Vacancy-induced bound states in topological insulators

Wen-Yu Shan, Jie Lu, Hai-Zhou Lu, and Shun-Qing Shen
Department of Physics, The University of Hong Kong, Pokfulam Road, Hong Kong
(Received 24 May 2011; published 20 July 2011)

We present an exact solution of a modified Dirac equation for topological insulator in the presence of a hole or vacancy to demonstrate that vacancies can induce bound states in the band gap of topological insulators. They arise due to the $\mathbb{Z}_2$ classification of time-reversal invariant insulators. Coexistence of the in-gap bound states and the edge or surface states in topological insulators suggests that imperfections may affect transport properties of topological insulators via additional bound states near the system boundary.

DOI: 10.1103/PhysRevB.84.035307 PACS number(s): 73.20.Hb, 73.43.−f

I. INTRODUCTION

Topological insulators are narrow-band semiconductors with band inversion generated by strong spin-orbit coupling.[1] They are distinguished from the ordinary band insulators according to the $\mathbb{Z}_2$ invariant classification of the band insulators that respect time-reversal symmetry. The variation of the $\mathbb{Z}_2$ invariant at their boundaries will lead to the topologically protected edge or surface states with the gapless Dirac energy spectrum.[2–7] Imperfections, such as impurity, vacancy, and disorder, are inevitably present in topological insulators. Owing to the time-reversal symmetry, an exciting feature of topological insulator is that its boundary states are expected to be topologically protected against weak nonmagnetic impurities or disorders.[8,9] This provoked much interest on the single impurity problem on the surface of a topological insulator, starting with gapless Dirac model.[10–14] However, reminding that the boundary state is only a manifestation of the topological nature of bulk bands, one should also start with the examination of the host bulk to know how the imperfections affect the electronic structure. It is well known that single impurity or defect can induce bound states in many systems, such as the Yu-Shiba state in s-wave superconductor[15,16] and in d-wave superconductors.[17] Topological defects were discussed in the B-phase of ^{3}He superfluid[18] and topological insulators and superconductors.[19] Here we report that bound states can form around a single vacancy in the bulk energy gap of topological insulators. These bound states are found to have the same origin as boundary states due to the $\mathbb{Z}_2$ classification.

The formation of the in-gap bound states can be readily illustrated by reviewing the quantum spin Hall effect in two-dimensional (2D) topological insulators,[20–22] in which strong spin-orbit coupling twists the bulk conduction and valence bands, leading to a nontrivial $\mathbb{Z}_2$ index. As the $\mathbb{Z}_2$ index varies across the edge, edge states arise in the gap with the gapless Dirac dispersion. Unlike the quantum Hall effect in a magnetic field, spin-orbit coupling respects the time-reversal symmetry, so the resulting edge states appear in helical pairs, of which one state is the time-reversal counterpart of the other, propagating along opposite directions and with opposite spins [Fig. 1(b)]. Now imagine that the system edge is bent into a hole, the edge states will circulate around the hole as the periodic boundary conditions along the propagating direction remain unchanged [Fig. 1(d)]. While shrinking the radius of the hole, most of the edge states will be expelled into the bulk bands as the energy separation among the states becomes larger and larger, and it is found that at least two degenerate pairs of the states will be trapped to form the bound states in the gap as the hole evolves into a point defect. This mechanism of the formation of the bound states can be realized in topological insulator in all dimensions.

The paper is organized as follows. In Sec. II, we review the modified Dirac model, which can be used to characterize topological insulators. In Sec. III, we show explicitly the existence of in-gap bound states in the presence of vacancy in 2D topological insulators, and that the bound states can be manipulated by a magnetic flux threading the vacancy. In Sec. IV, the generalization to three-dimensional (3D) topological insulators is discussed with the bound-state solutions presented. Finally, in Sec. V, we discuss the possible implications of these in-gap bound states on transport properties.

II. MODIFIED DIRAC MODEL

We will employ a modified Dirac model to provide a unified description of topological insulators in various dimensions

$$H_0 = v\mathbf{p}\cdot\boldsymbol{\alpha} + (mv^2 - Bp^2)\beta. \tag{1}$$

The modification comes from the quadratic correction in momentum $-Bp^2$ to the band gap mv^2 term. $p_i = -i\hbar\partial_i$ is the momentum operator ($i \in \{x,y,z\}$), $p^2 = p_x^2 + p_y^2 + p_z^2$, v and m have the dimensions of the speed and mass, respectively. B has the dimension of m^{-1}. The Dirac matrices satisfy the anticommutation relations $\alpha_i\alpha_j = -\alpha_j\alpha_i$ ($i \neq j$), $a_i\beta = -\beta\alpha_i$, and $\alpha_i^2 = \beta^2 = 1$. One representation of the Dirac matrices in three spatial dimensions can be expressed as a set of 4×4 matrices

$$\alpha_i = \sigma_x \otimes \sigma_i, \qquad \beta = \sigma_z \otimes \sigma_0, \tag{2}$$

where $\sigma_{i=x,y,z}$ are the Pauli matrices, σ_0 is the 2×2 unit matrix, and $\otimes$ represents the Kronecker product. This modified Dirac Hamiltonian preserves the time-reversal symmetry $\hat{\Theta}H_0\hat{\Theta}^{-1} = H_0$ under the time-reversal operation $\hat{\Theta} = -i\alpha_x\alpha_z\hat{K}$, where $\hat{K}$ is the complex conjugate operator. This model has the identical mathematical structure as the effective models for the quantum spin Hall effect and 3D topological insulator.[21,23–25] Following Kane and Mele for a lattice model,[2] we can study the $\mathbb{Z}_2$ classification of the modified Dirac model in the continuous limit. The $\mathbb{Z}_2$ index can be obtained by counting the number of pairs of complex zeros of

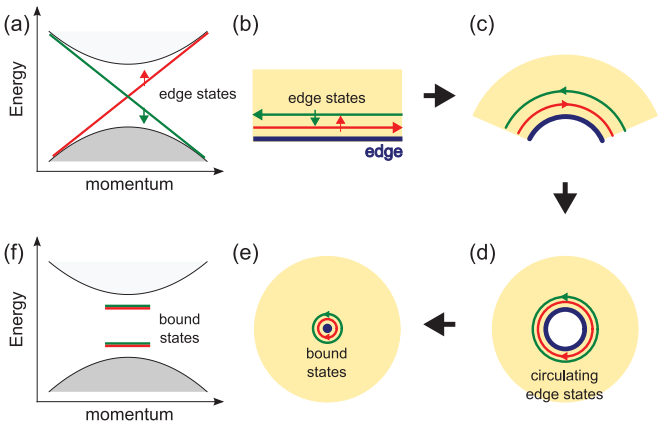

FIG. 1. (Color online) Schematic description of the formation of vacancy-induced in-gap bound states in 2D topological insulators. (a) and (b) A pair of helical edge states traveling along the edge of a 2D topological insulator with the gapless Dirac dispersion. (c) and (d) When the edge is bent into a hole, the helical edge states evolve to circulate around the hole. (e) and (f) The circulating edge states may develop into bound states as the hole shrinks into a point or being replaced by a vacancy. The same physics is expected to happen in one and three dimensions.

the Pfaffian Pf$[A(\mathbf{k})]$,

$$\mathrm{Pf}[A(\mathbf{k})] = \frac{1}{2^n n!}\sum_P \mathrm{sgn}(P)\prod_{i=1}^{n} A_{P(2i-1)P(2i)}, \qquad (3)$$

in which P refers to all permutations of $\{1,\ldots,2n\}$, and $A(\mathbf{k})$ is a $2n$-order antisymmetric matrix defined by the overlaps of time reversal

$$A_{ij}(\mathbf{k}) = \langle u_i(\mathbf{k})|\Theta|u_j(\mathbf{k})\rangle, \qquad (4)$$

with i and j running over all the bands below the Fermi surface. After some algebra,[26] the Pfaffian is found to have the form,

$$\mathrm{Pf}[A(\mathbf{k})] = \frac{mv^2 - Bp^2}{\sqrt{(mv^2 - Bp^2)^2 + v^2p^2}}. \qquad (5)$$

Given an isotropic system, the Pfaffian Pf$[A(\mathbf{k})]$ has odd pairs of zeros only when $mB > 0$. Therefore, the modified Dirac model becomes topologically nontrivial if $mB > 0$, and topologically trivial if $mB < 0$. Similar property also lives in the effective model of the B-phase of ^{3}He superfluidity.[27] Due to the bulk-boundary correspondence of topological insulator,[28,29] there always exist topologically protected boundary states at the open boundaries, where the $\mathbb{Z}_2$ invariant changes from nontrivial to trivial. This feature can be well described by the modified Dirac model when $mB > 0$. Starting from this modified Dirac model, we are now ready to explore the existence of the in-gap bound states induced by a single vacancy by presenting an exact solution to the modified Dirac model in the presence of vacancy boundary conditions.

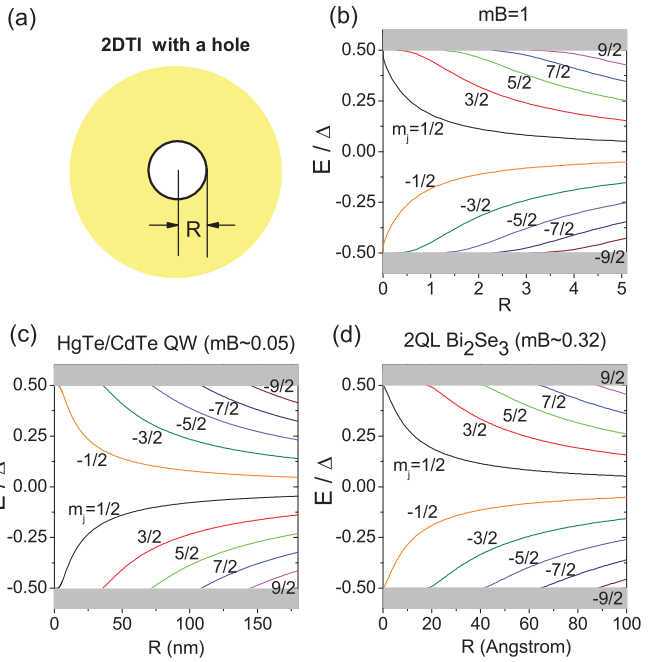

FIG. 2. (Color online) Two-dimensional in-gap bound states. (a) A 2D topological insulator with a hole of radius R at the center. (b) and (c) Energies (E in units of the band gap Δ) of in-gap bound states circulating around the hole as functions of the hole radius. m_j is the quantum number for the z-component of the total angular momentum of the circulating bound states. In (b), $m = v = B = \hbar = 1$; in (c), $mv^2 = -10$ meV, $B\hbar^2 = -686$ meV nm^2, and $\hbar v = 364.5$ meV nm, adopted from Ref. 21; in (d), $mv^2 = 0.126$ eV, $B\hbar^2 = 21.8$ eV Å^2, $\hbar v = 2.94$ eV Å, adopted from Ref. 30. $\Delta = 2mv^2$ for $0 < mB < 1/2$, and $\Delta = (v^2/|B|)\sqrt{4mB - 1}$ for $mB > 1/2$.The gray areas mark the bulk bands.

III. VACANCY IN 2D TOPOLOGICAL INSULATORS

A. Vacancy without magnetic flux

In two dimensions ($p_z = 0$), the modified Dirac model can be reduced into two independent 2×2 Hamiltonians

$$h_\pm = (mv^2 - Bp^2)\sigma_z + \hbar v(p_x\sigma_x \pm p_y\sigma_y), \qquad (6)$$

with h_- the time-reversal counterpart of h_+.[21,24,25] It is convenient to adopt polar coordinates $(x,y) = r(\cos\varphi, \sin\varphi)$ in two dimensions. Here these equations are solved under the vacancy boundary conditions [Fig. 2(a)], i.e., the center of the 2D topological insulator is punched with a hole of radius R, thus the wave function is required to vanish at $r = R$ and $r = +\infty$. Due to the rotational symmetry of h_+, the z-component of the total angular momentum $j_{z+} = -i\hbar\partial_\theta + (\hbar/2)\sigma_z$ provides a good quantum number, labeled by a half-integer $m_j \in \{\pm 1/2, \pm 3/2, \ldots\}$, which can be used to characterize the bound states. In this way, the equation is reduced to a set of one-dimensional (1D) radial equations, which can be solved exactly. The trial wave function has the form $(\psi_1, \psi_2)^{\mathrm{T}} e^{-\lambda r}$. The secular equations of the undetermined coefficients $(\psi_1, \psi_2)^{\mathrm{T}}$ give four roots of λ_n $(= \pm\lambda_1, \pm\lambda_2)$ as

functions of E,

$$\lambda_{1,2}^2 = \frac{v^2}{2B^2\hbar^2}\left[1 - 2mB \pm \sqrt{1 - 4mB + \frac{4B^2E^2}{v^4}}\right]. \quad (7)$$

Using the boundary conditions at $r = R$ and $r = +\infty$, we finally arrive at the transcendental equation for the bound-state energies

$$\frac{\lambda_1^2 + \frac{mv^2-E}{B\hbar^2}}{\lambda_1}\frac{K_{m_j+\frac{1}{2}}(\lambda_1 R)}{K_{m_j-\frac{1}{2}}(\lambda_1 R)} = \frac{\lambda_2^2 + \frac{mv^2-E}{B\hbar^2}}{\lambda_2}\frac{K_{m_j+\frac{1}{2}}(\lambda_2 R)}{K_{m_j-\frac{1}{2}}(\lambda_2 R)}, \quad (8)$$

and the wave function $\Psi^+_{m_j}(r,\theta)$ for h_+ turns out to have the form

$$\begin{bmatrix} \frac{K_{m_j-\frac{1}{2}}(\lambda_1 R)}{K_{m_j+\frac{1}{2}}(\lambda_1 R)} f_{m_j-\frac{1}{2}}(r) e^{i(m_j-\frac{1}{2})\theta} \\ i\frac{\lambda_1^2 + \frac{mv^2-E}{B\hbar^2}}{(\lambda_1 v/B\hbar)} f_{m_j+\frac{1}{2}}(r) e^{i(m_j+\frac{1}{2})\theta} \end{bmatrix}, \quad (9)$$

with

$$f_{m_j\pm\frac{1}{2}}(r) = \frac{K_{m_j\pm\frac{1}{2}}(\lambda_1 r)}{K_{m_j\pm\frac{1}{2}}(\lambda_1 R)} - \frac{K_{m_j\pm\frac{1}{2}}(\lambda_2 r)}{K_{m_j\pm\frac{1}{2}}(\lambda_2 R)}, \quad (10)$$

where $K_n(x)$ is the modified Bessel function of second kind. The solution for h_- can be derived following the same procedure, and here we use the symmetry analysis to find the solution. The z-component of the total angular momentum $j_{z-} = -i\hbar\partial_\theta - (\hbar/2)\sigma_z$ commutes with h_-, and we label this good quantum number with $m_j' \in \{\pm 1/2, \pm 3/2, \ldots\}$. The time-reversal pair $\Psi^+_{m_j}(r,\theta)$ and $\Psi^-_{m_j'}(r,\theta)$ must have opposite angular momenta, i.e., $m_j' = -m_j$. Substituting this relation into Eq. (8), we obtain the transcendental equation:

$$\frac{\lambda_1^2 + \frac{mv^2-E}{B\hbar^2}}{\lambda_1}\frac{K_{m_j'-\frac{1}{2}}(\lambda_1 R)}{K_{m_j'+\frac{1}{2}}(\lambda_1 R)} = \frac{\lambda_2^2 + \frac{mv^2-E}{B\hbar^2}}{\lambda_2}\frac{K_{m_j'-\frac{1}{2}}(\lambda_2 R)}{K_{m_j'+\frac{1}{2}}(\lambda_2 R)}. \quad (11)$$

The form of wave function $\Psi^-_{m_j'}(r,\theta)$ can be found by acting time-reversal operator Θ on $\Psi^+_{m_j}(r,\theta)$.

In Figs. 2(c)–2(d), we show the bound-state energies as functions of R for an ideal case [Fig. 2(b), $mB = 1$], for the HgTe quantum well [Fig. 2(c), $mB = 0.05$],[21] and for a two-quintuple layer of Bi_2Se_3 thin film [Fig. 2(d), $mB = 0.32$].[30] For a macroscopically large R, we found an approximated solution for the energy spectrum of h_+ as $E = m_j\hbar v\,\mathrm{sgn}(B)/R$. As the time-reversal copy of h_+, h_- has an approximated spectrum $E = -m_j\hbar v\,\mathrm{sgn}(B)/R$. They form a series of paired helical edge states, in good agreement with the edge-state solutions in the 2D quantum spin Hall system[31] if we take $k = m_j/R$ for a large R. When shrinking R, the energy separation of these edge states $\Delta E = \pm\hbar v/R$ increases with shrunk R, and the edge states with higher m_j will be pushed out of the energy gap gradually. However, we observe that for $mB > 0$, two pairs of states with $m_j = \pm 1/2$ always stay in the energy gap, and as $R \to 0$, their energies approach to $E = \pm(v^2/2|B|)\sqrt{4mB - 1}$ for $mB > 1/2$ or $\pm mv^2$ for $0 < mB < 1/2$ (a detailed proof is given in Appendix A). When comparing the details of Fig. 2(c) with Fig. 2(d), we find that the two pairs of states for $m_j = \pm 1/2$ have quite different asymptotic behaviors in the spectrum when R decreases to zero. This finding agrees with the analytical result obtained by taking small R expansion of the transcendental equation [Eq. (8)] for $0 < mB < 1/2$ (the proof is demonstrated in Appendix A),

$$E \approx mv^2\left[1 - \frac{\chi^2}{2}\exp(\chi C)(\Lambda_1 R)^\chi\right], \quad (12)$$

with $\chi = (1 - 2mB)/mB$, $\Lambda_1 = \frac{v}{|B|\hbar}\sqrt{1 - 2mB}$, and $C = -\ln 2 + \gamma$. γ is Euler's constant. The equation shows that for smaller mB, the E–R relation for $m_j = \pm 1/2$ tends to become flat before entering the bulk. This can be explained by noting the fact that there is no in-gap bound state when $mB < 0$, suggesting $mB = 0$ is the phase transition point. The bound state with smaller mB is closer to the transition point, and thus tends to enter the bulk more easily, which results in the quite flat E–R relation.

The solutions verify the formation of the in-gap bound states as shown in Fig. 1. Therefore, considering the symmetry between h_+ and h_-, we conclude that there always exist *at least* two pairs of bound states in the energy gap in the 2D quantum spin Hall system in the presence of vacancy.

B. Vacancy with magnetic flux

When a magnetic flux is applied through the hole, the energy levels of the bound states can be continuously manipulated. To consider a Φ flux that threads through the hole, we perform the Peierls substitution $\mathbf{p} \to \mathbf{p} + e\mathbf{A}$ in h_+ by taking the gauge $\mathbf{A} = (\Phi/2\pi r)\mathbf{e}_\theta$, which still keeps m_j a good quantum number. Therefore, the eigenfunctions of this new Hamiltonian can be readily expressed as $\exp(-i\nu\theta)\Psi^+_{m_j}(r,\theta)$ after a gauge transformation, with $\nu = \Phi/\Phi_0$ and the flux quantum $\Phi_0 = h/e$. In this case, the allowed value for m_j is no longer half-integer, but changed to $m_j = m_+ + \nu$, with the expectation value of j_{z+} being $m_+\hbar \in \{\pm 1/2, \pm 3/2, \ldots\}\hbar$. This is because that applying magnetic flux is equivalent to changing the boundary conditions. We have mentioned in the previous section that the approximated solution at large R for h_+ is $E \propto m_j/R$. When half-quantum flux $\nu = 1/2$ is introduced, m_j can be 0. This provides the possibility for the existence of zero-energy bound state. Similar analysis can be applied to h_- and one get $m_j' = m_- + \nu$, with the expectation value of j_{z-} being $m_-\hbar \in \{\pm 1/2, \pm 3/2, \ldots\}\hbar$. Half-quantum flux can also trap zero-energy bound state when we substitute the possible m_j' into Eq. (11), making the zero mode doubly degenerate. Numerical result for a 70 nm thick HgTe quantum well[21] is shown in Fig. 3. The existence of zero mode can be seen explicitly in Figs. 3(a) and 3(b) when half-quantum flux is introduced. In Fig. 3(a), the spectrum of h_+ with index m_j and h_- with index $-m_j$ overlaps, representing that the time-reversal symmetry is preserved. However, for arbitrary magnetic flux in Fig. 3(b), bound-state energies from block h_+ (black lines) and h_- (light orange lines) split off, as the time-reversal symmetry is broken by the magnetic flux. Compared with the bulk-state spectrum under magnetic flux,[32] we can see that the in-gap bound states from h_+ and h_- blocks move in opposite directions. The slopes of both series of curves are almost linear, supporting our approximated results that bound-state energy $E \propto m_j/R$ or $E \propto -m_j'/R$ for large

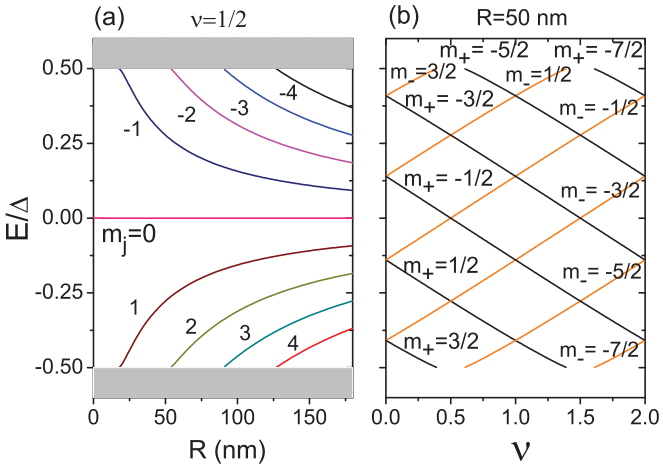

FIG. 3. (Color online) Effect of magnetic flux on in-gap bound states. Energies (E in units of the band gap Δ) of in-gap bound states circulating around the hole as functions of (a) the hole radius R when half-quantum flux $\nu = 1/2$ is applied and (b) the magnetic flux ν (in unit of flux quantum $\Phi_0 = h/e$) for fixed radius $R = 50$ nm. m_+ (m_-) is the quantum number for the z-component of the total angular momentum j_{z+} (j_{z-}) of the circulating bound states. $m_j = m_+ + \nu$. In (b), black (light orange) lines belong to h_+ (h_-) block. In (a) and (b), $mv^2 = -10$ meV, $B\hbar^2 = -686$ meV nm^2, and $\hbar v = 364.5$ meV nm, adopted from Ref. 21. $\Delta = 2mv^2$. The gray areas mark the bulk bands.

radius R. The existence of zero mode when half-quantum flux appears may not be surprising since it is well known that a vortex in a spinless p-wave superconductor can trap majorana zero mode[33] and we notice that the topological insulator shares a large similarity in the low-energy effective theory with $p_x + ip_y$ superconductor.[26] Some differences remain in the base space and symmetry. The space for superconductor is the Nambu space, whose redundancy makes the excitation of majorana fermions. However, in topological insulators the quasiparticles are Dirac fermions. In addition, spinless p-wave superconductor violates time-reversal symmetry, while topological insulator does not for half-quantum flux.

IV. VACANCY IN 3D TOPOLOGICAL INSULATORS

For 3D topological insulators,[23,34] the mechanism of the formation of the in-gap bound states is applicable. In 3D, the modified Dirac equation with a central potential becomes a classical problem, the hydrogen atom-like problem. For the Coulomb potential, it was exactly solved to give the fine structure of light spectra of hydrogen atom. Similarly, the eigenstates of the 3D modified Dirac equation with a central potential can be labeled by three good quantum numbers. The first two are the total angular momentum $\hat{\mathbf{J}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} + \frac{\hbar}{2}\hat{\mathbf{\Sigma}}$ and its z-component $\hat{J}_z$, where the spin operator $\hat{\mathbf{\Sigma}}_\alpha = \sigma_0 \otimes \sigma_\alpha$ ($\alpha = x, y, z$). The eigenvalues of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ are $j(j+1)\hbar^2$ and $m_j\hbar$, respectively, with $j \in \{\frac{1}{2}, \frac{3}{2}, \ldots\}$ and $m_j \in \{-j, \ldots, j\}$. The third conserved quantity is the spin-orbit operator $\hat{\kappa} = \beta(\hat{\mathbf{r}} \times \hat{\mathbf{p}} \cdot \hat{\Sigma} + \hbar)$. Note that $\hat{\kappa}^2 = \hat{\mathbf{J}}^2 + \hbar^2/4$, then the eigenvalues of $\hat{\kappa}$ is $\hbar\kappa = \pm\hbar(j + 1/2) = \pm\hbar, \pm 2\hbar, \ldots$. Thus, κ here is similar to the $\pm$ index that block diagonalize the Hamiltonian into $h_\pm$ in the 2D case. These conserved quantities also help to reduce the problem into a set of 1D radial equations.[35] In the presence of the vacancy or a cavity of radius R with the boundary conditions at $\Psi(R) = \Psi(\infty) = 0$, the radial part of the wave function can be solved in terms of the modified spherical Bessel function of the second kind $k_n(x)$. With the help of the recursive relation of $k_n(x)$, the transcendental equations for the bound state energies can be found as

$$\frac{\lambda_1^2 + \frac{mv^2 - E}{B\hbar^2}}{\lambda_1} \frac{k_{j\pm\frac{1}{2}}(\lambda_1 R)}{k_{j\mp\frac{1}{2}}(\lambda_1 R)} = \frac{\lambda_2^2 + \frac{mv^2 - E}{B\hbar^2}}{\lambda_2} \frac{k_{j\pm\frac{1}{2}}(\lambda_2 R)}{k_{j\mp\frac{1}{2}}(\lambda_2 R)}, \quad (13)$$

for $\kappa = j + \frac{1}{2}$ and $-(j + \frac{1}{2})$, respectively. The corresponding wave function $\Psi^{m_j}_{j,\kappa}(r,\theta,\phi)$ are of the form:

$$\Psi^{m_j}_{j,\kappa}(r,\theta,\phi) \propto \begin{bmatrix} \frac{i(\lambda_1 v/B\hbar)}{\lambda_1^2 + \frac{mv^2 - E}{B\hbar^2}} g_{j\mp\frac{1}{2}}(r) \phi^{A/B}_{j,m_j} \\ \frac{k_{j\pm\frac{1}{2}}(\lambda_1 R)}{k_{j\mp\frac{1}{2}}(\lambda_1 R)} g_{j\pm\frac{1}{2}}(r) \phi^{B/A}_{j,m_j} \end{bmatrix}, \quad (14)$$

where

$$g_{j\pm\frac{1}{2}}(r) = \frac{k_{j\pm\frac{1}{2}}(\lambda_1 r)}{k_{j\pm\frac{1}{2}}(\lambda_1 R)} - \frac{k_{j\pm\frac{1}{2}}(\lambda_2 r)}{k_{j\pm\frac{1}{2}}(\lambda_2 R)}, \quad (15)$$

$$\phi^A_{j,m_j}(\theta,\varphi) = \begin{bmatrix} \sqrt{\frac{j+m_j}{2j}} Y^{m_j - \frac{1}{2}}_{j-\frac{1}{2}}(\theta,\varphi) \\ \sqrt{\frac{j-m_j}{2j}} Y^{m_j + \frac{1}{2}}_{j-\frac{1}{2}}(\theta,\varphi) \end{bmatrix}, \quad (16)$$

$$\phi^B_{j,m_j}(\theta,\phi) = \begin{bmatrix} -\sqrt{\frac{j-m_j+1}{2(j+1)}} Y^{m_j - \frac{1}{2}}_{j+\frac{1}{2}}(\theta,\varphi) \\ \sqrt{\frac{j+m_j+1}{2(j+1)}} Y^{m_j + \frac{1}{2}}_{j+\frac{1}{2}}(\theta,\varphi) \end{bmatrix}, \quad (17)$$

and $Y^m_j(\theta,\varphi)$ is the spherical harmonics. ϕ^A_{j,m_j} and ϕ^B_{j,m_j} possess opposite parities.

Although the rotational symmetry simplifies the problem, it is believed that the presence of the bound states is not sensitive to the shape of the vacancy, because of their topological origin. As an example, we choose a set of isotropic parameters based on first-principle calculations for Bi_2Se_3, with $mv^2 = 0.28$ eV, $\hbar v = 3.2$ eV Å, and $B = 33$ eV Å^2. In this case, $mB \sim 1 > 1/2$. Similar to the 2D case, we find that the surface states around the cavity exist for a large radius R as expected by the bulk-boundary correspondence for a $\mathbb{Z}_2$-invariant topological insulators.[29] The states with larger orbital angular momentum are eventually expelled into the bulk band while the radius is shrinking. We plot several bound-state energies of small orbital angular momenta as a function of the radius R in Fig. 4. For convenience, the bound states are labeled by the quantum number κ for the spin-orbit operator. Each κ corresponds to $(2j + 1)$-fold degenerate states of different m_j. Note that when the vacancy radius is only several angstroms, two degenerate pairs of bound-state energies can survive. Detailed analysis of the solution indicates that the spatial distribution of a bound state is comparable with that of the edge or surface states (for a large R in the present case), which is determined by the model parameters and slightly depends on R. From the evolution of the edge or surface states into the in-gap bound states, we think their formation has the same topological origin.

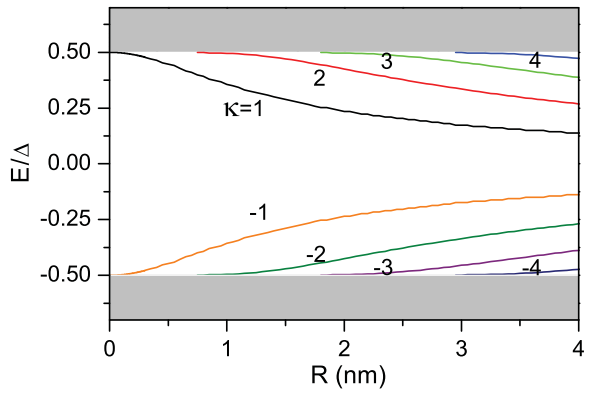

FIG. 4. (Color online) Three-dimensional in-gap bound states. Energies (E in units of the band gap Δ) of in-gap bound states surrounding a vacancy in a 3D topological insulator as functions of the vacancy radius R. κ is the quantum number of the spin-orbit operator. Parameters: $mv^2 = 0.28$ eV, $\hbar v = 3.2$ eV Å, and $B\hbar^2 = 33$ eV Å^2. $\Delta = (v^2/|B|)\sqrt{4mB - 1}$. The gray areas mark the bulk bands.

V. DISCUSSION

The model in Eq. (1) can be also applied to the B-phase of ^{3}He superfluidity, which is known to be topologically nontrivial, if one changes the notation mv^2 to $p_F^2/2m_3$, where m_3 is the mass of ^{3}He quasiparticle and p_F is the Fermi momentum. In liquids, there are no long-lasting vacancies, instead the defects in ^{3}He–B were modeled in forms of hard sphere, which cannot lead to zero electron wave function around the defect because of the Klein paradox of Dirac fermions. In other words, the hard sphere cannot approach the Dirichlet boundary conditions in this paper. Nevertheless, the bound states around impurities in ^{3}He–B are known for a long time. In the Dirac equation, the hard-sphere potential could generate a series of the bound states, which looks more like those of a hydrogen atom,[36] and may disappear under the continuous deformation of the potential. In this sense, the impurity problem of ^{3}He–B is distinct from this work.

Now we turn to possible implications of these solutions to topological insulators. Due to the overlapping in energy, when the vacancies are located close to the boundary, the induced in-gap bound states may sabotage the electronic transport through the boundary states. For a single hole or vacancy near the edge of the sample, when the in-gap bound state and the edge or surface states overlap in space, the distortion of their wave functions will cause change in their energies as well. As a result, we may also regard that there exist transition amplitudes between these states, and the helical edge states will be scattered by the in-gap bound states. If there is no other bound states in the bulk, the electrons in the edge states will not be further scattered away from the edge as what happens in the quantum Hall effect.[37] The situation will change if the concentration of the holes or vacancies is dense enough. In this case, the in-gap bound states will form an "impurity band" as in semiconductors and superconductors.[16,17] When the spatial size of the bound state is comparable with the average distance of the holes or vacancies, it becomes possible that the electrons are scattered from one edge of system to the other via the multiple scatterings by the in-gap bound states. In the HgTe/CdTe quantum wells,[22] a typical scale of the edge states and bound states is about 50–100 nm.[31] Thus, when the hole concentration is about $10^{10}/\text{cm}^2$, the in-gap bound states will destroy the quantum spin Hall effect. This may help to understand why the nonzero conductance is narrowed to a small region of gate voltage in the HgTe/CdTe quantum wells.

However, blessings usually come in disguise. The whole semiconductor business depends on how the positive and negative effects of impurities and vacancies are precisely balanced. The in-gap bound states for sure are essentially different from those we know before, as they are subjected to some topological nature and confined to a mesoscopic scale. Their possible impact and applications for topological insulators in future deserve further studies to explore.

ACKNOWLEDGMENTS

We would like to thank T. K. Ng for stimulating our interest on this topic. This work was supported by the Research Grant Council of Hong Kong under Grant Nos. HKU 7051/11P and HKUST3/CRF/09.

APPENDIX A: ASYMPTOTIC BEHAVIOR FOR SMALL RADIUS R

1. Leading-order expansion term

We only focus on $m_j = 1/2$, and the transcendental equation (8) becomes

$$\frac{\lambda_1^2 + \frac{mv^2 - E}{B\hbar^2}}{\lambda_1} \frac{K_1(\lambda_1 R)}{K_0(\lambda_1 R)} = \frac{\lambda_2^2 + \frac{mv^2 - E}{B\hbar^2}}{\lambda_2} \frac{K_1(\lambda_2 R)}{K_0(\lambda_2 R)}. \tag{A1}$$

By taking small R approximation and keeping only the leading term of R, we have

$$\frac{\lambda_2\left(\lambda_1^2 + \frac{mv^2 - E}{B\hbar^2}\right)}{\lambda_1\left(\lambda_2^2 + \frac{mv^2 - E}{B\hbar^2}\right)} \simeq 1, \tag{A2}$$

which gives $E = mv^2$ or $E = (v^2/2B)\sqrt{4mB - 1}$. Combining with the analysis in Appendix B, it is clear that for $m_j = 1/2$ branch, the bound state energy approaches to $E = (v^2/2B)\sqrt{4mB - 1}$ for $mB > 1/2$ or $E = mv^2$ for $0 < mB < 1/2$. The result for $m_j = -1/2$ branch can be obtained by reversing the sign of energy for the symmetry of the spectrum.

2. Higher-order expansion term

When higher-order expansion terms of R are kept, Eq. (A1) reduces to

$$\frac{\lambda_1^2 + \frac{mv^2 - E}{B\hbar^2}}{\lambda_1^2} \frac{1}{\ln(\lambda_1 R) + C} = \frac{\lambda_2^2 + \frac{mv^2 - E}{B\hbar^2}}{\lambda_2^2} \frac{1}{\ln(\lambda_2 R) + C}, \tag{A3}$$

where $C = -\ln 2 + \gamma$ and γ is Euler's constant. Here we assume that $m > 0$ and only focus on $m_j = 1/2$. (The result for $m < 0$ can be obtained by reversing the sign of energy and is the same for $m_j = -1/2$.) When $0 < mB < 1/2$, we

TABLE I. Possible cases of $\lambda_{1,2}$ in different ranges of energy E for parameters mB.

	$mB \in (-\infty,0)$	$mB \in (0,1/4)$	$mB \in (1/4,1/2)$	$mB \in (1/2,+\infty)$
$\lvert E\rvert \in [0, \frac{v^2}{2\lvert B\rvert}\sqrt{4mB-1})$	$\lambda_1>0, \lambda_2>0$	$\lambda_1>0, \lambda_2>0$	$\lambda_{1,2}=a\pm ib$	$\lambda_{1,2}=a\pm ib$
$\lvert E\rvert \in (\frac{v^2}{2\lvert B\rvert}\sqrt{4mB-1}, \lvert m\rvert v^2)$	$\lambda_1>0, \lambda_2>0$	$\lambda_1>0, \lambda_2>0$	$\lambda_1>0, \lambda_2>0$	$\lambda_1=i\eta_1, \lambda_2=i\eta_2$
$\lvert E\rvert \in (\lvert m\rvert v^2, \infty)$	$\lambda_1>0, \lambda_2=i\eta$	$\lambda_1>0, \lambda_2=i\eta$	$\lambda_1>0, \lambda_2=i\eta$	$\lambda_1>0, \lambda_2=i\eta$

substitute $E = mv^2 - \delta E$ into Eq. (A3) with $\delta E/mv^2 \ll 1$, and obtain an expression by neglecting terms comparable or higher-order smaller than $\delta E \ln(\delta E)$,

$$\ln(\Lambda_1 R) \backsimeq \frac{mB}{1-2mB}\ln(\delta E) + \frac{2mB}{1-2mB}\ln\frac{\sqrt{2m}B}{(1-2mB)v} - C, \tag{A4}$$

with $\Lambda_1 = \frac{v}{|B|\hbar}\sqrt{1-2mB}$. Let $\chi = (1-2mB)/mB$, Eq. (A4) reproduces Eq. (12).

APPENDIX B: DISCUSSION ON λ

From the definition of $\lambda_{1,2}$ in Eq. (7), we conclude all possible values of $\lambda_{1,2}$ for different parameters mB and different energy ranges in Table I. In this table, a, b, and η are positive. Considering that the density of wave distribution for bound states must vanish in infinite radius R, we can have bound states only if both λ_1 and λ_2 have real part. Thus, whatever boundary condition is given, the only possible energy range for the existence of bound states is $E \in (-|m|v^2, |m|v^2)$ for $mB \in (-\infty, 1/2)$ or $E \in (-\frac{v^2}{2|B|}\sqrt{4mB-1}, \frac{v^2}{2|B|}\sqrt{4mB-1})$ for $mB \in (1/2, +\infty)$.

[1] J. E. Moore, Nature **464**, 194 (2010).
[2] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[3] L. Fu and C. L. Kane, Phys. Rev. B **74**, 195312 (2006).
[4] J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306(R) (2007).
[5] T. Fukui, T. Fujiwara, and Y. Hatsugai, J. Phys. Soc. Jpn. **77**, 123705 (2008).
[6] X. L. Qi, T. L. Hughes, and S. C. Zhang, Phys. Rev. B **78**, 195424 (2008).
[7] R. Roy, Phys. Rev. B **79**, 195321 (2009).
[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] W. C. Lee, C. Wu, D. P. Arovas, and S. C. Zhang, Phys. Rev. B **80**, 245439 (2009).
[11] X. Zhou, C. Fang, W. F. Tsai, and J. P. Hu, Phys. Rev. B **80**, 245317 (2009).
[12] H. M. Guo and M. Franz, Phys. Rev. B **81**, 041102(R) (2010).
[13] Q. H. Wang, D. Wang, and F. C. Zhang, Phys. Rev. B **81**, 035104 (2010).
[14] R. R. Biswas and A. V. Balatsky, Phys. Rev. B **81**, 233405 (2010).
[15] L. Yu, Acta Phys. Sin. **21**, 75 (1965).
[16] H. Shiba, Prog. Theor. Phys. **40**, 435 (1968).
[17] A. V. Balatsky, I. Vekhter, and J. X. Zhu, Rev. Mod. Phys. **78**, 373 (2006).
[18] G. E. Volovik, *The Universe in a Helium Droplet* (Clarendon Press, Oxford, 2003).
[19] J. C. Y. Teo and C. L. Kane, Phys. Rev. B **82**, 115120 (2010).
[20] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[21] B. A. Bernevig, T. L. Hughes, and S. C. Zhang, Science **314**, 1757 (2006).
[22] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X. L. Qi, and S. C. Zhang, Science **318**, 766 (2007).
[23] H. J. Zhang, C. X. Liu, X. L. Qi, X. Dai, Z. Fang, and S. C. Zhang, Nat. Phys. **5**, 438 (2009).
[24] H. Z. Lu, W. Y. Shan, W. Yao, Q. Niu, and S. Q. Shen, Phys. Rev. B **81**, 115407 (2010).
[25] W. Y. Shan, H. Z. Lu, and S. Q. Shen, New J. Phys. **12**, 043048 (2010).
[26] S. Q. Shen, W. Y. Shan, and Hai-Zhou Lu, SPIN (2011).
[27] G. E. Volovik, JETP Lett. **91**, 55 (2010).
[28] X. L. Qi, Y. S. Wu, and S. C. Zhang, Phys. Rev. B **74**, 045125 (2006).
[29] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).
[30] Y. Zhang, K. He, C. Z. Chang, C. L. Song, L. L. Wang, X. Chen, J. F. Jia, Z. Fang, X. Dai, W. Y. Shan, S. Q. Shen, Q. Niu, X. L. Qi, S.-C. Zhang, X. C. Ma, and Q. K. Xue, Nat. Phys. **6**, 584 (2010).
[31] B. Zhou, H. Z. Lu, R. L. Chu, S. Q. Shen, and Q. Niu, Phys. Rev. Lett. **101**, 246807 (2008).
[32] P. Michetti and P. Recher, Phys. Rev. B **83**, 125420 (2011).
[33] N. Read and D. Green, Phys. Rev. B **61**, 10267 (2000).
[34] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nat. Phys. **5**, 398 (2009).
[35] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).
[36] E. V. Thuneberg, J. Kurkijarvi, and D. Rainer, Physica B **107**, 43 (1981).
[37] M. Buttiker, Phys. Rev. B **38**, 9375 (1988).