Finite size effects on helical edge states in HgTe quantum wells with the spin–orbit coupling due to bulk- and structure-inversion asymmetries*

Cheng Zhi(成 志) and Zhou Bin(周 斌)†

Department of Physics, Hubei University, Wuhan 430062, *China*

(Received 21 September 2013; revised manuscript received 26 November 2013; published online 14 January 2014)

There is a quantum spin Hall state in the inverted HgTe quantum well, characterized by the topologically protected gapless helical edge states lying within the bulk gap. It has been found that for a strip of finite width, the edge states on the two sides can couple together to produce a gap in the spectrum. The phenomenon is called the finite size effect in quantum spin Hall systems. In this paper, we investigate the effects of the spin–orbit coupling due to bulk- and structure-inversion asymmetries on the finite size effect in the HgTe quantum well by means of the numerical diagonalization method. When the bulk-inversion asymmetry is taken into account, it is shown that the energy gap E_g of the edge states due to the finite size effect features an oscillating exponential decay as a function of the strip width of the HgTe quantum well. The origin of this oscillatory pattern on the exponential decay is explained. Furthermore, if the bulk- and structure-inversion asymmetries are considered simultaneously, the structure-inversion asymmetry will induce a shift of the energy gap E_g closing point. Finally, based on the roles of the bulk- and structure-inversion asymmetries on the finite size effects, a way to realize the quantum spin Hall field effect transistor is proposed.

Keywords: quantum spin Hall state, finite size effect, spin–orbit coupling

PACS: 73.43.–f, 72.25.Dc, 85.75.–d **DOI:** 10.1088/1674-1056/23/3/037304

1. Introduction

The fact that the quantum spin Hall (QSH) state can appear in an inverted HgTe quantum well (QW) has been predicated theoretically[1] and later also verified experimentally.[2] The main feature of the QSH system is that there are topologically protected gapless helical edge states with different spins moving in opposite directions in the bulk energy gap. Recently, remarkable progress has been made in the experimental study of the QSH state in HgTe QWs.[3–7] The spin polarization of QSH edge states has been established,[4] and direct imaging of QSH edge states has been reported.[7] These interesting experiments demonstrate the potential applications of the QSH effect for spintronic devices. In fact, the first two-dimensional (2D) QSH system model proposed by Kane and Mele is a model for the graphene system with spin–orbit coupling (SOC).[8,9] Later, besides the inverted HgTe QWs, several other candidates of 2D QSH systems were also proposed, such as GaAs under shear strain,[10] a multilayer Bi thin film,[11] and an inverted InAs/GaSb/Alsb semiconductor QW.[12] On the other hand, the study of QSH has been generalized to three-dimensional (3D) topological insulators (TI) with the time reversal symmetry, characterized by the topologically protected spin helical surface states lying within the bulk gap.[13–17] Note that recently some authors have also investigated the time-reversal-symmetry-broken quantum spin Hall effect.[18–20]

For a strip HgTe QW with a finite narrow width L, the finite size effect in the 2D QSH system has been predicated by Zhou *et al.*[21] It was found that the edge states on the two sides can couple together to produce a gap in the spectrum. The gap due to the finite size effect was shown to become experimentally measurable around $L \simeq 200$ nm, which has been confirmed experimentally, suggesting that it destroys the QSH effect.[22] Thus the finite size effect should be seriously considered in the potential use of the topologically protected edge states of HgTe QWs in applications on the nanometer scale. The same finite size effect has also been predicated theoretically[23–26] and verified experimentally[27] in sufficiently thin 3D TIs films, in which the top and the bottom surface states couple to open an energy gap at the Dirac point.

Note that there are two kinds of SOCs induced respectively by the bulk-inversion asymmetry (BIA) and the structure-inversion asymmetry (SIA) in the HgTe QWs. The BIA is determined by the HgTe QW structure, whereas the SIA depends on internal and external electric fields of the material and can be tuned by a gate voltage. However, most theoretical discussions based on the four-band Bernevig–Hughes–Zhang (BHZ) model[1] usually neglect the effects of BIA and SIA, considering that the strength of the BIA is small and the SIA can be minimized to zero. In fact, several recent theoretical works have noticed the important roles of BIA and SIA

*Project supported by the National Natural Science Foundation of China (Grant No. 11274102), the Program for New Century Excellent Talents in Universities, China (Grant No. NCET-11-0960), and the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20134208110001).

†Corresponding author. E-mail: binzhou@hubu.edu.cn

© 2014 Chinese Physical Society and IOP Publishing Ltd

on the physical properties of HgTe QWs.[28–38] Moreover, it has been shown that the strong BIA and SIA are essential in the 2D QSH effects of an inverted InAs/GaSb/Alsb QW.[12] The InAs/GaSb/Alsb QW can be characterized by an effective four-band model, which is similar to the BHZ model for the HgTe QW but contains terms describing BIA and SIA.[12] In this paper, we study how the SOCs due to BIA and SIA affect the finite size effects in the QSH states of the inverted HgTe QWs. Firstly, we only take the BIA into account, and find that due to the finite size effect, the energy gap E_g of the edge states features an oscillating exponential decay as a function of the strip width of the HgTe QWs. Following the argument of 3D TIs thin films, we explain the origin of this oscillatory pattern on the exponential decay. Then, the SIA is also considered in addition to the BIA. It is shown that the SIA induces a shift of the energy gap closing point. Finally, based on the effects of the BIA and the SIA on the finite size effects, we propose a way to realize a QSH field effect transistor (FET).

2. Model

We use the effective four-band Hamiltonian to describe the electronic states in HgTe QWs[1,2,29]

$$\boldsymbol{H}=\begin{pmatrix} \varepsilon_k+\mathcal{M}_k & Ak_+ & -\mathrm{i}Rk_- & -\Delta \\ Ak_- & \varepsilon_k-\mathcal{M}_k & \Delta & 0 \\ \mathrm{i}Rk_+ & \Delta & \varepsilon_k+\mathcal{M}_k & -Ak_- \\ -\Delta & 0 & -Ak_+ & \varepsilon_k-\mathcal{M}_k \end{pmatrix}, \quad (1)$$

which is defined on a four-band basis

$$\{|E_1+\rangle, |H_1+\rangle, |E_1-\rangle, |H_1-\rangle\},$$

where $|E_1\pm\rangle$ and $|H_1\pm\rangle$ denote two electron-like subbands and two heavy hole subbands, respectively, and $\pm$ refers to the spin orientation. Here $k_\pm = k_x \pm \mathrm{i}k_y$, $\varepsilon_k = C - D(k_x^2+k_y^2)$, $\mathcal{M}_k = M - B(k_x^2+k_y^2)$, and A, B, C, D, and M are the expansion parameters, which are determined by the thickness of the HgTe QW and the material parameters. Additionally, in the effective four-band Hamiltonian (1), two SOC terms due to the BIA and the SIA are taken into account. The BIA and the SIA terms have coefficients Δ and R, respectively, and introduce off-diagonal matrix elements to couple the two diagonal blocks with opposite spin orientations. The BIA term describes the mixing between $|E_1+(-)\rangle$ and $|H_1-(+)\rangle$, and the strength Δ is determined by the HgTe QW structure. The SIA term can be recognized as the electron k-linear Rashba term connecting $|E_1+\rangle$ and $|E_1-\rangle$ subbands, while the heavy-hole k-cubic Rashba term is neglected.[12] The strength R of the Rashba term is related to the symmetries of the grown HgTe QWs and can be tuned by a gate voltage. For symmetric HgTe QWs, the strength R is negligibly small.[30] Note that the inverted InAs/GaSb/Alsb QW can also be characterized by an effective four-band model similar to Hamiltonian (1).[12]

Without the BIA and the SIA terms (i.e., $\Delta = R = 0$), Hamiltonian (1) is block diagonal and each block is exactly a massive Dirac Hamiltonian, which consists of an s-like electron subband and a p-like hole subband. There is a topological phase transition in this model, when the material parameters satisfy the condition $MB > 0$, the system is in a topologically nontrivial QSH phase protected by the time-reversal symmetry of this system; while for $MB < 0$, it is in a normal insulator phase.[1,21] For later use in numerical simulations, we choose model parameters describing the QSH state regime:[2] $A = 364.5$ meV·nm, $B = -686$ meV·nm^2, $D = -512$ meV·nm^2, $M = -10$ meV. Without loss of generality, we also assume $C = 0$. Note that the BIA and the SIA terms preserve the time-reversal symmetry, thus the two SOC terms taken into account in Hamiltonian (1) do not affect the topological stability of the QSH phase. In the following, we will discuss the effect of the SOCs on the finite size effects of the helical edge states in HgTe QWs.

For the purpose of numerically studying the finite size effects of this system, we work with a lattice regularization of this model. Thus the effective four-band Hamiltonian (1) can be rewritten as a lattice model as follows:

$$\begin{aligned} \boldsymbol{H} = \sum_{\boldsymbol{k}} \Big\{ & -\frac{2D}{a^2}[2-\cos(k_x a)-\cos(k_y a)]\boldsymbol{\sigma}_0\boldsymbol{\tau}_0 + M\boldsymbol{\sigma}_0\boldsymbol{\tau}_z \\ & -\frac{2B}{a^2}[2-\cos(k_x a)-\cos(k_y a)]\boldsymbol{\sigma}_0\boldsymbol{\tau}_z \\ & +\frac{A}{a}[\sin(k_x a)\boldsymbol{\sigma}_z\boldsymbol{\tau}_x - \sin(k_y a)\boldsymbol{\sigma}_0\boldsymbol{\tau}_y] + \Delta\boldsymbol{\sigma}_y\boldsymbol{\tau}_y \\ & +\frac{R}{2a}[\sin(k_x a)\boldsymbol{\sigma}_y - \sin(k_y a)\boldsymbol{\sigma}_x](\boldsymbol{\tau}_0+\boldsymbol{\tau}_z)\Big\} c_{\boldsymbol{k}}^\dagger c_{\boldsymbol{k}}, \end{aligned} \quad (2)$$

where a is a lattice constant, $\boldsymbol{\sigma}_0$ ($\boldsymbol{\tau}_0$) is the 2×2 identity matrix, $\boldsymbol{\sigma}_i$ ($i = x,y,z$) denote the Pauli matrices acting on the spin basis, and $\boldsymbol{\tau}_i$ ($i = x,y,z$) represent the Pauli matrices acting on the electron or heavy hole subbands. We now solve model (2) in a finite strip geometry of width L with the periodic boundary condition in the x direction and the open boundary condition in the y direction. In this case, k_y is no longer a good quantum number, while k_x is still a good quantum number. We must Fourier transform k_y into the position space to obtain a one-dimensional tight-binding model, which depends on the other momentum k_x, thus we make the substitution[28]

$$c_{\boldsymbol{k}} = \frac{1}{\sqrt{L}}\sum_n \mathrm{e}^{\mathrm{i}k_y n} c_{k,n}, \quad (3)$$

where $k \equiv k_x$. This gives the Hamiltonian

$$\boldsymbol{H} = \sum_{k,n}\left(\mathcal{E}c_{k,n}^\dagger c_{k,n} + \mathcal{T}c_{k,n}^\dagger c_{k,n+1} + \mathcal{T}^\dagger c_{k,n+1}^\dagger c_{k,n}\right), \quad (4)$$

with

$$\begin{aligned} \mathcal{E} = & -\frac{2D}{a^2}[2-\cos(ka)]\boldsymbol{\sigma}_0\boldsymbol{\tau}_0 - \frac{2B}{a^2}[2-\cos(ka)]\boldsymbol{\sigma}_0\boldsymbol{\tau}_z + M\boldsymbol{\sigma}_0\boldsymbol{\tau}_z \\ & +\frac{A}{a}\sin(ka)\boldsymbol{\sigma}_z\boldsymbol{\tau}_x + \Delta\boldsymbol{\sigma}_y\boldsymbol{\tau}_y + \frac{R}{2a}\sin(ka)\boldsymbol{\sigma}_y(\boldsymbol{\tau}_0+\boldsymbol{\tau}_z), \end{aligned} \quad (5)$$

$$\mathcal{T} = \frac{D}{a^2}\boldsymbol{\sigma}_0\boldsymbol{\tau}_0 + \frac{B}{a^2}\boldsymbol{\sigma}_0\boldsymbol{\tau}_z - \frac{\mathrm{i}A}{2a}\boldsymbol{\sigma}_0\boldsymbol{\tau}_y - \frac{\mathrm{i}R}{2a}\boldsymbol{\sigma}_x(\boldsymbol{\tau}_0+\boldsymbol{\tau}_z). \quad (6)$$

3. Results and discussion

In this section, we obtain the energy spectrum by numerical diagonalization of Hamiltonian (4) with a lattice constant $a = 2$ nm. As mentioned above, for symmetric HgTe QWs, the strength R of the Rashba term is negligibly small, thus we firstly address how the BIA term affects the finite size effect on the helical edge states of the HgTe QWs without the SIA (i.e., $R = 0$ in Hamiltonian (4)).

Figure 1 shows the energy spectra of the edge states of the HgTe QWs with width $L = 200$ nm and $L = 392$ nm, respectively. The dashed lines present the results without the SOC terms (i.e., $\Delta = R = 0$), an energy gap opens near $k = 0$ due to the finite size effect, consistent with the previous work.[21] When the BIA term is taken into account (here the strength Δ of 1.6 meV is adopted[28]), due to the SOC effect, spin splitting of the edge states occurs in the energy spectra (plotted by the solid lines in Fig. 1). As a result, for a narrow width (e.g., $L = 200$ nm, see Fig. 1(a)), the bottom of the conduction band and the top of the valance band are not located at the Γ point ($k = 0$) any more, but shift to finite k values. Thus the energy gap of the edge states (denoted by E_g in Fig. 1(a)) will generically open at this finite k rather than at the Γ point. In the case of $L = 392$ nm (see Fig. 1(b)), it is noted that without the SOC terms, the energy gap due to the finite size effect still appears, although it is very small, however, in the presence of the BIA term, the energy gap of the edge states nearly vanishes and the gapless edge states appear.

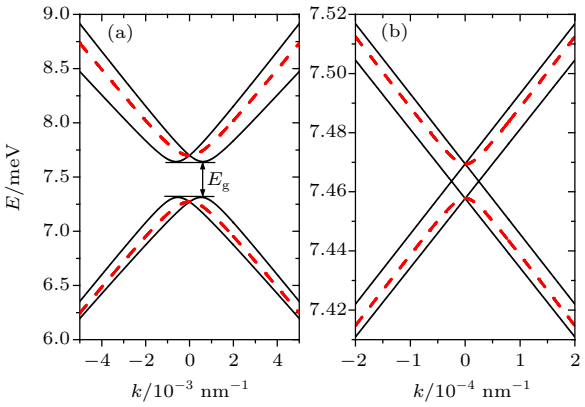


Fig. 1. (color online) Energy spectra of the edge states of HgTe QWs obtained by numerical diagonalization of Hamiltonian (4) without the SIA term (i.e., $R = 0$). The width is (a) $L = 200$ nm and (b) $L = 392$ nm. The (black) solid and the (red) dashed lines correspond to the models with ($\Delta = 1.6$ meV) and without ($\Delta = 0$) the BIA term, respectively. The other model parameters are given in the text.

The size-dependence of the energy gap E_g is plotted in Fig. 2. Figure 2(a) presents the variation of the energy gap with the width L in the absence of the SOC terms, in which the energy gap decays in an approximate exponential law of L. In this case, the analytical solution of the gap has been given by Zhou *et al.*[21] Figure 2(b) shows the result of the case with $\Delta = 1.6$ meV. Compared with the result of the case without the BIA plotted in Fig. 2(a), a qualitatively different scenario is shown in Fig. 2(b) for the case with the BIA: the variation of E_g with L is nonmonotonous and the energy gap appears to close at $L \simeq 392$ nm. This point has been noted in Ref. [31]. In general, the strength Δ of the BIA in HgTe QWs is determined by the material structure, and it was taken to be 1.6 meV for the QW of thickness $d = 7.0$ nm in Ref. [28]. In order to further understand the effect of the BIA on the finite size effect, in Figs. 2(c) and 2(d), we slightly overestimate the strength Δ of the BIA and take $\Delta = 3.2$ meV and 6.4 meV, respectively. From Figs. 2(c) and 2(d), it is seen that when the BIA term is taken into account, the nonmonotonous variation of E_g with L in fact features a superimposed oscillatory pattern on the exponential decay. The phenomenon is very similar to the variation pattern of the energy gap with the thickness of the thin film due to the finite size effect on the surface states in 3D TI Bi_2Se_3 thin films.[23–26] From Fig. 2, it is seen that with increasing Δ the energy gap E_g firstly vanishes at a narrower width. For instance, when $\Delta = 1.6$ meV, 3.2 meV, and 6.4 meV, the energy gap E_g firstly closes nearly at $L \simeq 392$ nm, 220 nm, and 128 nm, respectively, given by numerical calculations. Moreover, with increasing Δ the oscillation period of the gap decreases. For $\Delta = 3.2$ meV and 6.4 meV, the oscillation period of the gap is about 340 nm and 170 nm, respectively.

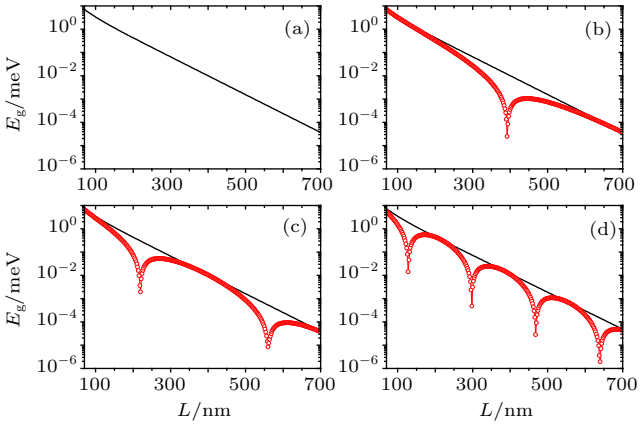


Fig. 2. (color online) The width dependence of the energy gap (E_g) of the edge states for different strength Δ with $R = 0$: (a) $\Delta = 0$; (b) $\Delta = 1.6$ meV; (c) $\Delta = 3.2$ meV; (d) $\Delta = 6.4$ meV. The (black) solid lines present the energy separation between conduction- and valence-band edges at $k = 0$. The (red) open circle symbol presents the energy gap of the edge states due to the finite size effect with the BIA term.

Aiming to explain the origin of the oscillatory decay of the gap E_g of the edge states in HgTe QWs with the BIA, we reconsider the effective four-band Hamiltonian (1) (neglecting the SIA term, i.e., $R = 0$) for a finite strip geometry of width L in the y direction. We keep the periodic boundary condition in the x direction. Owing to the open boundary condition in the y direction, we perform a Peierls substitution $k_y = -\mathrm{i}\partial_y$. By substituting a four-component trial solution $\psi = \psi_\lambda \mathrm{e}^{\lambda y}$ into the Schrödinger equation $\boldsymbol{H}(k, -\mathrm{i}\partial_y)\psi = E\psi$ (where E is the eigenvalue of energy), from the secular equation

$|\boldsymbol{H}(k,-\mathrm{i}\lambda)-E|=0$, we obtain the following equation of $\lambda(E)$:

$$0 = \left[\Delta^2-\left(E^2-M^2\right)-\left(2ED+2MB-A^2\right)\left(k^2-\lambda^2\right) + \left(B^2-D^2\right)\left(k^2-\lambda^2\right)^2\right]^2 - 4A^2\Delta^2\left(k^2-\lambda^2\right). \quad (7)$$

By solving Eq. (7), we can obatain eight solutions of $\lambda(E)$, denoted as $\pm\lambda_\alpha(E)$ ($\alpha=1,2,3,4$). The analytic expressions of $\lambda_\alpha(E)$ are too complex to be given here. Using the material parameters given above and numerical calculations, when the BIA term is taken into account and Δ is taken as a finite value, we can find that all λ_α are complex solutions for eigenenergy E of the edge states lying within the gap of the bulk (i.e., $-|M|<E<|M|$, here C is taken as zero). Moreover, four complex solutions λ_α satisfy $\lambda_1=\lambda_3^*$ and $\lambda_2=\lambda_4^*$. The real part of the complex λ_α (denoted as $\mathrm{Re}(\lambda_\alpha)$) corresponds to the inverse localization length of the edge states. The density distributions of the edge states located at the two boundaries of the strip sample are mainly determined by the larger one of the length scales $[\mathrm{Re}(\lambda_1)]^{-1}$ and $[\mathrm{Re}(\lambda_2)]^{-1}$. Here, we assume that $[\mathrm{Re}(\lambda_2)]^{-1}>[\mathrm{Re}(\lambda_1)]^{-1}$. When Δ in Eq. (7) is zero, that is, the BIA term is neglected, all λ_α of the edge states are real, then the decay pattern of the gap E_g of the edge states in HgTe QWs satisfies an approximate exponential law of L.[21,23] Just because not all λ_α of the edge states are real when the strength Δ is a finite value, we can observe a superimposed oscillatory pattern on the exponential decay of the gap E_g of the edge states in HgTe QWs with the BIA in Fig. 2. Thus the origin of the oscillatory decay here is consistent with that in 3D TI Bi_2Se_3 thin films.[23–26] According to the argument of Shan *et al*,[26] the oscillation period of the gap E_g with the width L in HgTe QWs with BIA is determined by the imaginary part of λ_2 (denoted as $\mathrm{Im}(\lambda_2)$), i.e., the oscillation period here is $\pi[\mathrm{Im}(\lambda_2)]^{-1}$, as can be verified by the results of numerical calculations.

The energy spectra of the edge states of HgTe QWs with $\Delta=6.4$ meV for different width L are plotted in Fig. 3. When the width of the HgTe QW is wide enough (e.g., $L=1000$ nm, see Fig. 3(f)), the effect of the BIA on the energy spectrum of the edge states is negligibly small; while for a narrow stripe, the effect of the BIA term is indeed worth considering.

Now we proceed to study the role of the SIA in the finite size effect on the helical edge states of the HgTe QWs. In fact, the analytical expressions for the wave functions of the edge states have been derived in the absence of BIA and SIA terms by Zhou *et al*.[21] We reexamine those wave functions of the edge states, and find that the components of the electron-like subbands ($|E_1\pm\rangle$) in the edge states are significantly smaller than those of the heavy hole subbands ($|H_1\pm\rangle$).[32] This result can be confirmed by numerical diagonalization of Hamiltonian (4). As an example, the density distributions of the edge states for $L=200$ nm are plotted for demonstration in Fig. 4. Since the Rashba term in Hamiltonian (1) describes the mixing between $|E_1+\rangle$ and $|E_1-\rangle$ and the strength R can be tuned by a gate voltage, we argue that only when R is tuned to a sufficiently large value, one might have an observable effect of the SIA on the finite size effect on the edge states.

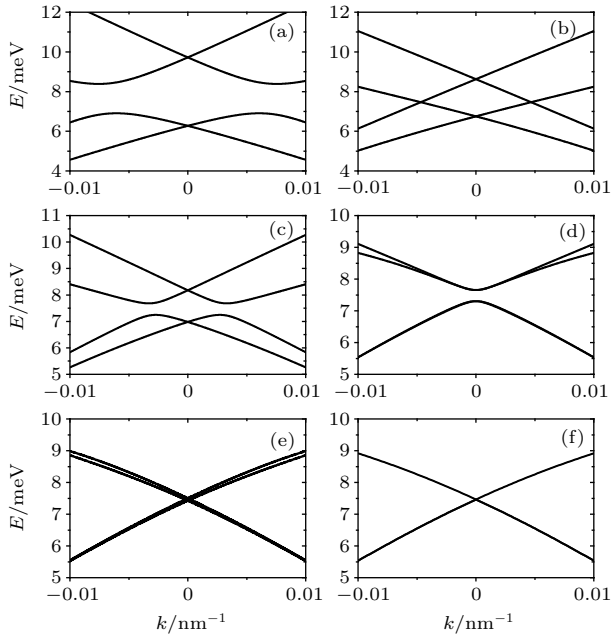


Fig. 3. Energy spectra of the edge states of HgTe QWs with $\Delta=6.4$ meV and different width L obtained by numerical diagonalization of Hamiltonian (4) without the SIA term (i.e., $R=0$): (a) $L=100$ nm; (b) $L=128$ nm; (c) $L=150$ nm; (d) $L=214$ nm; (e) $L=298$ nm; (f) $L=1000$ nm.

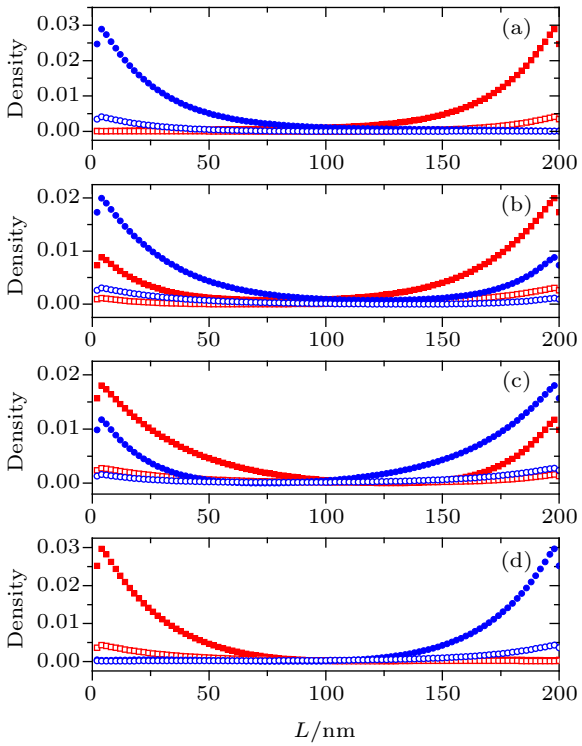


Fig. 4. (color online) Density distributions of the edge states of the HgTe QW with width $L=200$ nm obtained by numerical diagonalization of Hamiltonian (4). Here $k=0.001$ nm^{-1}, $\Delta=1.6$ meV, $R=160$ meV·nm, and the other model parameters are given in the text. The (red) open squares, (red) filled squares, (blue) open circles, and (blue) filled circles correspond to the components of $|E_1+\rangle$, $|H_1+\rangle$, $|E_1-\rangle$, and $|H_1-\rangle$ subbands in the edge-state wave function, respectively. Panels (a)–(d) are arranged in order from the highest to the lowest eigenenergy of the edge states.

When the BIA and the SIA are simultaneously taken into account, the size-dependence of the energy gap E_g of the edge states is plotted in Fig. 5 with the open square symbols, where $R = 160$ meV·nm and Δ takes four values, i.e., $\Delta = 1$ meV, 1.6 meV, 3.2 meV, and 6.4 meV, respectively. For the purpose of demonstration, the case without the SIA term is also plotted in Fig. 5 with the open circle symbols. From Fig. 5, it is seen that due to the SIA, the energy gap E_g vanishes at a narrower width and with increasing Δ this variation decreases. For example, for the case of $\Delta = 1$ meV, when $R = 0$ meV·nm and 160 meV·nm, the energy gap E_g vanishes at $L = 598$ nm and 502 nm, respectively; while for the case of $\Delta = 1.6$ meV, when $R = 0$ meV·nm and 160 meV·nm, the energy gap E_g vanishes at $L = 392$ nm and 362 nm, respectively. Specially, it is noted that for the case of $\Delta = 6.4$ meV, the effect of the SIA term with $R = 160$ meV·nm is nearly suppressed.

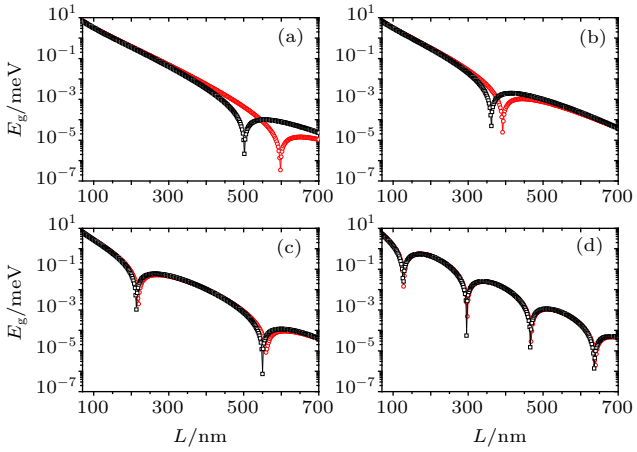


Fig. 5. (color online) The width dependence of the energy gap E_g of the edge states for different strength Δ with $R = 160$ meV·nm: (a) $\Delta = 1$ meV; (b) $\Delta = 1.6$ meV; (c) $\Delta = 3.2$ meV; (d) $\Delta = 6.4$ meV. The (red) open circle symbol and the (black) open square symbol correspond to the case without and with the SIA term, respectively.

When the edge state has an energy gap due to the finite size effect, the charge conductance of the edge states in a strip geometry is modified and has a form following the Landauer–Büttiker formula[21,39]

$$G = \frac{2e^2}{h}\left[\frac{1}{e^{[(E_g/2)-\mu]/k_BT}+1} - \frac{1}{e^{[-(E_g/2)-\mu]/k_BT}+1} + 1\right].$$

In Fig. 6, we plot the conductance G as a function of the chemical potential μ and the strength R of the SIA for $L = 350$ nm at temperature $T = 0.03$ K. Here Δ is taken to be 1.6 meV. When the chemical potential μ lies near the center of the gap E_g (i.e., $\mu = 0$), the nonmonotonous variation of G with R is seen. The main reason is that the variation of R induces a shift of the energy gap E_g closing point shown in Fig. 5(b). According to the result given in Ref. [29], the strength R of the SIA is taken to be $R \simeq -15.6\ \text{nm}^2 \times eE_z$, where E_z is the electric field perpendicular to the HgTe QW plane. For the case of $R = 500$ meV·nm shown in Fig. 6, we can estimate that the corresponding applied electric field is about 32 mV/nm. In Ref. [32], Virtanen *et al* proposed that an upper limit for the electric field that can be applied in practice is likely to be in the order of $E_z \sim 100$ mV/nm. They showed that for fields larger than this upper limit, the potential difference across the HgTe QW becomes comparable to the energy gap of the barrier material (CdTe).[32] Thus, based on the above numerical estimate, we argue that the variation of the strength R from 0 meV·nm to 500 meV·nm in Fig. 6 can be realized by tuning the applied electric field E_z in practice. On the other hand, the gate dependent shift of the Fermi level in the HgTe QW has been reported.[2] For the gate voltage V_g being within -1.9 V $< V_g < -1.4$ V, the Fermi level is adjusted in the bulk gap of the HgTe QW.[2] The strength R can be tuned by a gate voltage, however, tuning the gate voltage adjusts the asymmetry of the HgTe QW and the Fermi level simultaneously. Thus we need two gates to tune the strength R of the SIA and the Fermi level independently.[12] Note that such a dual-gate geometry has already been realized experimentally in an InAs/GaSb/Alsb QW.[12,40] If such a dual-gate geometry could also be realized experimentally in HgTe QWs, the phenomenon observed in Fig. 6 will be useful for realizing a QSH FET. When the Fermi level lies inside the gap E_g of the edge states due to the finite size effect, the conductance channel via the edge states is blocked, and thus the FET is in an OFF state. When the gate voltage is tuned to close the gap E_g due to the variation of the strength R of the SIA in HgTe QWs, the FET can be flipped to the ON state, where the current is carried only by the helical edge states. The realization of such a QSH FET needs very delicate experiments.

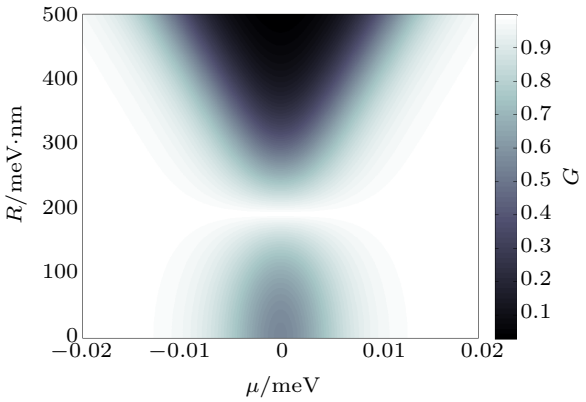


Fig. 6. (color online) The variation of conductance G as a function of the chemical potential μ and the strength R of the SIA for $L = 350$ nm at temperature $T = 0.03$ K. Here the strength Δ of the BIA is taken to be 1.6 meV, the unit of G is $2e^2/h$, and brighter color represent a larger G.

4. Conclusion

In a strip 2D QSH system (or a 3D TI thin film), the edge (or surface) states on the two different boundaries of the sample can hybridize to create an energy gap E_g. This finite size effect is an important issue that should be considered carefully

in potential applications of 2D QSH systems (or 3D TIs). In this paper, we focus on the effects of the SOC induced by BIA and SIA on the finite size effects in the inverted HgTe QW with finite width L. We firstly examine the role of the BIA. It is shown that different from the previous result of the case without the BIA, the energy gap E_g of the edge states due to the finite size effect features an oscillating exponential decay of the strip width of the HgTe QWs. The phenomenon is very similar to the scenario occurring in 3D TI thin films. Further taking the SIA into account, we find that the SIA can induce a shift of the energy gap E_g closing point. The strength of the SIA can be continuously tuned by a gate voltage, thus a way to realize a QSH FET is proposed based on the roles of BIA and SIA on the finite size effects in the inverted HgTe QWs. We notice that the effective four-band Hamiltonian (1) can also characterize the inverted InAs/GaSb/Alsb QW which has stronger BIA than the HgTe QW,[12] and moreover the helical edge modes of the QSH state in the inverted InAs/GaSb/Alsb QW have been confirmed experimentally.[41] Recently the gate tunable finite size effect in 3D TI Sb_2Te_3 thin films has been observed experimentally.[42] Thus we expect that the gate tunable finite size effect could also be observed experimentally in the inverted HgTe QWs or other 2D QSH systems.

References

[1] Bernevig B A, Hughes T L and Zhang S C 2006 *Science* **314** 1757
[2] König M, Wiedmann S, Brüne C, Roth A, Buhmann H, Molenkamp L W, Qi X L and Zhang S C 2007 *Science* **318** 766
[3] Roth A, Brüne C, Buhmann H, Molenkamp L W, Maciejko J, Qi X L and Zhang S C 2009 *Science* **325** 294
[4] Brüne C, Roth A, Buhmann H, Hankiewicz E M, Molenkamp L W, Maciejko J, Qi X L and Zhang S C 2012 *Nature Phys.* **8** 485
[5] Nowack K C, Spanton E M, Baenninger M, König M, Kirtley J R, Kalisky B, Ames C, Leubner P, Brüne C, Buhmann H, Molenkamp L W, Goldhaber-Gordon D and Moler K A 2013 *Nature Mater.* **12** 787
[6] König M, Baenninger M, Garcia A G F, Harjee N, Pruitt B L, Ames C, Leubner P, Brüne C, Buhmann H, Molenkamp L W and Goldhaber-Gordon D 2013 *Phys. Rev. X* **3** 021003
[7] Ma Y, Kundhikanjana W, Wang J, Calvo M R, Lian B, Yang Y, Lai K, Baenninger M, König M, Ames C, Brüne C, Buhmann H, Leubner P, Tang Q, Zhang K, Li X, Molenkamp L W, Zhang S C, Goldhaber-Gordon D, Kelly M A and Shen Z X 2012 arXiv:1212.6441v1 [cond-mat]
[8] Kane C L and Mele E J 2005 *Phys. Rev. Lett.* **95** 146802
[9] Kane C L and Mele E J 2005 *Phys. Rev. Lett.* **95** 226801
[10] Bernevig B A and Zhang S C 2006 *Phys. Rev. Lett.* **96** 106802
[11] Murakami S 2006 *Phys. Rev. Lett.* **97** 236805
[12] Liu C X, Hughes T L, Qi X L, Wang K and Zhang S C 2008 *Phys. Rev. Lett.* **100** 236601
[13] Hasan M Z and Kane C L 2010 *Rev. Mod. Phys.* **82** 3045
[14] Qi X L and Zhang S C 2011 *Rev. Mod. Phys.* **83** 1057
[15] Yang W M, Lin C J, Liao J and Li Y Q 2013 *Chin. Phys. B* **22** 097202
[16] He K, Ma X C, Chen X, Lü L, Wang Y Y and Xue Q K 2013 *Chin. Phys. B* **22** 067305
[17] Wang J and Zhu B F 2013 *Chin. Phys. B* **22** 067301
[18] Yang Y Y, Xu Z, Sheng L, Wang B G, Xing D Y and Sheng D N 2011 *Phys. Rev. Lett.* **107** 066602
[19] Sheng L, Li H C, Yang Y Y, Sheng D N and Xing D Y 2013 *Chin. Phys. B* **22** 067201
[20] Guo H M and Feng S P 2012 *Chin. Phys. B* **21** 077303
[21] Zhou B, Lu H Z, Chu R L, Shen S Q and Niu Q 2008 *Phys. Rev. Lett.* **101** 246807
[22] Brüne C, Roth A, Novik E G, König M, Buhmann H, Hankiewicz E M, Hanke W, Sinova J and Molenkamp L W 2010 *Nature Phys.* **6** 448
[23] Linder J, Yokoyama T and SudbøA 2009 *Phys. Rev. B* **80** 205401
[24] Liu C X, Zhang H J, Yan B H, Qi X L, Frauenheim T, Dai X, Fang Z and Zhang S C 2010 *Phys. Rev. B* **81** 041307
[25] Lu H Z, Shan W Y, Yao W, Niu Q and Shen S Q 2010 *Phys. Rev. B* **81** 115407
[26] Shan W Y, Lu H Z and Shen S Q 2010 *New J. Phys.* **12** 043048
[27] Zhang Y, He K, Chang C Z, Song C L, Wang L L, Chen X, Jia J F, Fang Z, Dai X, Shan W Y, Shen S Q, Niu Q, Qi X L, Zhang S C, Ma X C and Xue Q K 2010 *Nature Phys.* **6** 584
[28] König M, Buhmann H, Molenkamp L W, Hughes T L, Liu C X, Qi X L and Zhang S C 2008 *J. Phys. Soc. Jpn.* **77** 031007
[29] Rothe D G, Reinthaler R W, Liu C X, Molenkamp L W, Zhang S C and Hankiewicz E M 2010 *New J. Phys.* **12** 065012
[30] Krueckl V and Richter K 2012 *Semicond. Sci. Technol.* **27** 124006
[31] Takagaki Y 2012 *J. Phys.: Condens. Matter* **24** 435301
[32] Virtanen P and Recher P 2012 *Phys. Rev. B* **85** 035310
[33] Zou Y L, Zhang L B and Song J T 2013 *J. Phys.: Condens. Matter* **25** 075801
[34] Krueckl V and Richter K 2011 *Phys. Rev. Lett.* **107** 086803
[35] Takagaki Y 2012 *Phys. Rev. B* **85** 155308
[36] Michetti P, Penteado P H, Egues J C and Recher P 2012 *Semicond. Sci. Technol.* **27** 124007
[37] Ostrovsky P M, Gornyi I V and Mirlin A D 2012 *Phys. Rev. B* **86** 125323
[38] Weithofer L and Recher P 2013 *New J. Phys.* **15** 085008
[39] Datta S 1995 *Electronic Transport in Mesoscopic Systems* (Cambridge: Cambridge University Press)
[40] Cooper L J, Patel N K, Drouot V, Linfield E H, Ritchie D A and Pepper M 1998 *Phys. Rev. B* **57** 11915
[41] Knez I, Du R R and Sullivan G 2011 *Phys. Rev. Lett.* **107** 136603
[42] Zhang T, Ha J, Levy N, Kuk Y and Stroscio J 2013 *Phys. Rev. Lett.* **111** 056803