Cancelation of confinement effect by spin-orbit coupling in narrow strips of two-dimensional topological insulators

Y. Takagaki
Paul-Drude-Institut für Festkörperelektronik, Hausvogteiplatz 5-7, 10117 Berlin, Germany
(Received 20 August 2014; revised manuscript received 26 September 2014; published 9 October 2014)

The opening of the energy gap at the Dirac point when the helical edge states of two-dimensional topological insulators are hybridized in narrow channels is investigated. While the energy gap widens exponentially as the width is reduced within the block-diagonal effective four-band model, the dependence is no longer monotonic when off-diagonal terms representing spin-orbit coupling are taken into account. The energy gap vanishes periodically as the width is varied with the period being almost inversely proportional to the strength of the spin-orbit coupling. Both the bulk inversion asymmetry and Rashba terms cause the cancellation of the confinement effect. The disappearance of the hybridization takes place when the overlap integral becomes zero in the presence of the spin-orbit coupling due to the oscillatory decay of the transverse wave function.

DOI: 10.1103/PhysRevB.90.165305 PACS number(s): 73.20.At, 73.21.Hb, 73.25.+i

For a two-dimensional (2D) topological insulator (TI) [1], helical edge states appear at its boundary even in the absence of an external magnetic field [2]. This is in contrast to the edge states in the quantum Hall effect of 2D electron and hole gases, where the states are generated by the cyclotron orbital motion in a magnetic field. The dispersion of the helical edge states in the 2D TIs is approximately linear and gapless. That is, the conduction and valence bands are connected by the edge states. The boundaries of strips of ideal 2D TIs are hence metallic while the bulk states are insulating. The spin and momentum associated with the TI edge states are locked to each other as a consequence of strong spin-orbit coupling. The locking gives rise to the absence of backscattering from nonmagnetic impurities, which may be utilized for applications. Experimentally, the suppression of scattering was demonstrated using CdTe/HgTe/CdTe quantum wells [3]. Quantized transport associated with the topologically protected edges states was observed when the well thickness was adjusted to a range for which the system became a TI.

In narrow strips of 2D TIs, the wave functions of the edge states at the two boundaries overlap with each other. The consequential hybridization of the wave function results in an opening of an energy gap at the Dirac point. As is the issue for using graphene for transistors, the lack of an energy gap in the Dirac-particle-like dispersion leads to serious shortcomings for applications [4]. It is, therefore, important to understand the characteristics of the energy gap for possibly realizing a large energy gap. It was derived using the block-diagonal effective four-band model that the size of the energy gap changes exponentially with the channel width [5]. In a numerical study where the "off-diagonal" terms (spin-orbit terms) were also taken into consideration, the dependence of the energy gap size on the channel width was revealed not to be monotonic [6]. The energy gap was even suggested to disappear at a width of about 380–390 nm for a HgTe-CdTe quantum well.

In this paper, we investigate how the disappearance of the energy gap depends on the values of the spin-orbit parameters. We demonstrate that the disappearance, in fact, occurs periodically as the channel width is varied. Both the bulk inversion asymmetry and Rashba terms are found to cause the removal of the confinement effect in a similar fashion. The disappearance period is identified to change almost in proportion to the inverse of the strength of the spin-orbit coupling. These properties are shown to arise as the transverse wave function of the edge states decays in an oscillatory manner when the spin-orbit coupling is present.

Our analysis is based on the effective four-band Hamiltonian derived for the quantum wells in HgTe-CdTe heterostructures [7]

$$H = \begin{bmatrix} C_k + M_k & Ak_+ & -iRk_- & -\Delta \\ Ak_- & C_k - M_k & \Delta & 0 \\ iRk_+ & \Delta & C_k + M_k & -Ak_- \\ -\Delta & 0 & -Ak_+ & C_k - M_k \end{bmatrix}, \quad (1)$$

where $k_\pm = k_x \pm ik_y$, $C_k = -D(k_x^2 + k_y^2)$, and $M_k = M - B(k_x^2 + k_y^2)$. The model is defined on a basis $(|e+\rangle, |h+\rangle, |e-\rangle, |h-\rangle)$, where e and h denote, respectively, the electron and hole bands and $\pm$ refers to the spin orientation. Unless otherwise stated, the following parameter values [2,8] are assumed: $M = -0.01$ eV, $D = -0.5$ eV nm^2, $B = -0.7$ eV nm^2, $A = 0.365$ eV nm, $R = -0.016$ eV nm, and $\Delta = 0.0016$ eV. The negative band edge parameter M implies that the bulk conduction and valence bands overlap with each other when their mixing is absent, i.e., $A = 0$. Lifting the degeneracy for nonzero A produces the bulk band gap with an emergence of the Dirac-particle-like helical edge states. In comparison to the block-diagonal version considered by Zhou *et al.* [5], we take into account also the bulk inversion asymmetry term Δ and the Rashba term, which is in proportion to R, for the spin-orbit coupling.

In Fig. 1, we show the dispersions for a 384-nm-wide quasi-one-dimensional channel. The calculations were performed using the tight-binding method assuming a square lattice having a lattice parameter of $a = 6$ nm [9–11]. The doubly degenerate dispersion in Fig. 1(a) corresponds to the block-diagonal case of $R = \Delta = 0$. The blue and red curves correspond to the bulk electron and hole states, respectively. For the helical edge states, shown by the green curves, the Dirac point energy is given by $E = -MD/B = 7.14$ meV in the limit of infinitely large channel width. The energy gap opening at the Dirac point is evidenced by the magnified plot in the inset. The energy gap is small as the channel width is

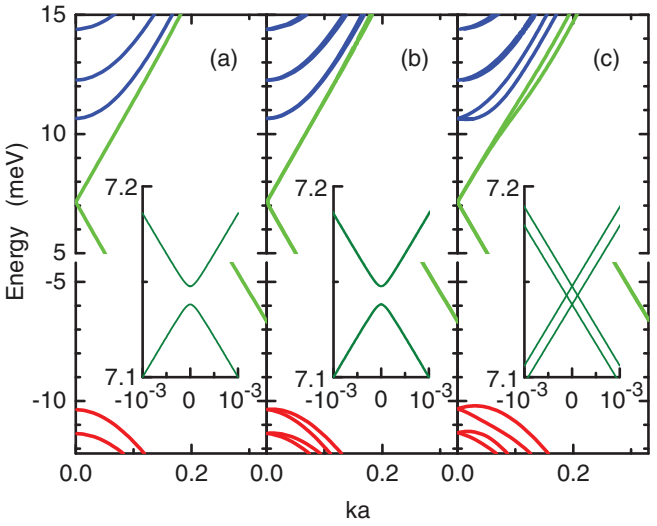

FIG. 1. (Color online) Dispersions in a 384-nm-wide channel of a CdTe/HgTe/CdTe topological insulator quantum well. The strengths of the spin-orbit coupling were set to be (a) $R = \Delta = 0$, (b) $R = -16$ meV nm and $\Delta = 0$, and (c) $R = -16$ meV nm and $\Delta = 1.6$ meV. The blue, green, and red curves correspond to the conduction band, edge states, and valence band, respectively. A square tight-binding lattice having a lattice parameter of $a = 6$ nm was used in calculating the wave number k for a given energy. The dispersion around the Dirac point is shown with magnified scales in the insets.

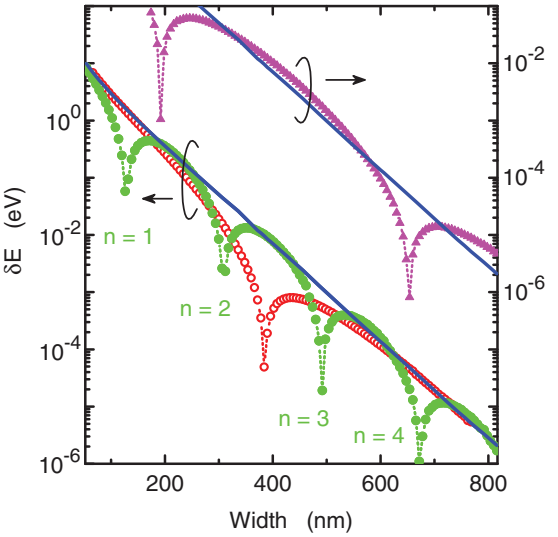

FIG. 2. (Color online) Dependence of magnitude δE of the energy gap at the Dirac point on the width of quasi-one-dimensional TI channel. For the effective four-band Hamiltonian describing the quantum well in a HgTe-CdTe heterostructure, $R = \Delta = 0$ was assumed for the solid curve. (Note that the two solid curves are identical.) For the open circles, filled circles, and filled triangles, the parameter values were assumed to be $R = -16$ meV nm and $\Delta = 1.6$ meV, $R = -16$ meV nm and $\Delta = 6$ meV, and $R = -600$ meV nm and $\Delta = 0$, respectively. The successive energy gap disappearances are indexed using n (=1, 2, ...) as indicated.

large. One notices that the bulk states are quantized due to the transverse confinement in the channel. The hybridization gap is seen to be much smaller than the separations between the quantized energy levels of the bulk states.

As shown in Figs. 1(b) and 1(c), the spin degeneracy is lifted when the spin-orbit terms are included in the simulations. In order to compare the magnitudes of the spin splitting arising from the bulk inversion asymmetry and Rashba contributions, Δ was set to be zero in Fig. 1(b). The bulk inversion asymmetry contribution is manifested to be dominant over the Rashba contribution in the present case.

We now focus our attention on the energy gap that opened at the Dirac point due to the hybridization of the edge states. In Fig. 2, we show the dependence of the magnitude δE of the energy gap on the width of the channel. The solid line corresponds to the case of $R = \Delta = 0$, for which the energy gap shrinks almost exponentially as the width increases [5,12]. (The solid line was redrawn for the curve shown by the triangles with an offset, i.e., the two solid lines in Fig. 2 are identical.) The open circles show the situation previously reported in Ref. [6], i.e., corresponding to a TI HgTe-CdTe quantum well. The dependence is no longer monotonic in the presence of the spin-orbit coupling ($|R|, |\Delta| > 0$). Moreover, the energy gap nearly vanishes at a width of about 384 nm. We emphasize that the bulk states remain quantized even when the confinement effect for the TI edge states is canceled by the spin-orbit coupling, as evidenced in Fig. 1(c). In other words, while the spin-obit coupling do not affect the quantization on the bulk states, as one would expect, they appear to be able to suppress the hybridization of the helical edge states.

For further investigating the influence of the spin-orbit coupling on the hybridization energy gap, the value of Δ was increased for the filled circles in Fig. 2 from 1.6 to 6 meV. We note that the values of the parameters in the effective four-band model change as the thickness of the HgTe/CdTe quantum well is varied [8]. The simulated situation can thus be experimentally realized, at least, to a certain extent. The width for the vanishing energy gap shifts to be smaller as Δ is increased. Moreover, the energy gap is seen to vanish repeatedly as the width is varied. For convenience, we index the series of the energy gap disappearance using n, as illustrated in Fig. 2.

We plot in Fig. 3 the channel widths W_n for which the energy gap vanishes. The data obtained by varying the value of Δ between 2.4 and 16 meV are shown by the circles. The curves show fittings of the n dependence of W_n assuming linear dependencies. The disappearance of the energy gap when the channel width is varied is demonstrated to take place almost exactly periodically. The period decreases with increasing Δ. As manifested in the inset, the straight lines nearly merge at a common point. We, therefore, find an approximate relationship $W_n = f(\Delta)(n - n_0) + W_0$ with $n_0 \approx 0.38$ and $W_0 \approx 18$ nm. We note that slight deviations from these values of n_0 and W_0 are noticed as Δ approaches zero.

The dependence of W_n on Δ is shown in Fig. 4(a). Approximate power-law dependencies of W_n on Δ are found, where the power-law exponent is suggested to be nearly independent of n. In fact, the slope $f(\Delta)$ for the lines in Fig. 3 is almost exactly in proportion to Δ^{-1}, see Fig. 4(b). (We would like to remind the readers that the dependence found in Fig. 3 is completely linear

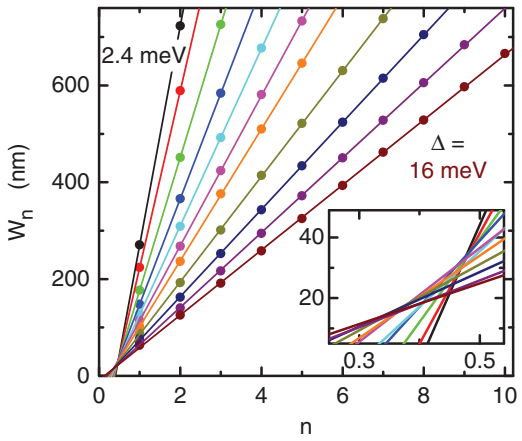

FIG. 3. (Color online) Widths W_n for the nth energy gap disappearance at the Dirac point. The numerical data are shown by the circles. The lines show fits to linear dependencies. The slope of the lines decreases monotonically with increasing Δ. The value of Δ was varied as 2.4, 3, 4, 5, 6, 7, 8, 10, 12, 14, and 16 meV for the cases from left to right, respectively. The approximate merging of the lines at a common point is shown in detail in the inset.

within numerical accuracies.) Here, the line manifests a dependence $\propto\Delta^{-1}$. Our numerical results thus obey the following relationship

$$W_n = cX^{\nu}(n - n_0) + W_0 \qquad (2)$$

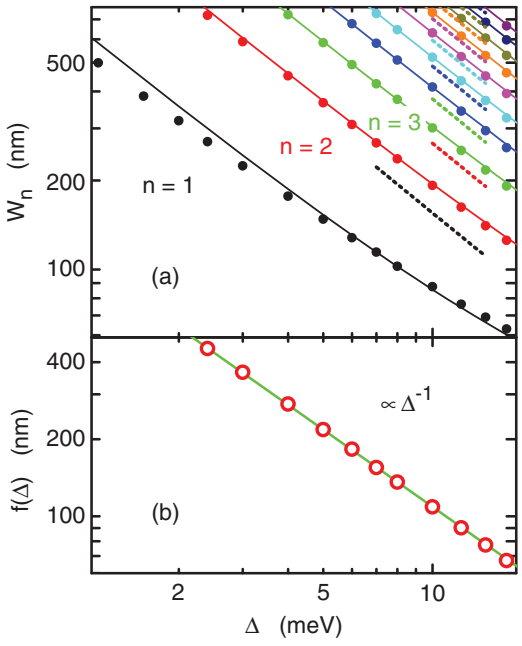

FIG. 4. (Color online) (a) Variation of widths W_n for the nth energy gap disappearance at the Dirac point with strength Δ of spin-orbit coupling originating from bulk inversion asymmetry. The circles show the numerical data for $n = 1, 2, 3, \ldots 10$. The behavior described by Eq. (2) is shown by the solid curves. The dotted lines show the prediction by Eq. (7). (b) Dependence of slopes $f(\Delta)$ for lines in Fig. 3 on Δ. The line shows a relationship $f(\Delta) \propto \Delta^{-1}$.

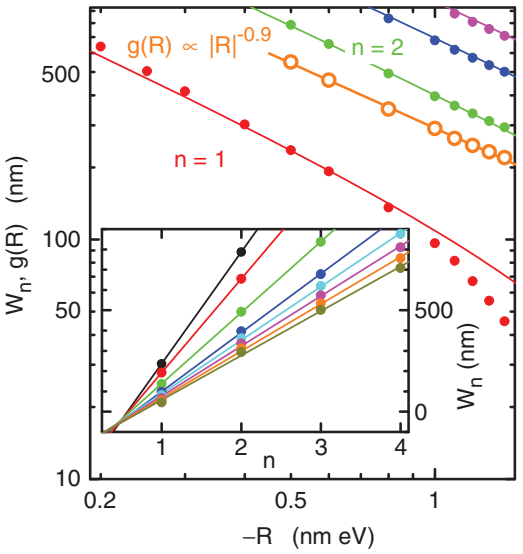

FIG. 5. (Color online) Dependencies of W_n and $g(R)$ on strength R of Rashba spin-orbit coupling. The filled circles show the widths W_n for the nth energy gap disappearance at the Dirac point with $n = 1, 2, 3$, and 4. The thin solid curves show the behavior described by Eq. (2). The open circles show the slopes $g(R)$ of the lines in the inset. The thick solid line shows a relationship $g(R) \propto |R|^{-0.9}$. Variation of n dependence of W_n when R is varied is shown in the inset. The numerical data are shown by the circles. The lines show fits to linear dependencies. The value of $-R$ is 0.5, 0.6, 0.8, 1, 1.1, 1.2, 1.3, and 1.4 eV nm for the cases from top to bottom, respectively.

with $X = \Delta$ being the strength of the spin-orbit coupling and $\nu = -1$. The solid curves in Fig. 4(a) show the behavior described by Eq. (2). It may be noteworthy that the power-law dependence of W_n on Δ becomes less accurate for smaller n in Fig. 4(a) as a consequence of W_0 not being zero.

The elimination of the hybridization energy gap by the spin-orbit coupling occurs similarly also when the Rashba contribution is significant. The value of R was set to -0.6 eV nm for the triangles in Fig. 2 instead of -0.016 eV nm. Here, $\Delta = 0$ was assumed to suppress the bulk inversion asymmetry contribution. ($\Delta = 0$ is kept throughout the discussion regarding the Rashba contribution below.) The characteristics of W_n associated with the Rashba term are summarized in Fig. 5. One can again extract an approximate relationship given by Eq. (2), similar to the case associated with the bulk inversion asymmetry term. For the numerical results in Fig. 5, we obtain $n_0 \approx 0.5$ and $W_0 \approx -37$ nm with $X = |R|$. We emphasize that the power-law exponent for the slopes $g(R)$ was gentler than for $f(\Delta)$. The thick solid line in Fig. 5 shows a dependence where $\nu = -0.9$. In comparison to the prediction by Eq. (2), which is shown by thin solid curves in Fig. 5, the deviation increases rapidly for $n = 1$ when $|R| > 1$ eV nm.

Regarding the curve shown by the triangles in Fig. 2, δE at widths away from the energy gap disappearances is larger than that when the spin-orbit coupling is absent. This is a general behavior when the spin-orbit coupling is strong. The exponential decay of δE with width becomes milder as the spin-orbit coupling is made stronger. The trend is recognizable also for the curve shown by the filled circles in Fig. 2. Due to the exponential decay of the edge wave function into

the interior of the channel, both the hybridization energy gap and the matrix element associated with the spin-orbit interaction exhibit nearly the same exponential dependence on the channel width [12]. The energy gap is, as a consequence, comparable in size with the spin splitting associated with the spin-orbit interaction [13], as one finds in the insets of Fig. 1.

In the remainder of the paper, we show that the disappearance of the energy gap occurs as the hybridization of the edge states can be absent in the presence of the spin-orbit coupling when the overlap integral vanishes. We solve the wave function analytically to investigate the overlap of the edge states at the two boundaries of a channel. Using the trial function $\varphi(y) \propto e^{\lambda y}$, the secular equation is obtained as

$$(B^2 - D^2)\lambda^4 + [2(MB + DE) - A^2]\lambda^2 + M^2 - E^2 + \Delta^2 = \pm i\sqrt{4A^2\Delta^2\lambda^2 + [M + E + (B - D)\lambda^2]^2 R^2\lambda^2}. \quad (3)$$

If the spin-orbit coupling is turned off ($\Delta = R = 0$), there exist solutions corresponding to the edge states having real λ [5]. However, λs acquire imaginary parts even for these edge states when $|\Delta|, |R| > 0$. The wave functions associated with the edge states no longer decay monotonically in the transverse direction but oscillate sinusoidally.

Let us now ignore the Rashba term, for simplicity. For small $|\Delta|$, $\lambda = \alpha + i\beta$ is obtained as

$$\alpha^2 = F \pm \sqrt{F^2 - \frac{M^2 - E^2}{B^2 - D^2}}, \quad (4)$$

$$\frac{1}{\beta} = \frac{2(B^2 - D^2)}{A\Delta}\sqrt{F^2 - \frac{M^2 - E^2}{B^2 - D^2}}, \quad (5)$$

where $F = [A^2 - 2(MB + DE)]/[2(B^2 - D^2)]$. We employ model wave functions $\varphi_1(y) \propto e^{-\alpha y}\sin(\beta y)$ and $\varphi_2(y) \propto e^{-\alpha(W-y)}\sin[\beta(W - y)]$ for the two edges of a channel with width W. The overlap is then estimated to be

$$\int_0^W \varphi_1\varphi_2 dy \propto e^{-\alpha W}\left[\frac{\sin(\beta W)}{\beta W} - \cos(\beta W)\right] = e^{-\alpha W}\Gamma(\beta W). \quad (6)$$

The overlap integral vanishes quasi-periodically as W is varied. The widths W_n for the vanishing energy gap are given by

$$W_n = \frac{\sqrt{B^2 - D^2}}{\Delta}\sqrt{\frac{A^2}{B^2 - D^2} - 4\frac{M}{B}}\gamma_n, \quad (7)$$

where $\gamma_1 = 4.49$, $\gamma_2 = 7.73$, $\gamma_3 = 10.9$, $\gamma_4 = 14.07$, $\gamma_5 = 17.22$, ... are the solutions of $\Gamma(\gamma) = 0$. Note that $E = -MD/B$ is satisfied when the energy gap vanishes. The dotted lines in Fig. 4 show the predictions. Given the simplifications introduced to derive Eq. (7), the nearly quantitative agreement with the numerical results is satisfactory. In particular, the periodicity of W_n with n is in excellent agreement.

For a certain range of material parameters, the imaginary part of λ is not zero even in the absence of the off-diagonal spin-orbit terms [14]. The energy gap at the Dirac point shrinks similarly in this case exhibiting a periodic dependence on the channel width. These special edge states are, however, absent in HgTe quantum wells [14]. The mechanism presented in the present paper can provide gap closing, in contrast, regardless of the material parameters so long as the off-diagonal spin-orbit terms cannot be ignored.

The energy gap opening due to hybridization takes place not only in narrow wires of 2D TIs but also in thin films of three-dimensional (3D) TIs, such as Bi_2Se_3. In Bi_2Se_3 films, the hybridization becomes crucial when the thickness is smaller than 6 nm [1]. The band structure in thin Bi_2Se_3 films can be described by a model resembling Eq. (1) [14,15], and so the off-diagonal spin-orbit terms are expected to alter the energy gap size also for 3D TIs. We note that the energy gap in 3D TIs, as a matter of fact, has been shown to depend on the film thickness in an oscillatory manner [15–17]. This phenomenon, however, presumably results from the special edge states as the oscillation period is independent on the off-diagonal spin-orbit terms.

In conclusion, we have demonstrated that there exist situations where the energy gap induced by the hybridization of the edge states in narrow strips of a 2D TI is almost completely erased by the spin-orbit interactions. The cancellation of the confinement effect occurs periodically as the channel width is varied. For the bulk inversion asymmetry and Rashba terms, the period has been evaluated to be almost inversely proportional to the strength of the spin-orbit coupling. We have identified that the wave function of the edge states decays sinusoidally in the transverse direction in the presence of the spin-orbit interactions. The overlap between the edge states at the two boundaries of the channel thereby vanishes at certain channel widths. Our work thus demonstrates that the off-diagonal spin-orbit terms affect the physics in 2D TIs in a qualitative manner.

[1] M. Z. Hazan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[2] M. König, H. Buhmann, L. W. Molenkamp, T. L. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, J. Phys. Soc. Jpn. **77**, 031007 (2008).

[3] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).

[4] M. C. Lemme, Solid State Phenomena **156**, 499 (2010).

[5] B. Zhou, H.-Z. Lu, R.-L. Chu, S.-Q. Shen, and Q. Niu, Phys. Rev. Lett. **101**, 246807 (2008).

[6] Y. Takagaki, J. Phys.: Condens. Matter **24**, 435301 (2012).

[7] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).

[8] D. G. Rothe, R. W. Reinthaler, C.-X. Liu, L. W. Molenkamp, S.-C. Zhang, and E. M. Hankiewicz, New J. Phys. **12**, 065012 (2010).

[9] T. Ando, Phys. Rev. B **44**, 8017 (1991).
[10] Y. Takagaki and Y. Tokura, Phys. Rev. B **54**, 6587 (1996).
[11] Y. Takagaki, Phys. Rev. B **85**, 155308 (2012).
[12] V. Krueckl and K. Richter, Phys. Rev. Lett. **107**, 086803 (2011).
[13] The energy gap δE is approximately proportional to $e^{-\lambda W}$, and the matrix element associated with the spin-orbit interaction is approximately given as $\eta \propto e^{-\lambda W}\lambda W\Delta$ with W being the width of the channel and λ the decay parameter of the edge state wave function [5,12]. The relationship Eq. (2) thus implies that the cancellation of the confinement effect takes place when conditions of $n\delta E/\eta = \text{const}$ are satisfied.
[14] F. Lu, Y. Zhou, J. An, and C.-D. Gong, Europhys. Lett. **98**, 17004 (2012).
[15] C.-X. Liu, H.-J. Zhang, B. Yan, X.-L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. B **81**, 041307(R) (2010).
[16] J. Linder, T. Yokoyama, and A. Sudbø, Phys. Rev. B **80**, 205401 (2009).
[17] W.-Y. Shan, H.-Z. Lu, and S.-Q. Shen, New J. Phys. **12**, 043048 (2010).