Stability of $Z_2$ topological order in the presence of vacancy-induced impurity band

Shi-Ting Lee$^1$, Shin-Ming Huang$^1$ and Chung-Yu Mou$^{1,2,3}$

$^1$ Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan
$^2$ Institute of Physics, Academia Sinica, Nankang 11529, Taiwan
$^3$ Physics Division, National Center for Theoretical Sciences, PO Box 2-131, Hsinchu, Taiwan

E-mail: mou@phys.nthu.edu.tw

Received 26 January 2014, revised 8 April 2014
Accepted for publication 24 April 2014
Published 29 May 2014

Abstract

Although topological insulators (TIs) are known to be robust against non-magnetic perturbations and exhibit edge or surface states as their distinct feature, experimentally it is known that vacancies often occur in these materials and impose strong perturbations. Here, we investigate effects of vacancies on the stability of $Z_2$ topological order using the Kane–Mele (KM) model as a prototype of topological insulators. It is shown that even though a vacancy is not classified as a topological defect in the KM model, it generally induces a pair of degenerate mid-gap bound states only in the TI phase. Hence mid-gap bound states due to vacancies arise from the same $Z_2$ classification of topological insulators. Furthermore, we show that in the presence of many vacancies, an impurity band is induced and coexists with edge states until a phase transition occurs when the spectral weights of Dirac cones near Dirac points are depleted. Our analyses indicate that the same scenario holds for point vacancies or lines of vacancies in 3D TIs as well.

Keywords: vacancy, topological insulators, stability

(Some figures may appear in colour only in the online journal)

1. Introduction

Building low dimensional electronic systems has been a focus of intense interest since the discovery of low-dimensional materials, such as carbon nanotubes and graphene. Recent theoretical and experimental works on topological order of materials [1–3] have led to the discovery of a different route to construct low dimensional electronic systems through a new class of materials, called topological insulators (TIs). In these materials, surface states or edge states arise in the boundaries of bulk insulating materials and can host electrons in low dimensions. Depending on the dimension of the materials, a two-dimensional TI hosts one-dimensional gapless edge states while a three-dimensional TI hosts two-dimensional gapless Dirac Fermions. These low-dimensional surface states or edges arise from $Z_2$ topological order of bulk states [1] and are protected by symmetries of bulk states.

Theoretically, the fundamental reason for the emergence of surface states or edge states in TI is due to underlying $Z_2$ topological structure in the bulk state. Since topological order is protected by associated symmetries [4], surface states and edge states in TIs are considered to be robust against bulk disorders that do not break the time-reversal symmetry that is associated with $Z_2$ topological order. The robustness against disorders is the key to potential technological applications and hence great effort has been dedicated to understanding the behavior of the topological materials in the presence of disorders [5–10], [12]. Indeed, surface states or edge states are shown to persist in the presence of weak disorders [5]. However, it is also found that in the presence of disorders, anomalous transmutation between TI and insulators with trivial topology may happen [9, 10, 12]. In particular, in computer simulations of a HgTe quantum well, it is discovered that an ordinary insulating state can be transformed into a topological insulator, called topological Anderson insulator [10, 11]. The mechanism behind the transformation is shown to be related to the renormalization of $Z_2$ topological index by disorders near the Dirac points [12] and the transformation may also depend on type of disorders [13]. Hence while TI is robust against perturbations, the exact phase boundary for TI may subject to change in the presence of disorders.
Experimentally, it is found that strong perturbations such as lattice defects often occur in topological insulators [14]. From the experience in other similar materials characterized by Dirac Hamiltonian, the presence of lattice defects often induces peculiar properties. In the case of graphene, it was found that vacancies can give rise to magnetic moments and turn graphene into a ferromagnet [15]. For topological insulators, it is shown that lattice defects with non-trivial topology also induce bound states [16]. However, in the tetradymite semiconductor Bi$_2$Se$_3$ that has been most extensively investigated, the most common observed lattice defects are selenium vacancies, which are not topological defects. These vacancies are believed to give rise to electron doping and increase the conductivity of bulk states dramatically [14]. Furthermore, there is also evidence that vacancies may induce an impurity band that affects current transportation [17]. Since vacancies do not break time-reversal symmetry, their presence is consistent with symmetries that are associated with $Z_2$ topological order. However, instead of being weak, vacancies are considered as strong perturbations and may change the topological structure of TI. It is therefore interesting and crucial to examine robustness of TI in the presence of vacancies.

In this paper, we will examine the stability of $Z_2$ topological order in the presence of vacancies using the Kane–Mele model. It is shown that even though vacancies are not classified as topological defects in the Kane–Mele model [18], the original $Z_2$ topological order insures that midgap bound states are induced only in topologically nontrivial phase. Based on Green's function analysis of edge states [19], we show that these midgap bound states are remnant edge states and form an impurity band in the presence of many vacancies. We shall show that the impurity band coexists with edge states when the $Z_2$ topological order persists until the spectral weights of Dirac cones near Dirac points are depleted. Finally, we briefly extend our analyses to three dimensions and show that there must be midgap bound states associated with point vacancies or line of vacancies in 3D TIs as well.

2. Theoretical formulation and midgap bound states near a vacancy

We start with the Kane–Mele model on a honeycomb lattice. The Hamiltonian is given by [1]

$$H = -t \sum_{\langle i,j \rangle} c_i^\dagger c_j + i \frac{2 \Delta_0}{3\sqrt{3}} \sum_{\langle i,j \rangle} \nu_0 \sigma^i c_i^\dagger c_j + \frac{2k_F}{3} \sum_{\langle i,j \rangle} c_i^\dagger \left( \sigma \times \hat{d}_i \right) c_j + \lambda_0 \sum_i \mu_i c_i^\dagger c_i + \sum_i V_i^c c_i^\dagger c_i,$$

where $c_i^\dagger = (c_i^\dagger_0, c_i^\dagger_1)$ creates an electron at lattice site $i$. The first term accounts for nearest-neighbor hoppings. The second term is a spin–orbit interaction between next nearest neighbors, in which $\nu_0 = (2/\sqrt{3}) (\hat{d}_i \times \hat{d}_j) \xi$ with $\hat{d}_1$ and $\hat{d}_2$ being two nearest-neighbor bonds that connect site $j$ to site $i$ and $\sigma$ are the Pauli matrices. The third term is the Rashba coupling. The fourth term characterizes the sublattice site energies with $\mu_i = +1 (-1)$ for $i \in A(B)$ sites. This term will be considered only in this section. The last term is the random disorder potential that models the impurities, where $V_i$ takes a nonvanishing value of $V$ only at impurity sites. In the usual treatment of disorder potential, it often assumes that $V_i$ is smooth and is characterized by its correlation function $\langle V_i V_j \rangle$. Here, however, by taking $V_i \rightarrow \infty$, $i$ site is excluded for electrons to visit and hence the limit $V_i \rightarrow \infty$ simulates a vacancy at $i$ site. As we shall see, the new ingredient of the lattice vacancy is the possibility of inducing bound states near a vacancy site, which can not be obtained perturbatively by using $\langle V_i V_j \rangle$. Note that vacancies can be also modeled by cutting all relevant couplings to sites of vacancies and it can be shown that this approach yields the same results in low energy sectors where energy bands occupy. In the following, all energies and lengths will be in units of $t$ and $a$ (lattice constant) respectively.

We shall first consider the case of a single vacancy located at the origin $\vec{r} = (0,0) \equiv \mathbf{0}$ that is located at $A$ site. The

![Figure 1](image-url)
configuration is shown in figure 1(a). According to the classification made [18], far from a vacancy in 2D, there is no non-trivial winding number associated with the hole introduced by a vacancy and hence a vacancy is not a topological defect. However, this does not imply that there is no midgap bound state associated with a vacancy. In fact, in the continuum limit for HgTe where the Dirac point is at Γ point, by approximating a vacancy as a circular hole with radius $R$, it is found that midgap bound states survive with energies being half of the gap magnitude in the limit $R \to 0$ [20]. In the continuum limit, electrons that occupy the midgap bound states can be viewed as Dirac fermions in a particular curved space [21]. Hence, a midgap bound state in a single vacancy results from a superposition of edge states that are curved into the circular surrounding of the vacancy. However, for a vacancy in a honeycomb lattice, the situation is quite different. As shown in figure 1(a), a vacancy has a finite size and its surrounding is not in a circular shape. In addition, Dirac points are located at finite wave vectors for honeycomb lattices. It is necessary to examine whether there are midgap bound states associated with vacancies.

To clarify the issue of whether midgap bound states will be induced by a single vacancy, we resort to method of the Green’s function, which has the advantage of not being restricted by finite size effects. In the presence of impurities or vacancies, the Green’s function $G_{\alpha\beta}(\vec{r}, \vec{r}', E)$ that describes the amplitude for the electron of energy $E$ to propagate from $\vec{r}'$ with spin component $\beta$ to the position $\vec{r}$ with spin component $\alpha$ satisfies

$$ (E-H)G(\vec{r}, \vec{r}', E) = \delta_{\vec{r}, \vec{r}'} $$

where we have collectively represented $G_{\alpha\beta}$ by the $2 \times 2$ matrix $G$. If we denote the Green’s function in the absence of vacancies by $g(\vec{r}, \vec{r}', E)$, we find that for a single impurity at $\vec{r} = 0$, $G$ satisfies

$$ G(\vec{r}, \vec{r}', E) = g(\vec{r}, \vec{r}', E) + g(\vec{r}, 0, E)VG(0, \vec{r}', E) $$

$$ = g(\vec{r}, \vec{r}', E) + g(\vec{r}, 0, E)Tv_{\zeta}G(0, \vec{r}', E), $$

where $T$ is the T-matrix for a single vacancy at $\vec{r} = 0$ and can be written as

$$ T = V + VgV + VgVgV + \ldots = V(1-gV)^{-1}. $$

If $H$ supports midgap bound states, $G$ must contain energies of the midgap bound states as poles in $E$. Therefore, the energies of midgap bound states, $E_0$, are determined by [15]

$$ \det(1-gV)_{E=E_0} = 0. $$

In the limit of $V \to \infty$, the impurity becomes a vacancy. Equation (5) reduces to equations that determine $E_0$ for midgap bound states of vacancies

$$ \det g(0, 0, E_0) = g_{11}(0, 0, E_0)g_{22}(0, 0, E_0) $$

$$ -g_{12}(0, 0, E_0)g_{21}(0, 0, E_0) = 0, $$

where 1 and 2 represent spin up and down respectively. We shall show in below that the existence of a midgap bound state for a single vacancy results from the existence of edge states.

To connect edge states with midgap bound states of a single vacancy, we shall start from a system that is a Z3 topological insulator. Hence there are helical edges states for any edges. Furthermore, the energy spectrum of edge states is a Dirac-like spectrum [1]. For a given edge along $\zeta$-direction, since it is translationally invariant along $\zeta$-direction, the system can be dimensionally reduced to one dimension by a partial Fourier transformation along $\zeta$-direction. Let the wave vector along $\zeta$ be $k_{\zeta}$. The existence of edge states thus implies that there are states with energies $E$ being inside the bulk energy gap and in the form [1]

$$ E = \mu_{\zeta} \pm v_{\zeta}k_{\zeta}, $$

where $v_{\zeta}$ is the speed of helical edge states and $\mu_{\zeta}$ is the intersecting energy of two helical modes (energy of the Dirac point). Note that the effective Hamiltonian that gives rise to the helical spectrum of equation (7) is a Dirac Hamiltonian and can be written as [1]

$$ H_{\text{eff}} = \begin{pmatrix} \mu_{\zeta} & v_{\zeta}k_{\zeta} \\ v_{\zeta}k_{\zeta} & -\mu_{\zeta} \end{pmatrix}. $$

where $H_{\text{eff}}$ is a $2 \times 2$ matrix in the spin space. In the following, we shall show that a midgap bound state for a single vacancy is essentially a weighted superposition of edge states of all possible orientations (denoted by $\zeta$) characterized by equation (7).

We now consider creating edges by introducing a line of point vacancies in an infinite honeycomb lattice as illustrated in figure 1(b). The vacancy line cuts the infinite honeycomb lattice into two semi-infinite honeycomb lattices with two edges. As we shall show in below, after partial Fourier transformation on coordinates along the vacancy line, for each Fourier mode $k_{\zeta}$, the vacancy line becomes a point. That is, through dimensional reduction, the vacancy line and edge states that are associated with two edges reduce to an effective vacancy with midgap bound states in one dimension. The connection of midgap bound states for a single vacancy to edge states is thus established. From this point of view, we may assume that the vacancy line passes through $\vec{r} = 0$ and its direction is denoted as $\zeta$-direction. Hence the resulting edge for each semi-infinite honeycomb lattice is along $\zeta$-direction. For later usage, we denote the direction perpendicular to $\zeta$ by $\eta$ with the corresponding coordinate being denoted by $x_{\eta}$. In the example shown in figure 1(b), one has $\zeta = x$, $\eta = y$ and $k_{\zeta} = k_x$. After the partial Fourier transformation along $\zeta$-direction, the Hamiltonian of the whole system (two semi-infinite honeycomb lattices plus a vacancy line) is reduced to one-dimensional subsystems of different $k_{\zeta}$, hence the total Hamiltonian $H$ can be written as

$$ H = \sum_{k_{\zeta}} h(k_{\zeta}). $$

Note that after the partial Fourier transformation, the vacancy line reduces to an effective vacancy for each $k_{\zeta}$ and each $h(k_{\zeta})$ contains a vacancy at the origin of A site, $x_{\eta} = 0$. The associated Green’s function is a function of $k_{\zeta}$ and $x_{\eta}$ and can be written as $G_{\alpha\beta}(x_{\eta}, x_{\eta}, k_{\zeta}, E)$. Clearly, following the derivation
that leads to equations (3) and (4), since there is a vacancy at $x_\eta = 0$, we find that for a given $k_\zeta$, the T-matrix is given by

$$T(k_\zeta, E) = V(1 - g(x_\eta = 0, x_\tau = 0, k_\zeta, E)^{-1}.$$  

In the limit of $V \to \infty$, one gets

$$T = -\left[g(x_\eta = 0, x_\tau = 0, k_\zeta, E)\right]^{-1}. \tag{11}$$

Since edge states are created by the vacancy line, according to equation (7), $\mu_\zeta \pm v_\zeta k_\zeta$ must be eigen-energies of $h(k_\zeta)$ and hence these energies must appear as poles of the T-matrix. In addition, using the fact that edge states are described by the effective Hamiltonian given in equation (8) and their energies must be poles of $T$, it implies that

$$T \sim (E - H_{\text{eff}})^{-1}. \tag{12}$$

As a result, by combining equations (11) and (12), we conclude that $g$ will take the following form

$$g(x_\eta = 0, x_\tau = 0, k_\zeta, E) = w(k_\zeta, E) \left( \frac{E - \mu_\zeta}{v_\zeta k_\zeta} \right)^{1/2} \left( \frac{E - \mu_\zeta}{v_\zeta k_\zeta} \right)^{-1/2}, \tag{13}$$

where the factor $w(k_\zeta, E)$ is a proportional constant and accounts for the weight associated with $k_\zeta$ mode. Note that $w(k_\zeta, E)$ is smooth, even in $k_\zeta$ and it has no zeros in $E$.

Using equation (13), the origin of the vacancy state is made clear. We first note that the bulk Green’s function in equation (6) satisfies

$$g(0, 0, E) = \sum_{k_\zeta} g(x_\eta = 0, x_\tau = 0, k_\zeta, E). \tag{14}$$

Combining equations (13) and (14), we find

$$g(0, 0, E) = \left( \begin{array}{cc} w E - \mu_\zeta & 0 \\ 0 & w E - \mu_\zeta \end{array} \right). \tag{15}$$

where $w_i = \sum k_\zeta w_i$, $\mu_i = \sum k_\zeta w_i$, and we have made use the fact that the summation $\sum w k_\zeta$ vanishes due to that $w(k_\zeta, E)$ is an even function of $k_\zeta$. Equation (15) then yields two midgap bound states with a degenerate energy at

$$E_0 = \mu_i = \sum k_\zeta w \mu_i \sum k_\zeta w. \tag{16}$$

Here $\mu_i$ is the weighted average of the intersecting energy for two helical modes and hence it must lie inside the gap

$$|\mu_i| < \Delta_\zeta. \tag{17}$$
where $\Delta_c$ is the gap along $\zeta$-direction. In the TI phase, since there are edge states inside the energy gap for all orientations of edges, $E$ must satisfy equation (17) for all possible $\zeta$, including the minimum of $\Delta_c$, which is the energy of the system. Hence one concludes that in consistent with the Kramers degeneracy theorem, there must be a pair of degenerate vacancy states inside the energy gap in the TI phase.

To verify the above conclusion, we evaluate energies of midgap bound states induced by a single vacancy through solving equation (6) numerically. Figure 2 shows energies of midgap bound states induced by a single vacancy. It is seen that generally the midgap bound states degenerate at one midgap energy. When there are $M$ vacancies at random positions, $M$ midgap energies are induced. Figure 3(b) shows that as the number of vacancies increases, these midgap bound states start to form an impurity band inside the gap. The appearance of an impurity band is in agreement with the observation made in [17]. These midgap states are localized states. Note that, since the system is composed by spin-1/2 electrons with time-reversal symmetry and without spin-rotation symmetry, disordered topological insulators belong to a symplectic symmetry class [22]. The level statistics of midgap energies is thus characterized by the so-called Gaussian symplectic ensemble [23]. In particular, for a Gaussian symplectic ensemble, the level-spacing distribution follows Wigner-Dyson type distribution [23]. Hence, for midgap energies in the impurity band, the distribution of level spacings also follows the Wigner–Dyson distribution.

From figure 3(a), we have the following counting for level distribution. In a honeycomb lattice with the number of unit cells being $N$, there are $2N$ lattice points and the total number of states is $4N$. If there are $M$ vacancies, the distribution of energies is as follows

\[
\begin{align*}
\text{Level counting} & \quad \text{valence band: } 2N-M \\
& \quad \text{impurity band: } M \\
& \quad \text{conduction band: } 2N-M \\
& \quad \text{high energy states: } M.
\end{align*}
\]

By using the above level distribution, one can characterize the $Z_2$ topological index for each band. For this purpose, we first note that the defining property of the non-trivial $Z_2$ topological index is the quantum spin Hall effect (QSHE). Therefore, the $Z_2$ topological index can be evaluated by computing quantum spin Hall conductance (QSHC). Computing the $Z_2$ index by evaluating QSHC is particularly useful when there is no translational invariance in the presence of vacancies and one can not define electronic Bloch states. To calculate the QSHC, generally one imposes spin-dependent twisted boundary conditions [24] as

\[
\begin{align*}
c^\alpha_{j+N_y} &= e^{i\alpha\pi} c^\alpha_j, \quad c^\beta_{j+N_y} = e^{i\beta\pi} c^\beta_j. \quad (19)
\end{align*}
\]

Here $\alpha, \beta = \pm 1$ are the spin indices that represent the spin state of $\uparrow$ and $\downarrow$ respectively, $j$ specifies the site, $N_y$ and $N_x$ are numbers of sites along $x$ and $y$ directions respectively (refer to figure 1(a)) and $\theta^x$ ($\theta^y$) are phases acquired whenever an electron with spin component $\alpha (\beta)$ goes across the boundary along $x (y)$ directions respectively. Both $\theta^x$ and $\theta^y$ are in the range $[0, 2\pi]$. Following [24], when the degeneracy between spin state $\uparrow$ and $\downarrow$ is lifted, the Hall conductances are generally represented as $\sigma^xy$, which characterizes the Hall voltage in the $x$ direction due to electrons of spins component $\alpha$ when
the spin-polarized current with spin component $\beta$ flows in the $y$ direction. The topological Chern number that is associated with $\sigma_{xy}^\alpha$ is thus defined as [24]

$$C_{\alpha\beta} = \frac{i}{4\pi} \int d\theta^x d\theta^y \left[ \left( \frac{\partial \psi}{\partial \theta^x} \right)^* \frac{\partial \psi}{\partial \theta^y} - \left( \frac{\partial \psi}{\partial \theta^y} \right)^* \frac{\partial \psi}{\partial \theta^x} \right],$$  \hspace{1cm} (20)

where $\psi$ is the many-particle ground state wavefunction of the system. $C_{\alpha\beta}$ can be collectively represented as a 2 × 2 matrix, representing the non-Abelian nature of the Hall conductance. For the quantum Hall effect (QHE), it is entirely due to charges. For the quantum Hall effect (QHE), it is entirely due to charges. The Hall conductance of the QHE is given by $\sigma_\| = \sum_{\alpha,\beta} \sigma_{xy}^\alpha \delta_{\alpha\beta}$. Hence it is clear that $\Sigma_{\alpha,\beta} C_{\alpha\beta}$ is the corresponding Chern number associated with the QHE. On the other hand, the QSHE is defined as the difference of Hall conductances between spin $\uparrow$ and $\downarrow$.

$$\sigma_\| = \sum_{\alpha} \sigma_{xy}^\alpha = \sum_{\alpha} \sum_{\beta} \sigma_{xy}^\alpha \delta_{\alpha\beta},$$  \hspace{1cm} (21)

As a result, the corresponding spin Chern number that is associated with the QSHE is given by

$$C_s = \sum_{\alpha} \sigma_{xy}^\alpha.$$  \hspace{1cm} (22)

From equations (20) and (22), it is seen that the minus sign associated with down spin can be absorbed into the twisted phase $\theta_i^\|$, so that one can set $\theta_i^\| = -\theta_i$. Therefore, to focus on the spin Chern number, one sets $\theta_i^\| = \theta_i$, and $\theta_i^\| = -\theta_i$, and imposes the spin-dependent twisted boundary conditions [24, 26] as follows

$$C_{\alpha\beta} = e^{i\theta_\|} c_{\alpha+j\beta}, \quad C_{\alpha+j\beta} = e^{-i\theta_{\bar{\alpha}}} c_j,$$

where $\sigma_{\bar{\alpha}}$ is the $z$ component of the Pauli matrices and both $\theta_i$ and $\theta_i^\| \in [0, 2\pi]$. The gauge potential $\Phi$ imposed by twisted boundary conditions is

$$\Phi(i_x, i_y) = \frac{\theta_i}{N_x} + \frac{\theta_i^\|}{N_y},$$  \hspace{1cm} (24)

Due to the presence of $\sigma_{\bar{\alpha}}, \Phi$ does not commute with $H$, which reflects the non-Abelian nature of the problem. Note that the spin Chern number computed by the spin-dependent twisted boundary conditions has been rigorously shown to be equivalent to the $Z_2$ topological index [26, 27].

Following [25] and [26], using the energy eigenkets of $H$ with spin-dependent twisted boundary conditions, one can compute the spin Chern number which yields the same classification as that of $Z_2$ topological order. For this purpose, we discretize the space $[0, 2\pi] \times [0, 2\pi]$ into $M \times M$ mesh points so that a general twisted boundary condition is represented by $\theta \equiv (\theta_p, \theta_q)$ with $\theta_p = 2\pi p/M, \theta_q = 2\pi q/M$ and $p, q = 0, 1, 2, 3, \ldots, M-1$. For each $\theta$, since hopping across boundaries depends on $\theta$, the Hamiltonian depends on $\theta, H(\theta)$, with eigenstates being $\ln(\theta)$. For each bond that connects $\theta$ and $\theta + \delta \theta^{\uparrow}$ or $\theta + \delta \theta^{\downarrow}$ with $\delta \theta = 2\pi/M$ and $\| = \times$ or $\gamma$, a non-Abelian link variable $U_{ij}(\theta)$ is defined as the overlap matrix

$$U_{ij}(\theta) = \langle m(\theta) | \ln(\theta + \delta \theta^{\uparrow}) \rangle,$$

where $n$ and $m$ are indices of energy bands that are included for computing the spin Chern number. From the link variable $U_{ij}(\theta)$, a U(1) link $u_{ij}(\theta)$ can be formed

$$u_{ij}(\theta) = \text{det} U_{ij}(\theta) / \text{det} U_{ij}(1).$$  \hspace{1cm} (26)

One can then find the lattice field strength $F_{xy}$ for each plaquette in the $(\theta_i, \theta_j)$ lattice by computing

$$F_{xy}(\theta) = \ln \left[ u_{ij}(\theta) u_{ij}(\theta + \delta \theta^{\uparrow}) \right] \times u_{ij}^{-1}(\theta + \delta \theta^{\uparrow}) u_{ij}^{-1}(\theta + \delta \theta^{\downarrow}) u_{ij}^{-1}(\theta + \delta \theta^{\downarrow}),$$  \hspace{1cm} (27)

where the principle branch of logarithm is taken. The spin Chern number is the summation of $F_{xy}$ over all plaquettes given by

$$C_s = \frac{1}{2\pi N} \sum_{\theta} F_{xy}(\theta).$$  \hspace{1cm} (28)

The spin Chern number thus obtained is always an integer. Furthermore, it has the advantage of being accurate even when the computation is done with small sizes of honeycomb lattice [25, 26].
To characterize TI with vacancies at random positions, we compute $C_i$ for valence band, impurity band and conduction band. For a given density of vacancies, it is found that depending on the configuration of positions for vacancies, the spin Chern number of each band can be either 2 or 0, depending on the configuration of positions for vacancies, the spin Chern number ($C_i$) of the impurity band always vanishes. Hence $C_i$ of the conduction band is always opposite to that of the valence band. To confirm the validity of computed $C_i$ based on equation (28), for a given vacancy configuration, we examine the bulk-edge correspondence by computing the energy spectrum for honeycomb ribbons with the armchair edges. To this end, it is convenient to compute the spectral function $A(k, \omega)$ defined by

$$A(k, \omega) = \frac{1}{Nz} \text{Im} \left\{ \sum_{j=1}^{4} \sum_{n,m=1}^{4N} \frac{|n|\psi(k,j,m)|^2}{(\omega - E_n + i\Gamma)} \right\},$$

where $\Gamma$ is an energy scale characterizing energy resolution, $ln$ is the $n$-th energy eigenstate of the impure system with eigen-energy $E_n$ and $\psi(k,j,m)$ is the energy eigenstate in the absence of vacancy with $k$ being the wave vector after a Fourier transformation. Figure 4 shows the spectral function $A$ of a honeycomb ribbon with armchair edges in the presence of vacancies for different vacancy densities. Here the listed spin Chern numbers are computed by using equation (28) based on twisted boundary conditions. The armchair edges are along the $x$ direction so we have the conserved wave vector $k_x$. It is seen that for small vacancy densities when $C_i$ does not vanish for both valence and conduction bands, edges states exist and coexist with the impurity band at the center. Furthermore, consistent with the bulk-edge correspondence, edge states reside between the impurity band and the valence band or the conduction band. As the density of vacancies increases, the strength of the impurity band increases and eventually $C_i$ for all energy bands vanish. In this case, as shown in figure 4(c), edge states diminish as well.

As indicated in the above, to explore the nature of transition, it is necessary to compute the averaged spin Chern number. In figure 5, we show the averaged spin Chern number for the valence band of TIs in the presence vacancies which positions are only at $A$ sites or distribute equally at $A$ sites and $B$ sites. The validity of computations based on equation (28) is also checked by direct computation of the corresponding systems with open boundaries. It is seen that the phase transition from TI to a trivial insulator occurs as density of vacancies increases. However, depending on positions of vacancies, the transition can be either first order when vacancies are randomly chosen at $A$ sites or continuous when vacancies are randomly chosen equally at $A$ sites or $B$ sites. As we shall explore in the following, the mechanism that causes different transition behaviors lies in the speed of the Dirac cones being depleted. Since the formation of Dirac cones involves with both $A$ sites and $B$ sites, removing points in $A$ sites and $B$ sites simultaneously will start to deplete the Dirac cone and the depletion is proportional to the number of points being removed. As a result, the transition is continuous when vacancies are distributed equally at $A$ sites and $B$ sites. On the other hand, removing points solely at $A$ sites does not deplete the Dirac cone immediately at first until when too many $A$ points are removed, the bi-partite nature is destroyed. At that point, further increase of vacancies induces a discontinuous transition to a trivial insulator with $C_i = 0$.

To explore the mechanism of how $Z_2$ topological order is destroyed by vacancies, we first note that near the phase boundary when the energy gap vanishes, the dominant contribution of the spin Chern number is from $k$ points near the Dirac points due to the singularity induced by band touching of conduction and valence bands. It is known that the Dirac cones continue to dominate the contribution of $Z_2$ topological order even when the system is away from the phase boundary [28]. Therefore, it is essential to examine how Dirac cones evolve with the introduction of vacancies. For this purpose, we examine the spectral function $A(k, \omega)$ near Dirac cones using the periodic boundary conditions so that the spectral weight only reflects effects due to vacancies not due to edge states. In figure 6(a), we show side view of $A(k, \omega)$ of a honeycomb lattice in the TI phase with a size of $100 \times 8$. A clear peak located right at the Dirac point is exhibited. The peak that reflects the spectral weight near Dirac cones is depleted quickly once vacancies are introduced, as shown in figure 6(b). Here the number of vacancies is 90, i.e., density of vacancies is 6.25%. Clearly, it shows that the diminishing of $C_i$ correlates with the depletion of Dirac cones. To further differentiate effects due to distribution of vacancies on different sub-lattices, we compute the density of states for 20 vacancies in a $50 \times 50$ honeycomb lattice with vacancies being solely distributed at $A$ sites, denoted by DOS(A) and vacancies be distributed equally at $A$ sites and $B$ sites, denoted as DOS(AB). The difference of densities of states is shown in figure 6(c). Clearly, the difference, DOS(A)-DOS(AB), is pronounced at several energies. Here dashed
and dotted lines indicate energies of gaps at Dirac points (K point), while dot-dashed lines indicate energies at M points. It is seen that removing points solely at A sites tends to keep a higher spectral weight at symmetry points (K and M points) in k space. As a result, the Dirac cones are less depleted when vacancies are distributed solely at A sites. The analyses shown here show that direct depletion of the Dirac cones is the main mechanism for transition from the TIs to trivial insulators for low concentrations of vacancies. For higher concentration of vacancies, as shown in figure 4, the impurity band starts to overlap with energy bands and destructs of $Z_2$ topological order eventually.

4. Discussion and conclusion

In summary, we have shown that vacancies in the TI phase of the Kane–Mele model always induce midgap bound states. These midgap bound states result from curving edge states into the surroundings of vacancies. The same reasoning can be generalized to investigate point vacancies or line vacancies in 3D TIs. In this case, the T-matrix still obeys equation (4). By introducing a plane of vacancies passing through $\mathbf{T} = 0$ that cuts infinite 3D lattice into two semi-infinite 3D lattices and following similar considerations that lead to equation (14), one arrives at

$$g(0, 0, E) = \sum_{k_a k_\beta} g(x_\gamma = 0, x_\gamma = 0, k_a, k_\beta, E),$$  

where $k_a$ and $k_\beta$ are wave vectors in two orthogonal directions of the plane with vacancies and $x_\gamma$ is the coordinate of the axis in perpendicular to the plane. For 3D TIs, since there are midgap surface states characterized by spectrum $\pm v_\parallel k_\parallel + \mu_\parallel$ with $k_\parallel = \sqrt{k_x^2 + k_y^2}$, the same reasoning leads to the conclusion that in the low energy sector, $g$ must take the following form

$$g(0_r, 0_r, k_a, k_\beta, E) = w(k_a, k_\beta, E) \begin{vmatrix} E - \mu_\parallel & v_\parallel (k_a - ik_\gamma) \\ v_\parallel (k_a + ik_\gamma) & E - \mu_\parallel \end{vmatrix},$$  

where $v_\parallel$ is the speed of electrons in surface states, $\mu_\parallel$ is the energy of the Dirac point and $w(k_a, k_\beta, E)$ is the weight associated with $k_\parallel$ and is a smooth function of $E$. Hence, we arrive at a conclusion similar to equation (17) with $\Delta_2$ being replaced by the gap $\Delta_\parallel$ of $k$ that lies in directions perpendicular to $x_\gamma$. As a result, one also concludes that there must be a pair of degenerated states inside the energy gap for a single vacancy in 3D topological insulators. This conclusion is in agreement

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Depletion of Dirac cones exhibited in $A(k, \omega)$ and the density of states. Here (a) and (b) show the side view of $A(k, \omega)$ along $\omega$ axis for a honeycomb lattice of size $100 \times 8$ with different number of vacancies: (a) 0% (b) 6.26%. A sharp peak at the Dirac point is clearly revealed in (a) but is depleted quickly when vacancies are introduced, as shown in (b). (c) Difference of densities of states for 20 vacancies in a $50 \times 50$ honeycomb lattice with vacancies being solely distributed at A sites, DOS(A) and vacancies being distributed equally at A sites and B sites, DOS(AB). Here $\lambda_\parallel = 0.075$ and $\lambda_\omega = 0.18\sqrt{3}$. Dashed and dotted lines indicate energy of gaps at Dirac points, while dot-dashed lines indicate energies at M points. It is seen that the Dirac cones are less depleted when vacancies are solely distributed at A sites.}
\end{figure}
with direct numerical simulations investigated in [29]. Similarly, for a single line of vacancies along \( \delta \) direction, performing a partial Fourier transform on the Hamiltonian along \( \delta \) reduces the problem to a 2D problem with a point vacancy. Hence there must be a pair of degenerated states inside the energy gap for every \( k_{B} \) mode along each line of vacancies in 3D TIs.

In addition to showing that midgap bound states must arise in the presence of a single vacancy, we also show that an impurity band emerges when the number of vacancies increases. The impurity band mixes directly with edge states. However, the impurity band always has trivial topological structure and the \( \mathbb{Z}_{2} \) topological order persists in both the conduction band and the valence band. Hence edge states exist between the impurity band and perturbed conduction or valence band. The mechanism behind the transition from TI to a trivial insulator due to vacancies is shown to result from the depletion of Dirac cones. Furthermore, due to different speeds of depleting Dirac cones, the transition can be either first order when vacancies are randomly chosen at \( A \) sites or continuous when vacancies are randomly chosen equally at \( A \) sites or \( B \) sites.

While so far in this work we have only considered the effects of vacancies, our results are not restricted to the limit \( V \rightarrow \infty \). For finite and large \( V \), equation (6) is modified with a correction of \( O(1/V) \). Hence midgap bound states induced by vacancies are shifted by the order of \( O(1/V) \), in agreement with results shown in figure 3. The phase boundary for midgap bound states shown in figure 2 then remains the same. As a result, the impurity band is also shifted by a similar amount. The majority of our results remains valid and provides useful characterization of electronic states induced by point defects in TIs.

Acknowledgments

We thank Professor M-C Chang for useful discussions. This work was supported by the National Science Council of Taiwan.

References

[1] Kane C L, Mele E J 2005 Phys. Rev. Lett. 95 146802
[2] Hasan M Z and Kane C L 2010 Rev. Mod. Phys. 82 3045
[3] Qi S X L and Zhang S C 2011 Rev. Mod. Phys. 83 1057
[4] Schnyder A P, Ryu S, Furusaki A and Ludwig A W W 2008 Phys. Rev. B 78 195125
[5] Sheng D N, Weng Z Y, Sheng L and Haldane F D M 2006 Phys. Rev. Lett. 97 036808
[6] Onoda M, Avishai Y and Nagaosa N 2007 Phys. Rev. Lett. 98 076802
[7] Essin A M and Moore J E 2007 Phys. Rev. B 76 165307
[8] Obuse H, Furusaki A, Ryu S and Mudry C 2007 Phys. Rev. B 76 075301
[9] Obuse H, Furusaki A, Ryu S and Mudry C 2008 Phys. Rev. B 78 115301
[10] Schubert G, Fehske H, Fritz L and Voita M 2012 Phys. Rev. B 85 201105
[11] Li J, Chu R L, Jain J K and Shen S Q 2009 Phys. Rev. Lett. 102 136806
[12] Prodan E 2011 Phys. Rev. B 83 195119
[13] Groth C W, Wimmer M, Akhmerov A R, Tworzydlo J and Beenakker C W J 2009 Phys. Rev. Lett. 103 196805
[14] Song J, Liu H, Jiang H, Sun Q F and Xie X C 2012 Phys. Rev. B 85 195125
[15] Hor Y S 2009 Phys. Rev. B 79 195208
[16] Cheng P et al 2010 Phys. Rev. Lett. 105 076801
[17] Huang B L, Mou C Y 2009 Europhys. Lett. 88 68005
[18] Huang B L, Chang M C and Mou C Y 2010 Phys. Rev. B 82 155462
[19] Rüegg A, Lin C 2013 Phys. Rev. Lett. 110 046401
[20] Ran Y 2014 arXiv: 1006.5454
[21] Ran Y, Zhang Y and Vishwanath A 2009 Nature Phys. 5 298
[22] Ren Z, Taskin A A, Sasaki S, Segawa K and Ando Y 2010 Phys. Rev. B 82 241306
[23] Teo J C Y, Fu L and Kane C L 2008 Phys. Rev. B 78 045426
[24] Wu S T, Mou C Y 2002 Phys. Rev. B 66 012512
[25] Mou C Y, Wortis R, Dorsey A T, Huse D A 1995 Phys. Rev. B 51 6575
[26] Wu S T and Mou C Y 2003 Phys. Rev. B 67 024503
[27] Shan W Y, Lu J, Lu H Z and Shen S Q 2011 Phys. Rev. B 84 035307
[28] Lee D H 2009 Phys. Rev. Lett. 103 196804
[29] Ryu S, Mudry C, Obuse H and Furusaki A 2007 Phys. Rev. Lett. 99 116601
[30] Mehta M L 2004 Random Matrices (San Diego: Elsevier)
[31] Sheng D N, Weng Z Y, Sheng L, Haldane F D M 2006 Phys. Rev. Lett. 97 036808
[32] Yang Y, Xu Z, Sheng L, Wang B, Xing D Y and Sheng D N 2011 Phys. Rev. Lett. 107 066602
[33] Fukui T, Hatsugai Y and Suzuki H 2005 J. Phys. Soc. Japan 74 1674
[34] Fukui T and Hatsugai Y 2007 Phys. Rev. B 75 121403
[35] Soluyanov A A and Vanderbilt D 2012 Phys. Rev. B 85 115415
[36] Ezawa M 2012 Phys. Rev. Lett. 109 055502
[37] Black A M, Balatsky A V 2012 Phys. Rev. B 85 121103
[38] Black-Schaffer A M and Balatsky A V 2012 Phys. Rev. B 86 115433

Shi-Ting Lee et al