \cos(\theta - \varphi)$ corresponds to a rotation of the electron spin by φ , and the term $\propto \cos(\theta + \varphi)$ corresponds to a rotation by $-\varphi$. The conductance is thus maximal for $\varphi_{\text{max}} = \pm \theta$ [Fig. 3(b)], manifesting the desired flux-induced spin-rotation effect. Physically, the $\varphi_{\text{max}} = \theta$ term corresponds to a process in which electrons traverse the device without undergoing spin flips [Fig. 3(a), darker trajectory] while the $\varphi_{\text{max}} = -\theta$ term corresponds to a process involving at least one TR breaking spin-flip reflection [Fig. 3(a), lighter trajectory]. As can be seen from Eq. (7), the relative intensity of the two contributions to the conductance is $I_{-\theta}/I_\theta = (1-t^2)^2$ which can be close to unity for strongly reflecting junctions $t \ll 1$. As both contributions are minimal for $\varphi = \pi$ at $\theta = 0$, one can consider $\varphi = \pi, \theta = 0$ as the “off” state of a spin transistor [Fig. 3(c), right] where the rotation

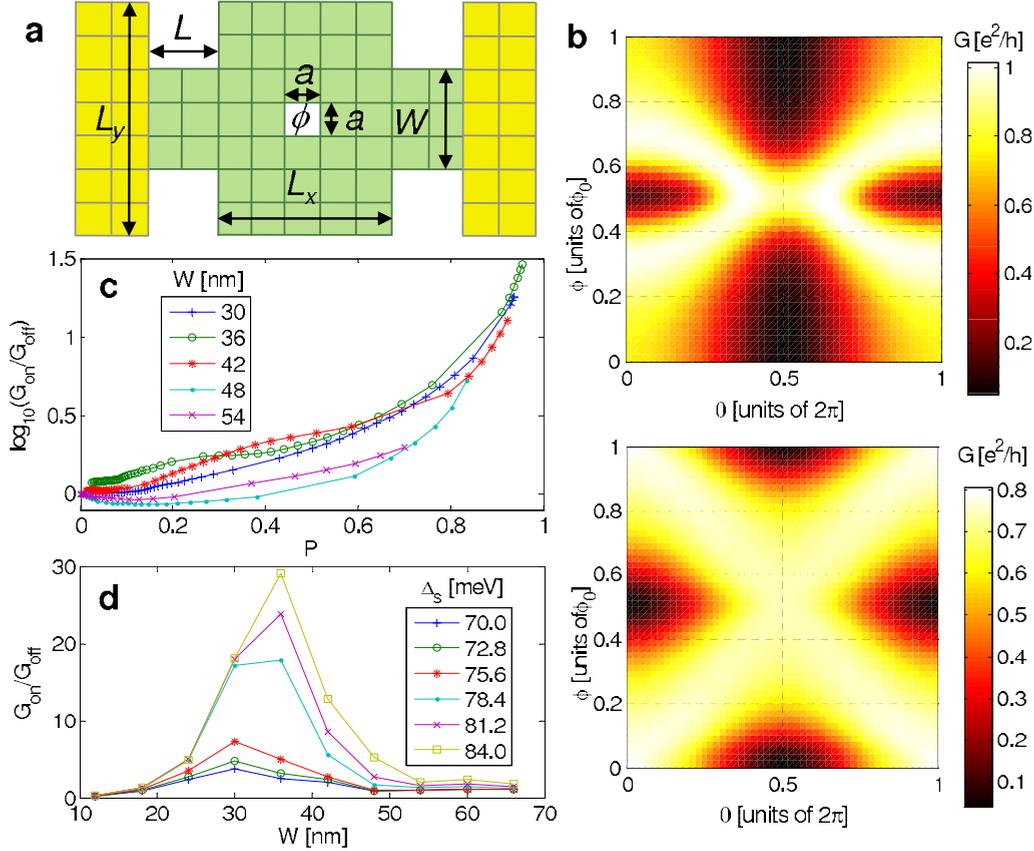


FIG. 4. (Color online) Numerical study of the spin AB effect in HgTe QWs. (a) Device geometry used for the numerical two-terminal conductance calculation: $a=30 \text{ \AA}$ is the lattice constant of the tight-binding model, $L=18 \text{ nm}$, $L_x=\ell=240 \text{ nm}$, $L_y=120 \text{ nm}$, ϕ is the AB flux, and W is the QPC width. (b) Intensity map of the conductance $G(\phi, \theta)$ for fixed chemical potential $\mu=0.06 \text{ eV}$ (top panel) and averaged chemical potential over energy range $\Delta\mu=5 \text{ meV}$ corresponding to an average over $\sim 2\pi$ dynamical phase (bottom panel). These two situations correspond to low and high temperature, respectively (see text). (c) Logarithmic plot of on/off ratio $G_{\text{on}}/G_{\text{off}}$ of topological spin transistor as a function of spin polarization P of injected carriers for fixed chemical potential $\mu=0.06 \text{ eV}$ and different values of the QPC width W . (d) Plot of on/off ratio as a function of QPC width W for fixed chemical potential $\mu=0.06 \text{ eV}$ and different values of the spin splitting Δ_s in the bulk leads.

of the spin is provided by a purely quantum-mechanical Berry phase effect. This is in contrast with the famous Datta-Das spin transistor¹⁸ where the rotation of the spin is achieved through the classical spin-orbit force. The “on” state corresponds to the absence of spin rotation for $\varphi=0$ [Fig. 3(c), left].

IV. EXPERIMENTAL REALIZATION IN HgTe QUANTUM WELLS

We now show that this proposal can, in principle, be realized experimentally in HgTe QWs. We model the device of Fig. 1 as a rectangular QSH region threaded by a magnetic AB flux through a single plaquette in the center, and connected to semi-infinite metallic leads on both sides by rectangular QSH constrictions modeling quantum point contacts (QPCs) [Fig. 4(a)]. The QSH region is described by an effective 4×4 tight-binding Hamiltonian^{7,19} with the chemical potential in the bulk gap, while the metallic leads are described by the same model with the chemical potential in the conduction band. The detailed form of the model is given in

Appendix C. The injection of spin-polarized carriers by the FM layers of Fig. 1 is mimicked by the inclusion of an effective Zeeman term in the Hamiltonian of the semi-infinite leads. We calculate numerically the two-terminal conductance through the device of Fig. 4(a) for a QW thickness $d=80 \text{ \AA}$. We use the standard lattice Green’s-function Landauer-Büttiker approach²⁰ in which the conductance is obtained from the Green’s function of the whole device, the latter being calculated recursively.²¹

The results of the numerical calculation are plotted in Figs. 4(b)–4(d). In the absence of phase-breaking scattering processes, one distinguishes two temperatures regimes $T \ll T_\ell$ and $T \gg T_\ell$ separated by a crossover temperature $T_\ell = \pi\hbar v/k_B\ell$ with v the edge-state velocity, defined as the temperature for which a thermal spread $\Delta\mu \sim k_B T$ in the energy distribution of injected electrons corresponds to a spread in the distribution of dynamical phases $\lambda = 2k_F\ell$ of $\Delta\lambda \sim 2\pi$. In the low-temperature regime $T \ll T_\ell$, $\Delta\lambda \ll 2\pi$ and the dynamical phase is essentially fixed such that $G(T \ll T_\ell) \approx G(T=0)$. In this regime, $G(T=0, \mu)$ is approximately periodic in μ for μ within the bulk gap, with period $\Delta\mu \sim k_B T_\ell$. A crossing pattern [Fig. 4(b), top] occurs periodically

and can be obtained by tuning the chemical potential. It corresponds to the flux-induced spin-rotation effect (Fig. 3). In the high-temperature regime $T \gg T_\ell$, one could expect that the crossing pattern, and thus the spin-rotation effect, would be washed out by thermal self-averaging of the dynamical phase. Surprisingly, the pattern remains [Fig. 4(b), bottom], and actually acquires a more symmetric structure through the self-averaging procedure. In both temperature regimes, the conductance pattern agrees qualitatively with the result of the simple 1D Hamiltonian model [Fig. 3(b)].

So far, our discussion has ignored the existence of phase-breaking processes. Such processes introduce an additional characteristic temperature T_φ , defined as the temperature above which the phase coherence length $\ell_\varphi(T)$ becomes smaller than the system size ℓ , that is $\ell_\varphi(T_\varphi) = \ell$ and $\ell_\varphi(T > T_\varphi) < \ell$. As explained in Sec. I, the stability of the QSH state is protected by Kramers' theorem. However, Kramers' theorem requires the quantum phase coherence of electronic wave functions, hence for $T > T_\varphi$ the QSH state can be destroyed.¹⁰⁻¹² Thus, the observation of the spin AB effect requires $T < T_\varphi$, with the precise value of T_φ depending on the particular nature of the phase-breaking mechanisms. With this first requirement satisfied, two scenarios are possible depending on the relative value of the two characteristic temperatures T_ℓ and T_φ . If $T_\ell < T_\varphi$, the scenario described in the previous paragraph applies, with the existence of a low-temperature regime $T \ll T_\ell < T_\varphi$ with well-defined dynamical phase and a high-temperature regime $T_\ell \ll T < T_\varphi$ with completely randomized dynamical phase. On the other hand, if $T_\ell > T_\varphi$, then since we require $T < T_\varphi$ for the observation of the spin AB effect the high-temperature regime $T \gg T_\ell$ can be never be achieved. This could correspond, for instance, to a very small, fully phase-coherent device $\ell \ll \ell_\varphi$, with no noticeable thermal fluctuation effects. In transport measurements on HgTe QWs (Refs. 10, 11, and 22) a robust QSH state has been observed in devices of size $\ell \approx 1 \mu\text{m}$ up to temperatures of 4.2 K. This gives us a lower bound estimate of a few kelvin for T_φ , for a device of such size. For a typical edge-state velocity $\hbar v \sim 3.5 \text{ eV \AA}$ one obtains a crossover temperature $T_\ell \sim 13 \text{ K} \geq T_\varphi$, which indicates that one would probably be in the low-temperature regime with weak thermal fluctuation effects and good tunability of the crossing pattern with chemical potential μ . The other scenario requiring $T_\ell < T_\varphi$ can be realized if the lower bound estimate of 4.2 K turns out to be too conservative and we actually have $T_\varphi > 13 \text{ K}$, or if the edge-state velocity is significantly smaller than the value of 3.5 eV \AA used above. The latter possibility can occur in type-II QWs (Ref. 8) where the edge state velocity is about one order of magnitude smaller, hence $T_\ell \sim 1 \text{ K}$ and the condition $T_\ell < T_\varphi$ would, in principle, be satisfied.

In our calculations, for simplicity we have assumed that electrons on both the top and bottom edges acquire the same dynamical phase λ . In a real system, the two arms of the ring are not perfectly symmetric and the electrons propagating on different arms can certainly acquire different dynamical phases $\lambda_{\text{bottom}} \neq \lambda_{\text{top}}$. However, the dynamical phase difference $\delta \equiv \lambda_{\text{bottom}} - \lambda_{\text{top}}$ only leads to an additional flux-independent rotation of the spin of the outgoing electrons, which leads to a shift of the conductance pattern in the angle

θ by an amount δ [see Eq. (A6) of Appendix A]. Thus the transistor remains effective if one uses $\theta = \delta$ instead of $\theta = 0$ in the right FM lead. If one prefers to use $\theta = 0$, one can cancel out the phase asymmetry by patterning an electrostatic gate on top of one given arm. By tuning the potential of this gate, one can adjust the Fermi wave vector locally and introduce a dynamical phase offset which cancels out the phase asymmetry δ .

In Figs. 4(c) and 4(d) we plot the on/off ratio $G_{\text{on}}/G_{\text{off}}$ of the topological spin transistor, which can be taken as the figure of merit of the device. We define $G_{\text{on}} \equiv G(\phi=0, \theta=0)$ and $G_{\text{off}} \equiv G(\phi=\frac{1}{2}\phi_0, \theta=0)$ [Fig. 3(c)]. We use two parameters, the junction spin polarization P and the bulk spin splitting Δ_s to quantify the degree of spin polarization of the injected carriers. An actual experimental implementation of the transistor concept described here will require optimization of these or similar parameters. The junction spin polarization P is obtained for a given junction geometry, i.e., a given choice of QPC width W and length L [Fig. 4(a)], by calculating the transfer matrix²³ of the junction directly from the TB model and using Eq. (3) with $P \equiv |\mathbf{P}|$. The spin splitting Δ_s is obtained from the continuum $\mathbf{k} \cdot \mathbf{p}$ HgTe QW Hamiltonian mentioned earlier, and is defined as the energy difference between "spin-up" ($E1+$) and "spin-down" ($E1-$) energy levels¹⁹ at the Γ point. The on/off ratio increases rapidly for a polarization P of order unity [Fig. 4(c)]. It is reasonable to expect that optimized junction designs, better than the simplistic proof-of-concept geometry used here, would yield even higher on/off ratios. There is also an optimal width $W_{\text{opt}} \approx 0.29L_y$ for the junction QPC [Fig. 4(d)]. For $W < W_{\text{opt}}$, interedge tunneling²⁴ strongly backscatters the incoming electrons and reduces G_{on} , which suppresses the on/off ratio. For $W > W_{\text{opt}}$, the edge states on opposite edges are too far apart to recombine coherently and to produce the desired spin-rotation effect, which increases G_{off} and also suppresses the on/off ratio. In our calculation, we did not take into account the possible structural inversion asymmetry (SIA) which induces Rashba spin-orbit coupling in the QW.²⁵ However, it should be noticed that the usual contributions of SIA to the AB effect, such as the Aharonov-Casher effect,²⁶ are absent because there are no two-dimensional bulk carriers in the QSH state. Since the only conducting channels in the QSH state are the 1D edge states, the only effect of SIA is some global rotation of the edge-state spin direction. The topological spin rotation induced by half of a flux quantum is simply a consequence of the spatial separation of opposite spins on opposite edges, which is determined by the topological properties of the QSH state and thus remains robust.

V. CONCLUSION AND OUTLOOK

In this work, we have shown the possibility of using a topologically nontrivial state of matter, the QSH insulator state, to manipulate the spin of the electron by purely nonlocal, quantum-mechanical means, without recourse to local interactions with classical electromagnetic fields. This spin AB effect, which is a spin analog of the usual charge AB effect, relies on the helical and topological nature of the QSH

edge states which is peculiar to that state of matter, combined with a Berry phase effect. In addition, we have shown that the spin AB effect can be used to design a different kind of spin transistor which is fundamentally different from the previous proposals, in that there is no classical force or torque acting on the spin of the electron. Furthermore, edge transport in the QSH regime being dissipationless,^{10–12} the proposed topological spin transistor would have the advantage of a lower power consumption in comparison to previous proposals for spin transistors. More generally, such a quantum manipulation of the electron spin, if observed, could open different directions in spintronics research and applications, and would at the same time demonstrate the practical usefulness of topological states of quantum matter.

Recently we became aware of a paper by Usaj²⁷ which discusses a similar effect in the spin-polarized edge states of graphene ribbons. We expect our effect to be more robust to external perturbations due to the topological protection of the QSH edge states. Indeed, the helical edge liquid of the QSH state is a novel state of matter which is topologically distinct^{14,15} from the edge states of graphene.

ACKNOWLEDGMENTS

We are especially grateful to S.-C. Zhang for many illuminating discussions and collaborations at the early stage of this project. We also thank C.-C. Chen, B. Huard, T. L. Hughes, M. König, C.-X. Liu, Y. Oreg, S. Raghu, and H. Yao for insightful discussions. J.M. is supported by the National Science and Engineering Research Council (NSERC) of Canada, the Fonds québécois de la recherche sur la nature et les technologies (FQRNT), and the Stanford Graduate Program (SGF). E.-A.K. is supported in part by the Nanoscale Science and Engineering Initiative of the National Science Foundation under NSF Award No. EEC-0646547. X.-L.Q. is supported by the Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering, under Contract No. DE-AC02-76SF00515.

APPENDIX A: S-MATRIX ANALYSIS

We wish to obtain an expression for the S -matrix \mathcal{S} relating outgoing b to incoming a current amplitudes,

$$\begin{pmatrix} b_l \\ b_r \end{pmatrix} = \mathcal{S} \begin{pmatrix} a_l \\ a_r \end{pmatrix} \quad \text{with} \quad \mathcal{S} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix}, \quad (\text{A1})$$

where a_l and b_l (a_r and b_r) are $p_L \times 1$ ($p_R \times 1$) column vectors of the current amplitudes outside the QSH region in the left (right) lead (see Fig. 1), and p_L (p_R) is the number of propagating channels at the Fermi energy in the left (right) lead. The matrix \mathcal{S} therefore has dimensions $(p_L + p_R) \times (p_L + p_R)$ and the submatrices r, r' and t, t' are reflection and transmission matrices, respectively. The two-terminal conductance G from left to right is given by the Landauer formula²⁰ $G = \frac{e^2}{h} \text{tr} \, t t^\dagger$. We assume that phase coherence is preserved throughout the sample so that \mathcal{S} can be obtained by combining S matrices for different portions of the device coherently.²⁰ We define the $(p_{L,R} + 2) \times (p_{L,R} + 2)$ scattering matrices S_L, S_R for the left (L) and right (R) FM/QSH junc-

tions [e.g., see Fig. 2(a) for the left junction],

$$\begin{pmatrix} b_l \\ b_{l'} \end{pmatrix} = S_L \begin{pmatrix} a_l \\ a_{l'} \end{pmatrix}, \quad \begin{pmatrix} b_{r'} \\ b_r \end{pmatrix} = S_R \begin{pmatrix} a_{r'} \\ a_r \end{pmatrix}, \quad (\text{A2})$$

where l' (r') is the QSH region immediately to the right (left) of the left (right) junction, such that $a_{l'}, a_{r'}$ and $b_{l'}, b_{r'}$ are the two-component spinors of edge-state current amplitudes. They are related through the geometric AB phase φ (different for each spin polarization) and the dynamical phase $\lambda = 2k_F \ell$ (identical for both spin polarizations) where ℓ is the distance traveled by the edge electrons from left to right junction and k_F is the edge-state Fermi wave vector,

$$\begin{Bmatrix} a_{r'\uparrow,\downarrow} \\ a_{l'\uparrow,\downarrow} \end{Bmatrix} = e^{i\lambda/2} e^{\mp i\varphi/2} \begin{Bmatrix} b_{l'\uparrow,\downarrow} \\ b_{r'\uparrow,\downarrow} \end{Bmatrix}, \quad (\text{A3})$$

where the upper sign for φ corresponds to spin up. Using Eqs. (A2) and (A3), we can write

$$\begin{pmatrix} e^{-i\lambda/2} \Phi^\dagger a_{l'} \\ b_r \end{pmatrix} = S_R \begin{pmatrix} e^{i\lambda/2} \Phi b_{l'} \\ a_r \end{pmatrix}, \quad (\text{A4})$$

where we define $\Phi \equiv e^{-i\varphi\sigma_z/2}$. Using the first equality in Eq. (A2) together with Eq. (A4), we can eliminate the intermediate amplitudes $a_{l'}, b_{l'}$ and obtain relations between the left lead amplitudes a_l, b_l and the right lead amplitudes a_r, b_r , which gives us \mathcal{S} [Eq. (A1)]. The 2×2 transmission matrix t , i.e., the lower left block of \mathcal{S} , is then obtained in the form

$$t = t_R S_L, \quad (\text{A5})$$

where t_L and t_R are the $2 \times p_L$ and $p_R \times 2$ transmission matrices for the left and right junctions, respectively [i.e., the lower left blocks of S_L, S_R following the notation of Eq. (A1)], and S is a 2×2 matrix defined in Eq. (1). The effective spin-density matrices ρ_L, ρ_R of the FM leads used in Eq. (2) are defined as $\rho_L = t_{L,L}^\dagger$ and $\rho_R = t_{R,R}^\dagger$.

If the arms of the ring are asymmetric, the dynamical phase λ is generally different for each arm and we have $\lambda_{\text{bottom}} - \lambda_{\text{top}} \equiv \delta \neq 0$. In this case, one can show that Eq. (2) still holds but with the substitutions

$$\rho_L(\theta_L) \rightarrow R_\delta \rho_L(\theta_L) R_\delta^{-1} = \rho_L(\theta_L + \delta),$$

$$r'_L(\theta_L) \rightarrow R_\delta r'_L(\theta_L) R_\delta^{-1} = r'_L(\theta_L + \delta),$$

where $R_\delta \equiv e^{-i\sigma_z \delta/2}$ rotates the spin about the z axis by an angle δ . In other words, a phase asymmetry is equivalent to a rigid flux-independent rotation of the electron spin, and simply shifts the conductance pattern by a constant angle δ ,

$$G(\phi, \theta \equiv \theta_R - \theta_L) \rightarrow G[\phi, \theta_R - (\theta_L + \delta)]$$

$$= G(\phi, \theta - \delta). \quad (\text{A6})$$

APPENDIX B: SCATTERING AT THE JUNCTION

In order to solve the 1D scattering problem at the FM/QSH interface, we first observe that the number of degrees of

freedom is equal on either side of the junction. If the Fermi level ε_F is chosen such that both spin subbands in the FM leads are occupied, there are four propagating modes on each side of the junction (two spins and two chiralities). The QSH spin states $\phi^{\text{QSH}(\pm)}$ are σ_z eigenstates while the FM spin states $\phi^{\text{FM}(\pm)}(\theta)$ are eigenstates of $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}$ and depend explicitly on θ . The Schrödinger equation for the junction is then solved by the following scattering ansatz:

$$\psi_{\sigma}^{(+)}(x) = \begin{cases} \frac{\phi_{\sigma}^{<(+)}}{\sqrt{v_{\sigma}^{<}}} e^{ik_{\sigma}^{<}x} + \sum_{\sigma'} r_{\sigma'\sigma} \frac{\phi_{\sigma'}^{<(-)}}{\sqrt{v_{\sigma'}^{<}}} e^{-ik_{\sigma'}^{<}x} & x < 0, \\ \sum_{\sigma'} t_{\sigma'\sigma} \frac{\phi_{\sigma'}^{>(+)}}{\sqrt{v_{\sigma'}^{>}}} e^{ik_{\sigma'}^{>}x} & x > 0, \end{cases}$$

for a right-moving scattering state, and with similar expressions for a left-moving scattering state $\psi_{\sigma}^{(-)}$. Spin is denoted by σ , chirality by \pm , and side of the junction by \langle, \rangle . The propagating modes are explicitly normalized to unit flux such that $r_{\sigma'\sigma}$ and $t_{\sigma'\sigma}$ are the desired reflection and transmission matrices. Requiring the continuity of $\psi_{\sigma}^{(\pm)}$ and $\hat{v}_x \psi_{\sigma}^{(\pm)}$ at the interface $x=0$ (with $\hat{v}_x \equiv \partial H / \partial k_x$ the velocity operator), we obtain a system of linear equations for the 16 matrix elements r_L, t_L, r'_L, t'_L constituting S_L . As illustrated in Fig. 1, the magnetization angle is set to zero in the left lead and to θ in the right lead and we obtain $r'_L(0)$ and $r_R(\theta)$ in Eq. (1).

APPENDIX C: TIGHT-BINDING MODEL

The effective tight-binding model describing HgTe QWs is defined as^{7,19}

$$\mathcal{H} = \sum_i c_i^\dagger V_i c_i + \sum_{ij} (c_i^\dagger T_{ij} e^{iA_{ij}} c_j + \text{H.c.}), \quad (\text{C1})$$

where $T_{ij} = T_{\hat{x}} \delta_{j,i+\hat{x}} + T_{\hat{y}} \delta_{j,i+\hat{y}}$ is the nearest-neighbor hopping matrix, $A_{ij} = \frac{e}{\hbar c} \int_i^j d\mathbf{r} \cdot \mathbf{A}$ is the Peierls phase with \mathbf{A} the electromagnetic vector potential, and V_i , $T_{\hat{x}}$, and $T_{\hat{y}}$ are 4×4 matrices containing the $\mathbf{k} \cdot \mathbf{p}$ parameters and the effective Zeeman term. The 4×4 matrices $T_{\hat{x}}$, $T_{\hat{y}}$, and V_i used in the tight-binding Hamiltonian (C1) are given by

$$T_{\hat{x}} = \begin{pmatrix} D_+ & -\frac{iA}{2} & -\frac{i\Delta_e}{2} & 0 \\ -\frac{iA}{2} & D_- & 0 & -\frac{i\Delta_h}{2} \\ -\frac{i\Delta_e}{2} & 0 & D_+ & \frac{iA}{2} \\ 0 & -\frac{i\Delta_h}{2} & \frac{iA}{2} & D_- \end{pmatrix},$$

$$T_{\hat{y}} = \begin{pmatrix} D_+ & \frac{A}{2} & \frac{\Delta_e}{2} & 0 \\ -\frac{A}{2} & D_- & 0 & -\frac{\Delta_h}{2} \\ -\frac{\Delta_e}{2} & 0 & D_+ & \frac{A}{2} \\ 0 & \frac{\Delta_h}{2} & -\frac{A}{2} & D_- \end{pmatrix}, \quad (\text{C2})$$

and

$$V_i = [C - 4D - \varepsilon_F + E_g(i)] \mathbb{1}_{4 \times 4} + (M - 4B) \mathbb{1}_{2 \times 2} \otimes \sigma_z + H_{Z\parallel}^{\text{eff}} + H_{Z\perp}^{\text{eff}}, \quad (\text{C3})$$

where $D_{\pm} \equiv D \pm B$ and $A, B, C, D, M, \Delta_e, \Delta_h$ are $\mathbf{k} \cdot \mathbf{p}$ parameters,¹⁹ and $\mathbb{1}_{n \times n}$ denotes the $n \times n$ unit matrix. The Fermi energy ε_F is uniform throughout the device. The gate potential $E_g(i)$ is different in the QSH and lead regions [Fig. 4(a)], and is used to tune the central region into the QSH insulating regime. The in-plane $H_{Z\parallel}^{\text{eff}}$ and out-of-plane $H_{Z\perp}^{\text{eff}}$ effective Zeeman terms, which are used to mimic the injection of spin-polarized carriers from a FM layer (Fig. 1), are given by¹⁹

$$H_{Z\parallel}^{\text{eff}} = g_{\parallel} \mu_B \begin{pmatrix} 0 & 0 & B_-^{\text{eff}} & 0 \\ 0 & 0 & 0 & 0 \\ B_+^{\text{eff}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$H_{Z\perp}^{\text{eff}} = \mu_B B_z^{\text{eff}} \begin{pmatrix} g_{E\perp} & 0 & 0 & 0 \\ 0 & g_{H\perp} & 0 & 0 \\ 0 & 0 & -g_{E\perp} & 0 \\ 0 & 0 & 0 & -g_{H\perp} \end{pmatrix}, \quad (\text{C4})$$

where $B_{\pm}^{\text{eff}} = B_x^{\text{eff}} \pm iB_y^{\text{eff}}$, $\mathbf{B}^{\text{eff}} = (B_x^{\text{eff}}, B_y^{\text{eff}}, B_z^{\text{eff}})$ is some effective magnetic field the role of which is to induce a spin polarization in the leads, μ_B is the Bohr magneton, and g_{\parallel} and $g_{E\perp}$, $g_{H\perp}$ are the in-plane and out-of-plane g factors, respectively.

- ¹I. Žutić, J. Fabian, and S. Das Sarma, *Rev. Mod. Phys.* **76**, 323 (2004).
- ²Y. Aharonov and D. Bohm, *Phys. Rev.* **115**, 485 (1959).
- ³D. Frustaglia, M. Hentschel, and K. Richter, *Phys. Rev. Lett.* **87**, 256602 (2001); M. Hentschel, H. Schomerus, D. Frustaglia, and K. Richter, *Phys. Rev. B* **69**, 155326 (2004); D. Frustaglia, M. Hentschel, and K. Richter, *ibid.* **69**, 155327 (2004); D. Frustaglia and K. Richter, *ibid.* **69**, 235310 (2004).
- ⁴C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 226801 (2005).
- ⁵B. A. Bernevig and S. C. Zhang, *Phys. Rev. Lett.* **96**, 106802 (2006).
- ⁶S. Murakami, *Phys. Rev. Lett.* **97**, 236805 (2006).
- ⁷B. A. Bernevig, T. L. Hughes, and S. C. Zhang, *Science* **314**, 1757 (2006).
- ⁸C. Liu, T. L. Hughes, X.-L. Qi, K. Wang, and S.-C. Zhang, *Phys. Rev. Lett.* **100**, 236601 (2008).
- ⁹A. Shitade, H. Katsura, J. Kuneš, X.-L. Qi, S.-C. Zhang, and N. Nagaosa, *Phys. Rev. Lett.* **102**, 256403 (2009).
- ¹⁰M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, *Science* **318**, 766 (2007).
- ¹¹A. Roth, C. Brüne, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, *Science* **325**, 294 (2009).
- ¹²M. Büttiker, *Science* **325**, 278 (2009).
- ¹³C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 146802 (2005).
- ¹⁴C. Wu, B. A. Bernevig, and S. C. Zhang, *Phys. Rev. Lett.* **96**, 106401 (2006).
- ¹⁵C. Xu and J. E. Moore, *Phys. Rev. B* **73**, 045322 (2006).
- ¹⁶A. D. Stone and A. Szafer, *IBM J. Res. Dev.* **32**, 384 (1988).
- ¹⁷J. C. Slonczewski, *Phys. Rev. B* **39**, 6995 (1989).
- ¹⁸S. Datta and B. Das, *Appl. Phys. Lett.* **56**, 665 (1990).
- ¹⁹M. König, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, *J. Phys. Soc. Jpn.* **77**, 031007 (2008).
- ²⁰D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures* (Cambridge University Press, Cambridge, 1997).
- ²¹S. Y. Wu, J. Cocks, and C. S. Jayanthi, *Phys. Rev. B* **49**, 7957 (1994).
- ²²M. König, Ph.D. thesis, University of Würzburg, 2007.
- ²³S. Sanvito, C. J. Lambert, J. H. Jefferson, and A. M. Bratkovsky, *Phys. Rev. B* **59**, 11936 (1999).
- ²⁴B. Zhou, H.-Z. Lu, R.-L. Chu, S.-Q. Shen, and Q. Niu, *Phys. Rev. Lett.* **101**, 246807 (2008).
- ²⁵R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems* (Springer-Verlag, Berlin, 2003).
- ²⁶Y. Aharonov and A. Casher, *Phys. Rev. Lett.* **53**, 319 (1984).
- ²⁷G. Usaj, *Phys. Rev. B* **80**, 081414(R) (2009).