Helical edge states induced by lateral spin-orbit coupling

A. Matos-Abiague*
Institute for Theoretical Physics, University of Regensburg, 93040 Regensburg, Germany
(Received 30 December 2012; published 15 April 2013)

The presence of edges locally breaks the inversion symmetry of heterostructures and gives rise to lateral (edge) spin-orbit coupling (SOC), which, under some conditions, can lead to the formation of helical edge states. If the edge SOC is strong enough, the helical edge states can penetrate the band gap and be energetically isolated from the bulk-like states. As a result, backward scattering is suppressed, dissipationless helical edge channels protected against time-inversion symmetric perturbations emerge, and the system behaves as a two-dimensional topological insulator (TI). However, unlike in previous work on TIs, the mechanism proposed here for the creation of protected helical edge states relies on the strong edge SOC rather than on band inversion.

DOI: 10.1103/PhysRevB.87.155306 PACS number(s): 72.20.−i, 71.70.Ej, 72.25.−b

I. INTRODUCTION

In contrast to quantum Hall systems, dissipationless edge states can be realized in two-dimensional (2D) topological insulators (TIs) without the need for an external magnetic field.[1,2] In 2D TIs, such as inverted HgTe/CdTe and inverted InAs/GaSb quantum wells (QWs), the TI phase (also known as a quantum-spin Hall insulator) is characterized by the existence of gapless, topologically protected helical edge states whose energies cross the bulk energy gap.[3–6] The edge channels consist of counterpropagating states with opposite spins. Since one of the edge states is the time reversal of the other, backscattering due to time-reversal symmetric perturbations is suppressed.[7–9] Topologically protected helical states can be found also at the surfaces of some 3D crystals with strong atomic spin-orbit coupling (SOC), the so-called 3D TIs.[1,2]

Until now the only mechanism recognized to drive TIs into the topological phase is band inversion. In the present work we theoretically explore the possibility of creating protected edge states by employing the lateral (edge) SOC, without resorting to band inversion.

II. THEORETICAL MODEL

We consider a heterostructure in which a two-dimensional electron gas (2DEG) is confined inside a symmetric QW of width d grown along the $\hat{\mathbf{z}} \parallel [001]$ direction. The QW is assumed to support a single bound state with energy ϵ_z. In addition, the step-like potential $V_{\text{edge}}(x)$ defines the edges (located at $x = 0$ and $x = L$) of the 2DEG along the $\hat{\mathbf{x}} \parallel [100]$ direction (see Fig. 1). When the distance, L, is large enough, the overlap between the wave functions localized at different edges can be neglected and each edge can be treated independently from the other. In this case it suffices to analyze one of the edges, say the edge at $x = 0$. In the vicinity of the edge at $x = 0$ the in-plane motion of the 2DEG is described by the effective 2D Hamiltonian,

$$H_{2D} = p_x\left[\frac{1}{2m^*(x)}p_x\right] + \frac{p_y^2}{2m^*(x)} + V_0\Theta(-x) + \epsilon_z + \frac{\alpha}{\hbar}\delta(x)p_y\sigma_z - \gamma\delta(x), \quad (1)$$

where p_x (p_y) is the x component (y component) of the momentum, σ_z is a Pauli matrix, V_0 is the height of the potential barrier at the edge, and $\delta(x)$ and $\Theta(x)$ denote the Dirac δ and unit step functions, respectively. The effective mass takes the values $m^* = m_{\text{in}}$ and $m^* = m_{\text{out}}$ in the regions $x > 0$ and $x < 0$, respectively. The presence of the edge at $x = 0$ locally breaks the structure inversion symmetry in the x direction. This gives rise to the *edge* SOC, which is described by the fifth term on the right-hand side of Eq. (1) and has an amplitude α. The last term accounts for the possible existence of localized Tamm-like edge states.[10–12] The effects of the edge SOC (also referred to as lateral SOC in the case of quantum wires) on the spin Hall effect and spin polarization in quantum wires and quantum point contacts have been investigated from both the theoretical[13–18] and the experimental points of view.[16,17] Spin-dependent scattering of bulk-like states caused by the edge SOC in a 2DEG[19] as well as the formation of interface states in 2DEGs with position-dependent Rashba and Dresselhaus SOCs[20] has also been theoretically investigated.

According to Eq. (1) both the y component of the momentum and the z component of the spin remain as good quantum numbers. The eigenfunctions of H_{2D} can then be written as

$$\Psi_\sigma(x,y) = \frac{1}{\sqrt{l}}e^{ik_y y}\psi_\sigma(x)\chi_\sigma, \quad (2)$$

where l is the length of the sample in the y direction and χ_σ denotes the eigenspinors of σ_z with $\sigma = 1$ and $\sigma = -1$ for up and down spins, respectively. The functions $\psi_\sigma(x)$ obey the effective Schrödinger equation,

$$\left\{p_x\left[\frac{1}{2m^*(x)}p_x\right] + V_\sigma(x)\right\}\psi_\sigma(x) = \epsilon_\sigma\psi_\sigma(x), \quad (3)$$

with the spin-dependent effective potential,

$$V_\sigma(x) = V_{1D}\Theta(-x) + (\sigma\alpha k_y - \gamma)\delta(x), \quad (4)$$

and $V_{1D} = V_0 - (1 - r_m)\epsilon_y$. Here we have introduced the effective mass ratio $r_m = m_{\text{in}}/m_{\text{out}}$. The eigenenergies, ϵ_σ, are measured with respect to the value $\epsilon_y + \epsilon_z$, where $\epsilon_y = \hbar^2 k_y^2/(2m_{\text{in}})$.

In samples with symmetric edges, the wave functions, $\Psi_\sigma^{(L)}(x,y)$, localized around the edge at $x = L$ can be obtained from those localized around $x = 0$ [see Eq. (2)] by using the symmetry relation, $\Psi_\sigma^{(L)}(x,y) = \Psi_{-\sigma}(L - x, y)$.

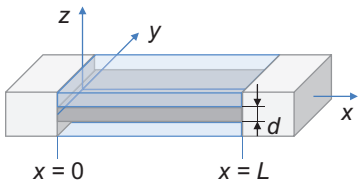

FIG. 1. (Color online) A 2DEG is formed in a symmetric QW of width d grown along the z axis. The edges are located at $x = 0$ and $x = L$. For symmetric edges, the system has a global structure inversion symmetry in the x direction. However, such a symmetry is locally broken by the presence of the edges. The local inversion asymmetry gives rise to lateral (edge) spin-orbit couplings with opposite strengths at opposite edges.

We first consider the case in which the edge states are induced exclusively from the edge SOC and take $\gamma = 0$. In such a case, the origin of the SOC-induced edge states can be qualitatively understood by analyzing the effective edge potential V_σ. The first term on the right-hand side of Eq. (4) represents an effective potential barrier at the edge, while for $\gamma = 0$ the second term can be attractive or repulsive, dependent on whether the sign of the product $\sigma\alpha k_y$ is negative or positive [see Figs. 2(a) and 2(b)]. Therefore, assuming that the constituent materials are such that $\alpha > 0$

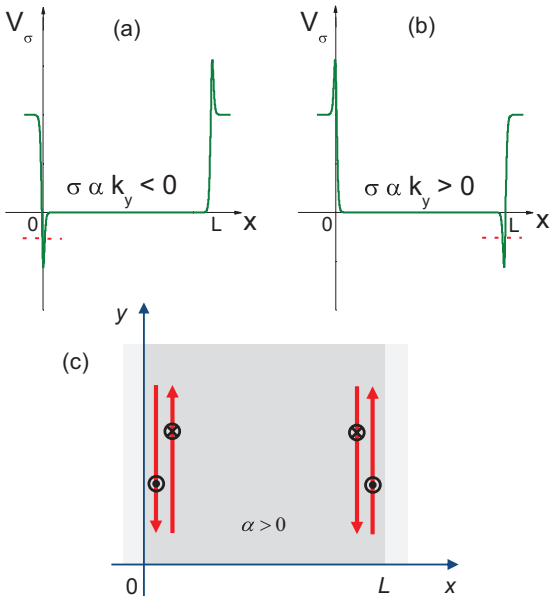

FIG. 2. (Color online) Formation of edge states. In the presence of the edge SOC, the effective edge potentials become spin dependent. Close to the edge at $x = 0$ the potential is attractive (a) or repulsive (b), dependent on whether the product $\sigma\alpha k_y$ is negative or positive. When the potential is strong enough, localized edge states can emerge. For $\alpha > 0$, V_σ is attractive for spin-up ($\sigma = 1$) carriers with $k_y < 0$ and spin-down ($\sigma = -1$) carriers with $k_y > 0$. This leads to the formation of two counterpropagating edge states with opposite spins. The effective edge potential at $x = L$, $V_\sigma^{(L)}$ is related to that at $x = 0$ by the symmetry relation $V_\sigma^{(L)}(x) = V_{-\sigma}(L - x)$. Therefore, edge states similar to those at $x = 0$ also appear at $x = L$, but with inverted spins. (c) Schematic of the spin-polarized, counterpropagating edge states. At equilibrium, the finite spin currents flowing along different edges compensate each other and the total spin current vanishes.

(the generalization to the case $\alpha < 0$ is straightforward), the spin-dependent potential V_σ forms an asymmetric QW for both spin-up ($\sigma = 1$) carriers with $k_y < 0$ and spin-down ($\sigma = -1$) carriers with $k_y > 0$, as shown schematically in Fig. 2(a). Since the resulting edge QW is asymmetric, it does not support bound states for arbitrary values of k_y. However, bound states can emerge when $|k_y|$ is larger than a certain critical value. These bound states are localized in the x direction around the edge but propagate freely in the y direction. They represent two counterpropagating spin-polarized edge channels with opposite spin polarization, one of which is the time reversal of the other [see Fig. 2(c)]. Note that the net spin polarization vanishes, as dictated by the time-reversal invariance of the Hamiltonian in Eq. (3). In a nonequilibrium situation, however, the edge SOC can lead to spin-dependent scattering of bulk-like states at the edge and produce a finite spin polarization.[19]

The eigenenergies and eigenfunctions of Eq. (3) corresponding to bound states are given, respectively, by

$$\epsilon_\sigma = -\frac{\hbar^2 k_\sigma^2}{2m_{\rm in}} \tag{5}$$

and

$$\psi_\sigma(x) = \sqrt{\frac{2\kappa_\sigma k_\sigma}{\kappa_\sigma + k_\sigma}} \times \begin{cases} e^{\kappa_\sigma x}, & x < 0, \\ e^{-k_\sigma x}, & x \geqslant 0, \end{cases} \tag{6}$$

where $\kappa_\sigma = \sqrt{(k_\sigma^2 + q_0^2)/r_m}$ and $q_0 = \sqrt{2m_{\rm in}V_{\rm 1D}/\hbar^2}$. The wave vector k_σ obeys the dispersion relation

$$k_\sigma + \sqrt{r_m\left(k_\sigma^2 + q_0^2\right)} + \sigma\lambda_{\rm SO}k_y = 0, \tag{7}$$

where $\lambda_{\rm SO} = 2m_{\rm in}\alpha/\hbar^2$ is a dimensionless parameter characterizing the strength of the edge SOC. In what follows we focus on the case $m_{\rm out} > m_{\rm in} > 0$ and assume that q_0 is real. The other cases can be treated in a similar way.

For Eq. (6) to represent bound states, k_σ must be a positive real number and therefore the appropriate solutions of Eq. (7) reduce to the form

$$k_\sigma = \frac{-\sigma\lambda_{\rm SO}k_y - \sqrt{r_m\lambda_{\rm SO}^2k_y^2 + r_m(1 - r_m)q_0^2}}{1 - r_m}, \tag{8}$$

where $0 < r_m < 1$. The requirement that k_σ must be a positive real number imposes constraints on the existence of the solutions in Eq. (8). For example, if $\lambda_{\rm SO} > 0$ ($\lambda_{\rm SO} < 0$), a necessary, although not sufficient, condition for the existence of bound states is $\sigma k_y < 0$ ($\sigma k_y > 0$), as inferred from Eq. (8). This confirms our early argument that the SOC-induced edge states represent two spin-polarized counterpropagating channels with opposite spins. On the other hand, positive real values of k_σ, and therefore edge states, do not exist for arbitrary values of k_y but only for those values for which the first term in the numerator in Eq. (8) becomes larger than the second one. A more general analysis of Eq. (8) reveals that if the SOC strength is such that $\lambda_{\rm SO}^2 < (1 - r_m)$, the domain of k_y values for which edge states exist is given by

$$\frac{r_m b}{r_m(1 - r_m) + \lambda_{\rm SO}^2} < (k_y d)^2 < \frac{b}{(1 - r_m) - \lambda_{\rm SO}^2}. \tag{9}$$

On the other hand, if $\lambda_{SO}^2 > (1 - r_m)$ the k_y domain for edge states is

$$(k_y d)^2 > \frac{r_m b}{r_m(1 - r_m) + \lambda_{SO}^2}. \tag{10}$$

In the equations above the dimensionless parameter $b = V_0/\epsilon_0$ characterizes the height of the edge potential in units of the energy $\epsilon_0 = \hbar^2/(2m_{in}d^2)$.

In contrast to the edge states of the integer quantum Hall phase and/or in inverted band 2D TIs, the edge states induced by the edge SOC are, in general, not protected against disorder. This is so because although scattering from one edge state to the other is forbidden by time-reversal symmetry, backscattering can occur through the bulk-like states which, in general, coexist with the edge states in the same energy range. However, protected edge states may appear in gapped materials with strong enough edge SOC. In such a case edge states can exist inside the band gap and be energetically isolated from the bulk-like states, as shown below.

The band gap of the material forming the 2DEG is located below the ground-state energy of the QW, ϵ_z. In order to induce bound states inside the gap, the SOC strength must be such that the total energy of the edge states, $E_\sigma = \epsilon_\sigma + \epsilon_y + \epsilon_z$, becomes lower than ϵ_z; i.e., the condition

$$\epsilon_\sigma + \epsilon_y = \frac{\hbar^2}{2m_{in}}\left(k_y^2 - k_\sigma^2\right) < 0 \tag{11}$$

must be satisfied. This condition imposes an additional constraint for the edge states to be protected against time-reversal invariant perturbations. Using Eqs. (8) and (11) one finds that a necessary condition for the emergence of protected edge states is

$$|\lambda_{SO}| > 1 + r_m. \tag{12}$$

When the edge SOC is weak, Eq. (12) is not fulfilled and only unprotected edge states exist. However, when the edge SOC is strong enough and Eq. (12) is satisfied, the dissipationless edge states emerge in the k_y domain specified by

$$|k_y d| > \sqrt{\frac{r_m b}{(1 + \lambda_{SO})^2 - r_m^2}}. \tag{13}$$

III. RESULTS

The emergence of the helical edge states computed for $r_m = 0.2$ and $V_0/\epsilon_0 = 100$ is displayed in Fig. 3(a), as a function of the spin-orbit parameter λ_{so} and the momentum. The white zones correspond to the absence of edge states, while the red and hatched blue regions correspond, respectively, to the unprotected and protected helical edge states. The upward-pointing arrows (downward-pointing arrows) indicate that the spin orientation of the states in the corresponding quadrants is parallel (antiparallel) to the $\hat{\mathbf{z}}$ direction. As discussed above, at $\lambda_{so} = 0$ there are no edge states. However, when $|\lambda_{so}|$ starts to deviate from 0, unprotected helical edge states emerge in a reduced region of the k_y space. As $|\lambda_{so}|$ increases, the domain of the k_y space corresponding to unprotected edge states becomes wider, and when the

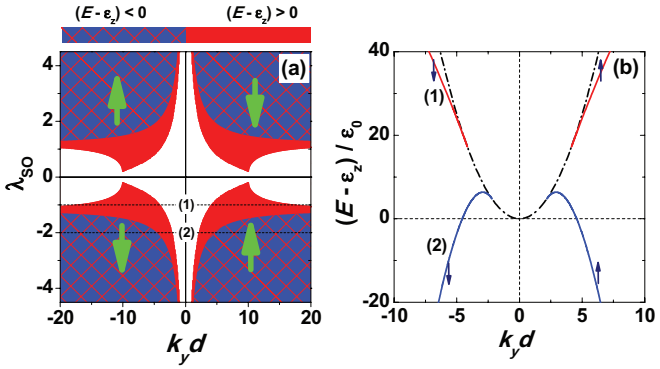

FIG. 3. (Color online) (a) Energy of edge states for $\gamma = 0$ as a function of the SOC parameter (λ_{so}) and the momentum along the edge (k_y). White regions correspond to the absence of edge states. Red and hatched blue zones represent, respectively, unprotected and protected edge states. Arrows indicate the spin orientation (up or down with respect to the z axis) in the corresponding quadrants. (b) Energy dispersion of the unprotected ($\lambda_{so} = -1$) and protected ($\lambda_{so} = -2$) edge states corresponding to dotted lines (1) and (2) in (a). The dash-dotted line represents the bulk-like states. The spin orientation of the different energy branches is indicated by the arrows.

condition in Eq. (12) is fulfilled, the protected edge states start to appear. The energy dispersions of the edge states corresponding to lines (1) and (2) in Fig. 3(a) (i.e., for $\lambda_{so} = -1, -2$) are shown in Fig. 3(b), where the energy of the spin-degenerate bulk-like states (dash-dotted line) is also included for comparison. The spin orientation of the different energy branches is indicated by the arrows. The pair of helical edge states labeled (1) coexists with the bulk-like states and is therefore not protected against disorder (backscattering can occur through transitions from the edge states to the bulk-like states). On the other hand, the helical edge states labeled (2) enter into the gap region ($E - \epsilon_z < 0$). Being isolated from the bulk-like states, these two counterpropagating spin-polarized states, one of which is the time reversal of the other, become protected against time-reversal perturbations.

The penetration of the protected edge states into the gap region can be easily understood in the extreme case $r_m \ll 1$ (i.e., $m_{out} \gg m_{in}$). In this case Eq. (8) can be approximated by $k_\sigma \approx -\sigma\lambda_{so}k_y$ and yields

$$E_\sigma - \epsilon_z \approx \frac{\hbar^2 k_y^2}{2\mu}, \tag{14}$$

with the SOC-dependent mass,

$$\frac{1}{\mu} = \frac{1}{m_{in}}\left(1 - \lambda_{so}^2\right). \tag{15}$$

Thus, if $|\lambda_{so}| > 1$ [note that for $r_m \to 0$, this is consistent with Eq. (12)], the mass μ changes sign and the sub-bands start to penetrate into the gap region.

In the 2D TIs relying on band inversion the protected edge states form a Dirac point inside the energy gap and the corresponding energy spectrum is almost linear. In contrast, as shown in Fig. 3(b), the protected edge states induced purely by the edge interface SOC do not form a Dirac point and exhibit a strongly nonlinear spectrum. A similar behavior of the energy

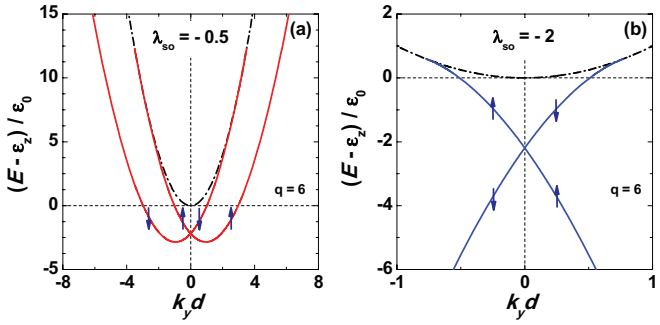

FIG. 4. (Color online) Energy spectrum of (a) Tamm-like unprotected edge states ($\lambda_{so} = -0.5$) and (b) Tamm-like protected edge states ($\lambda_{so} = -2$) for different values of the parameter $q = 2m_{in}\gamma d/\hbar^2$. Arrows indicate the spin orientation (up or down with respect to the z axis) of the corresponding energy branches. The bulk-like band (dash-dotted line) is also included for comparison.

dispersion has also been found for the edge states in Bi(111) ultrathin films.[21,22]

We now consider the case $\gamma \neq 0$ and conveniently introduce the dimensionless parameter $q = 2m_{in}\gamma d/\hbar^2$. As long as the inequality $0 < q < \sqrt{r_m}q_0(k_y = 0)$ holds, there are no qualitative differences from the case $q = 0$. The only effect is the reduction of the k_y domain in which edge states do not exist [i.e., the white zones in Fig. 3(a) are reduced with increases in q]. This is because as q increases, the energy branches corresponding to the counterpropagating edge states tend to match each other at the zone center. If the value of γ is such that $q > \sqrt{r_m}q_0(k_y = 0)$, a Dirac point at $k_y = 0$ emerges inside the energy gap and helical Tamm-like edge states appear. The energy dispersion of these states is shown in Fig. 4 for $q = 6$ [for $r_m = 0.2$ and $V_0/\epsilon_0 = 100$, one finds $\sqrt{r_m}q_0(k_y = 0) = 4.47$]. Unprotected and protected helical Tamm-like edge states are represented, respectively, by solid lines in Figs. 4(a) and 4(b). The dash-dotted lines describe the bulk-like states and the arrows indicate the spin orientation of the corresponding energy branch. The unprotected edge states exhibit a typical Rashba-type spectrum [see Fig. 4(a)],[23] while the energy spectrum of the protected edge states [see Fig. 4(b)] resembles that of an inverted band TI.

IV. CONCLUSIONS

Protected edge and surface states are known to appear in systems with strong atomic SOC in which a band inversion drives the system into a TI phase.[1,2] The mechanism proposed in this work does not rely on band inversion but rather on the use of the edge SOC. It represents a new alternative for the creation of protected helical edge states that could be relevant for engineering new types of TIs. Unfortunately, although the model presented here provides a qualitative description of the formation of such states, it is difficult at the present stage to predict which materials should be combined for the condition in Eq. (12) to be fulfilled and the protected helical states to appear. This is so because the parameter characterizing the strength of the edge SOC (α or its dimensionless version, λ_{so}) is not precisely known for such systems. In general, the value of α must be extracted from first-principles calculations or deduced from experiments. Nevertheless, rough estimations of α have been reported for some interfaces between III-V semiconductors.[24] We have performed similar estimations for other kinds of experimentally relevant interfaces between III-V semiconductors and found that for these interfaces the condition in Eq. (12) is, in general, not fulfilled. However, for $InSb/In_{0.65}Al_{0.35}Sb$ and $Ga_{0.47}In_{0.53}As/Al_{0.48}In_{0.52}As$ interfaces we found $|\lambda_{so}| \sim 0.3$, a value that is only about 4 to 5 times smaller than the value required for the emergence of the protected helical edge states. This suggests that the proposed mechanism may indeed exist in other kind of edges (e.g., with vacuum or between materials with noncommon atoms) where the interface SOC can, in principle, be enhanced even by orders of magnitude.

ACKNOWLEDGMENTS

I am grateful to J. Fabian and B. Scharf for fruitful discussions. This work was supported by the DFG (Grant No. SFB 689).

*alex.matos-abiague@physik.uni-r.de

[1] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
[2] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).
[3] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).
[4] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).
[5] C. Liu, T. L. Hughes, X.-L. Qi, K. Wang, and S.-C. Zhang, Phys. Rev. Lett. **100**, 236601 (2008).
[6] I. Knez, R.-R. Du, and G. Sullivan, Phys. Rev. Lett. **107**, 136603 (2011).
[7] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[8] C. Wu, B. A. Bernevig, and S.-C. Zhang, Phys. Rev. Lett. **96**, 106401 (2006).
[9] C. Xu and J. E. Moore, Phys. Rev. B **73**, 045322 (2006).
[10] E. L. Ivchenko, A. Y. Kaminski, and U. Rössler, Phys. Rev. B **54**, 5852 (1996).
[11] E. E. Takhtamirov and V. A. Volkov, J. Exp. Theor. Phys. **89**, 1000 (1999).
[12] I. V. Tokatly, A. G. Tsibizov, and A. A. Gorbatsevich, Phys. Rev. B **65**, 165328 (2002).
[13] Y. Xing, Q.-F. Sun, L. Tang, and J. P. Hu, Phys. Rev. B **74**, 155313 (2006).
[14] K. Hattori and H. Okamoto, Phys. Rev. B **74**, 155321 (2006).
[15] S. Bellucci and P. Onorato, Phys. Rev. B **75**, 235326 (2007).
[16] P. Debray, S. M. S. Rahman, J. Wan, R. S. Newrock, M. Cahay, A. T. Ngo, S. E. Ulloa, S. T. Herbert, M. Muhammad, and M. Johnson, Nature Nanotechnol. **4**, 759 (2009).
[17] A. T. Ngo, P. Debray, and S. E. Ulloa, Phys. Rev. B **81**, 115328 (2010).

[18] K. Hattori, J. Phys. Soc. Jpn. **79**, 104709 (2010).
[19] P. Bokes and F. Horváth, Phys. Rev. B **81**, 125302 (2010).
[20] A. A. Sukhanov and V. A. Sablikov, J. Phys.: Condens. Matter **23**, 395601 (2011).
[21] S. Murakami, Phys. Rev. Lett. **97**, 236805 (2006).
[22] M. Wada, S. Murakami, F. Freimuth, and G. Bihlmayer, Phys. Rev. B **83**, 121310(R) (2011).
[23] J. Fabian, A. Matos-Abiague, C. Ertler, P. Stano, and I. Žutić, Acta Phys. Slov. **57**, 565 (2007).
[24] A. Matos-Abiague, Phys. Rev. B **81**, 165309 (2010).