PHYSICAL REVIEW B **86**, 081303(R) (2012)

Surface states of topological insulators

Fan Zhang,* C. L. Kane, and E. J. Mele

Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

(Received 28 March 2012; published 10 August 2012)

We introduce a topological boundary condition to study the surface states of topological insulators within a long-wavelength four-band model. We find that the Dirac point energy, the band curvature, and the spin texture of surface states are crystal-face dependent. For an arbitrary termination of a bulk crystal, the energy of the symmetry protected Dirac point is determined by the bulk physics that breaks particle-hole symmetry in the surface normal direction and is tunable by surface potentials that preserve time reversal symmetry. For a model appropriate to Bi_2Se_3 the constant energy contours are generically elliptical with spin textures that are helical on the cleavage surface, collapsed to one dimension on any side face, and tilted out of plane otherwise. Our findings identify a route to engineering the Dirac point physics on the surfaces of real materials.

DOI: 10.1103/PhysRevB.86.081303 PACS number(s): 73.20.−r, 71.70.Ej, 73.22.Dj

Introduction. The discovery[1–6] of topological insulators (TIs) and the synthesis[7–9] of three-dimensional materials that realize their physics has opened up a new field in solid state physics. Particular interest has focused on the TI surface states with point degeneracies that are topologically protected by time reversal symmetry ($\mathcal{T}$) when a trivial insulator with topological index[1,2] $\mathcal{Z}_2 = 1$ is joined to a TI with $\mathcal{Z}_2 = -1$. On the cleavage surface of Bi_2Se_3, a single Dirac cone has been identified[9–11] by angle resolved photoemission (ARPES) and scanning tunneling microscope (STM) experiments. The surface states form a spin-momentum-locked helical metal with conduction and valence bands exhibiting opposite helicities near the Dirac point, and develop hexagonal warping[9,11–13] of its constant energy surfaces away from the Dirac point.

Interest in the topological surface bands has led to the derivation of various models to describe them using bulk continuum Hamiltonians terminated at the boundary to vacuum. This approach is attractive since it provides a compact and often analytic theory of the TI surface states. In practice this approach has been problematic because its solution requires, in addition, the specification of the surface boundary condition (SBC) that terminates the bulk-derived model. The nature of the SBC has remained unclear and different approaches to it are discussed in the current literature. The surface has been introduced either by an *ad hoc* fixed node boundary condition[7,13–17] which clamps the wave function to zero at the boundary or by a "natural" boundary condition[18] which posits an energy functional whose minimization yields the SBC. The choice of SBC has profound consequences for the spatial form of the surface state wave functions and their interactions with external fields and absorbed species. While these properties should be derivable from a complete continuum theory of the surface, a well-defined SBC for the termination of a TI has so far been missing.

In this Rapid Communication we point out that this difficulty has a unique resolution due to the topological character of the interface between a TI and vacuum. We identify the physically relevant *topological boundary condition* (TBC) at the surface which correctly accounts for the band inversion at the TI-vacuum interface responsible for its protected surface states. As an example we apply the method to Bi_2Se_3, a well-studied topological insulator controlled by a single band inversion near its bulk Γ point. For a momentum linearized theory of its cleavage surface our method reproduces its well-studied low-energy surface state dispersion and spin texture. Extending our method to include its symmetry allowed quadratic terms reveals electronic physics that is unexpectedly rich: The topological surface bands develop a pronounced curvature and the Dirac point is displaced from the gap center, as seen experimentally. Notably, we find that the both Dirac point energy and its associated spin texture are crystal-face dependent due to the $R\bar{3}m$ bulk symmetry of this material. We find that these properties can be compactly encoded in the algebraic structure of a general surface state Hamiltonian that couples the orbital and spin degrees of freedom for an *arbitrary* crystal face. We further identify the modification of these features by surface-localized potentials that are absent from previous bulk-derived theories. Of the possible surface potentials that are allowed by $\mathcal{T}$ symmetry we find that the four-band model admits only two that affect the TI surface states. The combination of the TBC and these two surface terms provides an analytically tractable two-parameter family of Hamiltonians that dictate the surface electronic spectra for an arbitrary crystal termination of a TI.

We start from a description of the low-energy model of Bi_2Se_3, which applies generally to other TIs with the same crystal structure with space group $R\bar{3}m$. At the origin of Bi_2Se_3 Brillouin zone Γ, the effective Hilbert space near the bulk gap is spanned by states with angular momentum $m_j = \pm\frac{1}{2}$ and parity $\mathcal{P} = \pm 1$. Because of the spin-orbit coupling (SOC), $|p_z\uparrow\rangle$ and $|p_z\downarrow\rangle$ states are mixed with $|p_+\downarrow\rangle$ and $|p_-\uparrow\rangle$ states, respectively, however, since the crystal-field splitting is much larger than the SOC, p_z orbitals dominate and m_j pseudospin is proportional to the electron spin. The $\mathcal{P} = \pm 1$ hybridized states can be labeled approximately by the $|+\rangle$ state from Bi atoms while $|-\rangle$ from Se, due to the large energy difference between $4p$ (Se) and $6p$ (Bi) principal quantum levels.

Besides $\mathcal{T}$ and the parity inversion ($\mathcal{P}$) symmetries, the Bi_2Se_3 crystal structure has threefold rotational ($\mathcal{C}_3$) symmetry along $\hat{z}$ perpendicular to the quintuple layers, and twofold rotational ($\mathcal{C}_2$) symmetry along the ΓM direction. By convention we choose the parity operator $\mathcal{P} = \tau_z$ and the time reversal operator $\mathcal{T} = iK\sigma_y$, where K is the complex conjugate operation. Therefore, to quadratic order in k, the $\boldsymbol{k}\cdot\boldsymbol{p}$ bulk

Hamiltonian that preserves the above four symmetries has a unique form:

$$\begin{aligned}\mathcal{H} =& \left(c_0 + c_z k_z^2 + c_\parallel k_\parallel^2\right) + \left(-m_0 + m_z k_z^2 + m_\parallel k_\parallel^2\right)\tau_z \\ &+ v_z k_z \tau_y + v_\parallel (k_y \sigma_x - k_x \sigma_y)\tau_x,\end{aligned} \tag{1}$$

where we assume $m_z, m_\parallel, v_z, v_\parallel > 0$ and $\hbar = 1$ hereafter. The first parentheses is a scalar term that preserves $\mathcal{T}$ and $\mathcal{P}$ symmetries but breaks the particle-hole (p-h) symmetry. The quadratic scalar terms intrinsically give rise to the curvature of Dirac surface bands and the rigid shift of Dirac point from the middle of the bulk gap, as we will demonstrate in the following. Equation (1) has both $\mathcal{T}$ and $\mathcal{P}$ symmetries with a topological index $\mathcal{Z}_2 = \mathrm{sgn}(-m_0)$.

Topological boundary condition. The symmetry protected physics at the TI surface occurs because of the sign reversal of the mass term in Eq. (1) across an interface with vacuum. Accordingly we study a model in which the TI side has $m_0 = m$ positive and half of the bulk gap at $\boldsymbol{k} = 0$, and $c_0 = 0$ to define the middle of the $\boldsymbol{k} = 0$ gap as energy zero. The vacuum side is a trivial insulator with an infinite gap, i.e., $m_0 = -M$, where $M \to +\infty$. Both c_z and $c_\parallel$ vanish since the vacuum side is dominated by M and can be regarded as p-h symmetric around c_0.

It is useful to consider the one-dimensional quantum mechanics of the $(00\bar{1})$ surface in which we turn off the quadratic terms in Eq. (1). Note that the two spin flavors are decoupled for this eigenvalue problem. The spin $\uparrow$ states in $\{\tau_z, \sigma\}$ representation are determined by the following coupled Dirac equations:

$$\begin{pmatrix} -m(z) - E & -v_z \partial_z \\ v_z \partial_z & m(z) - E \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = 0. \tag{2}$$

The wave function $\psi_\uparrow = (\psi_1, \psi_2)'$ is continuous across the surface, integrating Eq. (2) over the vicinity of the surface. This continuity condition leads to a nontrivial solution which is isolated in the middle of the bulk gap and is localized on the surface. This midgap ($E = 0$) state has a simple and elegant exact solution,

$$\begin{aligned}\psi_{k_\parallel,\uparrow}(x,y,z) &= \frac{1}{A} e^{i(k_x x + k_y y)} \phi(z) \begin{pmatrix} 1 \\ 0 \end{pmatrix}_\sigma \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\tau_x}, \\ \phi(z) &= \begin{cases} e^{-\kappa z}, & z > 0 \quad \text{(TI)}, \\ e^{\kappa_0 z}, & z < 0 \quad \text{(vac)}, \end{cases}\end{aligned} \tag{3}$$

where $\kappa = m/v_z$, $\kappa_0 = M/v_z$, and A is a normalization factor. $\phi(z)$ is evanescent on both sides of the surface where the mass m_0 changes sign, analogous to the Jackiw and Rebbi (JR) solution[19] of a two-band Dirac model. The spin $\downarrow$ solution can be obtained by $\psi_\downarrow = \mathcal{T}\psi_\uparrow$. This isolated midgap state at $k_\parallel = 0$ is identified as the Dirac point and at finite $k_\parallel$ spreads into a perfect Dirac cone which forms the ideal topologically protected surface bands. Note that the evanescent solutions of the JR model in the exterior region provide the appropriate boundary condition for terminating of the bulk states of Eq. (1). This TBC applies even when the quadratic terms are nonzero.

Surface state spin texture. On the $(00\bar{1})$ surface, the midgap solution of the boundary problem at $k_\parallel = 0$ is determined by the operators $\boldsymbol{\tau}$ and is free under any rotation of the operators $\boldsymbol{\sigma}$. This $\boldsymbol{\tau} \otimes \boldsymbol{\sigma}$ algebra of the boundary problem can be generalized to any surface with $\boldsymbol{\tau}$ replaced by $\boldsymbol{S}_1$ and

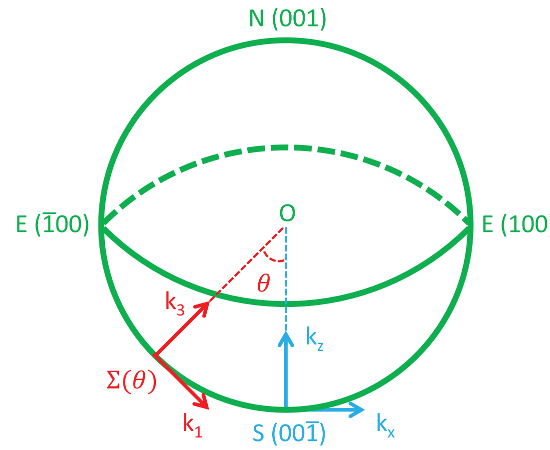

FIG. 1. (Color online) The definition of an arbitrary face of a Bi_2Se_3-type topological insulator. The south (S) pole denotes the $(00\bar{1})$ surface parallel to the quintuple layers and the equator represents the side faces perpendicular to the quintuple layers. For any face $\Sigma(\theta)$, $\hat{k}_3 \perp \Sigma$, and $k_2 = k_y$.

$\boldsymbol{\sigma}$ by $\boldsymbol{S}_2$. For an arbitrary surface $\Sigma(\theta)$ defined in Fig. 1, the algebraic structure is

$$\begin{aligned}\boldsymbol{S}_1 &= \{\alpha\tau_x + \beta\sigma_y\tau_y, \alpha\tau_y - \beta\sigma_y\tau_x, \tau_z\}, \\ \boldsymbol{S}_2 &= \{\alpha\sigma_x - \beta\sigma_z\tau_z, \sigma_y, \alpha\sigma_z + \beta\sigma_x\tau_z\},\end{aligned} \tag{4}$$

where $v_3 = \sqrt{(v_z\cos\theta)^2 + (v_\parallel \sin\theta)^2}$, $\alpha = v_z\cos\theta/v_3$, and $\beta = v_\parallel \sin\theta/v_3$. These pseudospins satisfy $[S_a^i, S_b^j] = 2i\delta_{ab}\epsilon^{ijk}S_a^k$. Rewritten in this pseudospin basis, Eq. (1) reads

$$\begin{aligned}\mathcal{H} =& -m_0 S_1^z + (v_3 k_3 + v_0 k_1) S_1^y \\ &+ \left(v_\parallel k_y S_2^x - v_1 k_1 S_2^y\right) S_1^x,\end{aligned} \tag{5}$$

where $v_0 = (v_\parallel^2 - v_z^2)\sin\theta\cos\theta/v_3$ and $v_1 = v_z v_\parallel/v_3$. By matching the eigensystems of the TI and vacuum sides, we obtain the $\Sigma(\theta)$ surface states similar to Eq. (3) where $\boldsymbol{\tau}$ and $\boldsymbol{\sigma}$ are replaced by $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$, respectively. Note that $\kappa = m/v_3 + i v_0 k_1/v_3$ in general. k_1 coupling to S_1^y neither influences the energy spectrum nor the spin texture, only making the evanescent states oscillate and giving a phase accumulation along $\hat{k}_1$ away from the Dirac point. Therefore, ignoring the quadratic corrections, we first derive the effective surface state Hamiltonian for an arbitrary face $\Sigma(\theta)$ to the linear order,

$$\mathcal{H}^{(1)}(\theta) = v_\parallel k_y S_2^x - v_1 k_1 S_2^y, \tag{6}$$

from which we can further explicitly demonstrate the surface state spin texture on $\Sigma(\theta)$:

$$\begin{aligned}\langle\sigma_x\rangle_\theta &= \pm\frac{v_z v_\parallel k_y \cos\theta}{v_3\sqrt{v_1^2 k_1^2 + v_\parallel^2 k_y^2}}, \\ \langle\sigma_y\rangle_\theta &= \pm\frac{-v_z v_\parallel k_1}{v_3\sqrt{v_1^2 k_1^2 + v_\parallel^2 k_y^2}}, \\ \langle\sigma_z\rangle_\theta &= 0,\end{aligned} \tag{7}$$

where $+$ $(-)$ denotes the conduction (valence) band. The electron real spin $\langle s\rangle$ is proportional to but always smaller than $\langle\boldsymbol{\sigma}\rangle$ due to the SOC.[20] In the local coordinates,

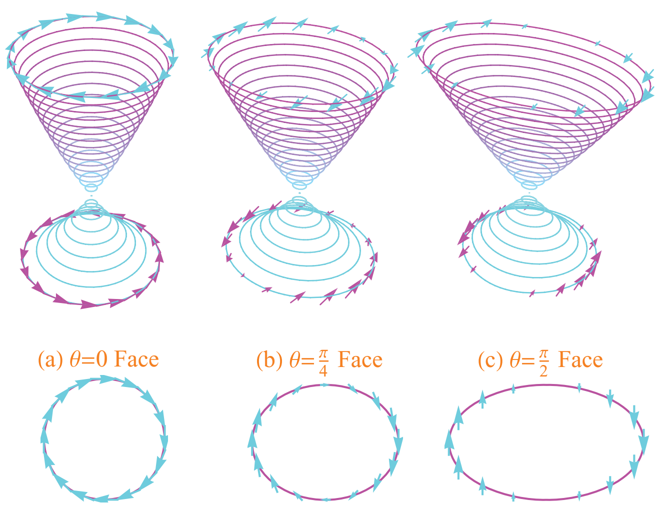

FIG. 2. (Color online) Dirac cones and spin textures of surface states on faces with (a) $\theta = 0$, (b) $\theta = \pi/4$, and (c) $\theta = \pi/2$. The equal-energy contours are from -80 to 160 meV with 10 meV increments relative to the Dirac point. Surface band curvatures are taken into account. All three panels are in the same scale and the parameters are adopted from Ref. 7. The lower panels only show the spin textures at 160 meV in the k_1-k_y planes, where k_y is the vertical axis.

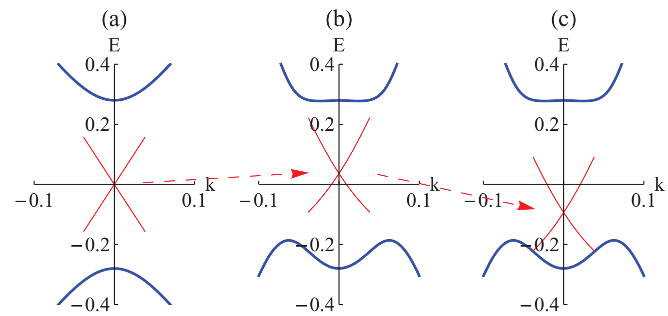

FIG. 3. (Color online) Band curvature and Dirac point position of the (001) surface state (red) in the bulk (blue) gap. (a) plots the linear theory Eq. (6), (b) includes quadratic corrections in Eqs. (8) and (9), and (c) further considers a surface potential with $\Delta_E = -m/8$ in Eq. (10). E and k are in units of eV and Å^{-1}. Parameters are adopted from Ref. 7.

$\langle\sigma_1\rangle_\theta = \langle\sigma_x\rangle_\theta \cos\theta$ and $\langle\sigma_3\rangle_\theta = \langle\sigma_x\rangle_\theta \sin\theta$. Equation (7) indicates that while the pseudospin (S_2) texture has a universal structure on the elliptic constant energy contour near the Dirac point, the surface state spin (σ) texture is rather different from face to face. The absence of spherical symmetry in the bulk Hamiltonian requires that the spin and orbital structures are crystal-face dependent on the surface of a TI. The spin texture is helical on the south and north poles, though compressed to a single dimension along the equator, and tilted out of plane for a general crystal surface. The surface state anisotropy and spin textures for different faces are compared in Fig. 2.

The surface state spin textures on the poles and the equator of a TI sphere can be understood by symmetries. Note that the *bulk* Hamiltonian Eq. (1) has no coupling to σ_z in the crystal frame due to the mirror symmetry with respect to $\Gamma M(\hat{x})$. For the $(\bar{1}00)$ face, the $\mathcal{C}_2$ symmetry along the ΓM direction forbids σ_x coupling to any in-plane momentum. In linear order $\mathcal{C}_3$ symmetry upgrades to continuous rotational symmetry, consequently, the surface normal spin is zero along the TI equator. On the two poles, the mirror symmetry with respect to ΓM and the $\mathcal{C}_3$ symmetry along the $\hat{z}$ direction ensure that spin and momentum lock into the form of $k_y\sigma_x$-$k_x\sigma_y$.

It is interesting to point out that the surface band is the positive eigenstate of S_1^x and the chiral counterpart is separated and localized on the opposite face. Thus the surface state Hilbert space is reduced by *half* and this pseudospin polarity (or chirality) blocks the backscattering on the surface as a result of the interplay between $\mathcal{T}$ symmetry and band inversion physics with TBC.

Dirac point energy and surface potentials. The absence of spherical symmetry in the bulk also requires that the Dirac point has different energies on different crystal faces. When the quadratic mass terms and *p-h* symmetry breaking terms in Eq. (1) are introduced, the surface state wave function and spectrum are changed. In the coupled Schrödinger-type Eq. (1), the components of the wave functions and their slopes are all continuous across the boundary. We find that the effective Hamiltonian for face $\Sigma(\theta)$ inherit two important[21] corrections:

$$\mathcal{H}^{(2)}(\theta) = c_\parallel k_y^2 + (c_z \sin^2\theta + c_\parallel \cos^2\theta)k_1^2, \tag{8}$$

$$\mathcal{H}^{\mathrm{DP}}(\theta) = \frac{c_z \cos^2\theta + c_\parallel \sin^2\theta}{m_z \cos^2\theta + m_\parallel \sin^2\theta} m. \tag{9}$$

The breaking *p-h* symmetry terms give the surface Dirac cone a parabolic curvature described by Eq. (8) and shifts the Dirac point from the midgap to a nonzero energy given by Eq. (9). These two effects are responsible for the shape of the cleavage surface bands observed in ARPES experiments,[8,9,22–24] as illustrated in Fig. 3.

Although the surface state solution obtained by Eq. (2) with TBC remains topologically stable, it is essential to understand how localized potentials that arise from surface reconstruction or adsorbed species can influence the surface spectra and the associated wave functions as well. We focus on potentials that preserve $\mathcal{T}$ symmetry as well as the lattice translational symmetry in the plane of the surface. Among these six types, I, τ_z, τ_x, and $\vec{\sigma}\tau_y$ listed in Table I, τ_x and $\vec{\sigma}\tau_y$ break $\mathcal{P}$ symmetry. It turns out that the same potential could play different roles on different crystal faces, as shown in Table I.

On the south pole of a TI sphere, a potential $\Delta_0\delta(z)2v_z/mI$ changes the wave function continuity condition obtained by integrating Eq. (2), including surface potentials over the vicinity of $z = 0$. The amplitudes of ψ_1 and ψ_2 are either weakened or enhanced across the surface, with changes always opposite to each other. Consequently, to match the evanescent solutions on the vacuum and the matter sides, the surface band energies are rigidly shifted by $\delta\mathcal{E}_{\mathrm{DP}}$. τ_x potentials have similar effects to the I type although τ_x breaks $\mathcal{P}$ symmetry. For a τ_z potential, it simply modifies the mass term on the surface, and it is not surprising that it does not affect the surface spectra. A τ_z potential tunes the amplitudes of ψ_1 and ψ_2 in the same manner on the matter side, but this gives no observable effect since $\psi_1(z) = \psi_2(z)$ and $\psi_{1,2}(z < 0) \to 0$ are still valid. Similarly, a potential such as $\sigma_n\tau_y$ ($\Delta_n\sigma_n \equiv \vec{\Delta}\cdot\vec{\sigma}$) does nothing to the surface spectra and ψ_1/ψ_2; however, it couples the two spin flavors and shifts their phase with opposite signs. Note that

TABLE I. Summary of the influence of $\mathcal{T}$ symmetry allowed momentum-independent surface potentials $\Delta\delta(r_3)2v_3/m$ on the inversion symmetry, and the wave-function (Ref. 25) continuity and the Dirac point energy of surface states. This Δ represents different surface potentials and their corresponding energy scales: $\Delta_0 I$, $\Delta_m\tau_z$, $\Delta_E S_1^x$, and $\Delta_n S_2^n S_1^y$. For the $\sigma_n\tau_y$ column, only the results for the $S_2^n = 1$ state are shown and their complex conjugates represent the results for the $S_2^n = -1$ state.

$\Sigma(0)$	I	τ_z	τ_x	$\sigma_n\tau_y$
$\Sigma(\theta)$	I	τ_z	S_1^x	$S_2^n S_1^y$
$\mathcal{P}$	$+$	$+$	$-$	$-$
$\delta\mathcal{E}_{DP}$	$\frac{4m^2\Delta_0(m^2-\Delta_0^2)}{(m^2-\Delta_0^2)^2+4m^2\Delta_0^2}$	0	$\frac{4m^2\Delta_E(m^2+\Delta_E^2)}{(m^2+\Delta_E^2)^2+4m^2\Delta_E^2}$	0
$\frac{\psi_1(0^+)}{\psi_2(0^+)}$	$\frac{m^2-2m\Delta_0-\Delta_0^2}{m^2+2m\Delta_0-\Delta_0^2}$	1	$\left(\frac{m-\Delta_E}{m+\Delta_E}\right)^2$	1
$\frac{\psi_1(0^+)}{\psi_1(0^-)}$	$\frac{m^2-2m\Delta_0-\Delta_0^2}{m^2+\Delta_0^2}$	$\frac{m+\Delta_m}{m-\Delta_m}$	$\frac{m-\Delta_E}{m+\Delta_E}$	$\frac{m-i\Delta_n}{m+i\Delta_n}$
$\frac{\psi_2(0^+)}{\psi_2(0^-)}$	$\frac{m^2+2m\Delta_0-\Delta_0^2}{m^2+\Delta_0^2}$	$\frac{m+\Delta_m}{m-\Delta_m}$	$\frac{m+\Delta_E}{m-\Delta_E}$	$\frac{m-i\Delta_n}{m+i\Delta_n}$

$\psi_1 = \psi_2$ (Ref. 25) is always true on the vacuum side since $M \to \infty$. The above results are summarized in Table I, providing sufficient information for constructing the surface state wave function and to engineer the Dirac point energy position.

On an arbitrary face $\Sigma(\theta)$, the types of surface potentials are the same but their combinations and the corresponding roles are rearranged. This can be fully understood by the fact that the spin-orbital structure $\boldsymbol{\tau}\otimes\boldsymbol{\sigma}$ on the south pole is replaced by a pseudospin-pseudospin structure $\boldsymbol{S}_1(\theta)\otimes\boldsymbol{S}_2(\theta)$ on $\Sigma(\theta)$. Note that the combination only occurs for potentials with the same parity.

Although the surface state solutions are stable in the presence of localized surface potentials that preserve $\mathcal{T}$ symmetry, Table I shows that the two terms proportional to I and S_1^x play an essential role in determining the energy *position* of the Dirac point,

$$\mathcal{E}_{DP} = \mathcal{H}^{DP} + \frac{4m^2(\Delta_0+\Delta_E)\left(m^2-\Delta_0^2+\Delta_E^2\right)}{4m^2(\Delta_0+\Delta_E)^2+\left(m^2-\Delta_0^2+\Delta_E^2\right)^2}, \qquad (10)$$

which implies that I and S_1^x potentials are able to tune the Dirac point from the midgap to the band edges $\pm m$. This effect is illustrated in Fig. 3.

Discussions. The interactions of the topologically protected bands with surface potentials provide a robust route to engineering and manipulating the topological surface states. In particular, an external surface potential can raise or lower the Dirac point to the middle of the bulk gap, providing experimental access to the topologically protected band. This goal could be achieved by surface oxidation,[26] or by other possible chemical processes,[27] perturbations,[28] and interactions[29] on the surface. We also point out that the electrostatic gating[22,27,30,31] $E_{eff}\langle r_3\rangle_{sf}$ can act to Stark shift the TI surface state into the gap, with a field strength determined by the penetration length of the surface states. Typically, the screened field $\sim$100 meV nm^{-1} is required for the surface states that decay in a couple of quintuple layers. Our present results provide a framework to study the self-consistent band bending physics of real TI materials[8,9,22–24] and mean-field models of surface state many-body interactions.

Our model is quite different from the fixed boundary condition (FBC) which arbitrarily clamps the surface state wave function to zero at the boundary. The FBC solution is actually *insensitive* to the mass inversion at the surface which topologically protects the surface bands, and consequently it cannot be used to calculate the energy of the symmetry protected degeneracy relative to the bulk bands or their interaction with surface-localized potentials. Indeed, the energy of the Dirac degeneracy is used as an *input* to these theories. Furthermore, FBC admits an infinite number of (physically spurious) solutions with nonzero energies for $k_1 = k_2 = 0$ that satisfy an (incorrect) surface boundary condition. By contrast TBC guarantees that there is only one isolated solution, i.e., the Dirac point of surface bands protected by the change in bulk topology. Moreover, TBC demonstrates that the energy position of the Dirac point is tunable in the bulk gap via the symmetry allowed scalar terms and surface potentials.

On a noncleavage surface, dangling bonds and their reconstruction may add complexity to the surface state spectrum which can be modeled as surface potentials encoded in the parameters Δ_0 and Δ_E. Our formulation for the surface potentials is applicable within the four-band model whenever these perturbations are spatially localized to the surface region (much smaller than the decay length for the TI evanescent states) and are energy independent. Since the bulk gap is small in these materials, this will generally be the case for most surface perturbations. However, for situations where the perturbing potential penetrates the bulk, either by the propagation of electrostatic or strain fields, one needs to replace the surface matching condition by a scheme in which the bulk and vacuum evanescent states are separately matched to electronic states calculated within a finite depth near the surface. In principle, these states can be obtained numerically, in which case Table I gives the symmetry allowed *spatially varying* potentials that can enter the theory. Note that if the penetration depth is sufficiently large, this can introduce a strongly energy-dependent matching condition, reflecting the presence of bound and resonant states in the intermediate region. This physics is responsible for the quantum well states that arise from strong band bending near a TI surface.[27]

The novel spin texture that we predict near the Dirac point on noncleavage surfaces is determined by the bulk symmetries along with the topological stability of the surface spectrum and may be accessible in ARPES and STM experiments. On an equatorial face, Zeeman coupling or magnetic disorder coupled to the spin degree of freedom do not generally open a gap, and the diamagnetic susceptibility is anticipated to be unusually anisotropic. More spectacularly, there is intrinsic charge redistribution in the surface bands near the corners of a TI that connect different crystal faces with intrinsically different Dirac point energies.

Finally, we point out that our theory can be applied to many other topological states of matter that are related to band inversion (or mass reversal), besides $\mathcal{T}$ symmetry protected strong TIs. A recent example is SnTe, a topological crystalline insulator,[32] where band inversion occurs at four bulk L points and its surface states are protected by mirror symmetry.

Acknowledgment. This work has been supported by DARPA under Grant No. SPAWAR N66001-11-1-4110.

*zhf@sas.upenn.edu

[1]C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).

[2]L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).

[3]J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306 (2007).

[4]R. Roy, Phys. Rev. B **79**, 195322 (2009).

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

[6]X. Qi and S. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).

[7]H. Zhang, C. Liu, X. Qi, X. Dai, Z. Fang, and S. Zhang, Nat. Phys. **5**, 438 (2009).

[8]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).

[9]Y. Chen, J. G. Analytis, J. Chu, Z. Liu, S. Mo, X. Qi, H. Zhang, D. Lu, X. Dai, Z. Fang, S. Zhang, I. R. Fisher, Z. Hussain, and Z. Shen, Science **325**, 178 (2009).

[10]P. Roushan, J. Seo, C. V. Parker, Y. S. Hor, D. Hsieh, D. Qian, A. Richardella, M. Z. Hasan, R. J. Cava, and A. Yazdani, Nature (London) **460**, 1106 (2009).

[11]Z. Alpichshev, J. G. Analytis, J. H. Chu, I. R. Fisher, Y. L. Chen, Z. X. Shen, A. Fang, and A. Kapitulnik, Phys. Rev. Lett. **104**, 016401 (2010).

[12]L. Fu, Phys. Rev. Lett. **103**, 266801 (2009).

[13]C. X. Liu, X. L. Qi, H. J. Zhang, X. Dai, Z. Fang, and S. C. Zhang, Phys. Rev. B **82**, 045122 (2010).

[14]J. Linder, T. Yokoyama, and A. Sudbo, Phys. Rev. B **80**, 205401 (2009).

[15]H. Z. Lu, W. Y. Shan, W. Yao, Q. Niu, and S. Q. Shen, Phys. Rev. B **81**, 115407 (2010).

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

[17]P. G. Silvestrov, P. W. Brouwer, and E. G. Mishchenko, arXiv:1111.3650.

[18]A. Medhi and V. B. Shenoy, arXiv:1202.3863.

[19]R. Jackiw and C. Rebbi, Phys. Rev. D **13**, 3398 (1976).

[20]O. V. Yazyev, J. E. Moore, and S. G. Louie, Phys. Rev. Lett. **105**, 266806 (2010).

[21]The crossing terms $\propto k_1 k_3 \sin 2\theta$ have negligible effects and thus are ignored for simplicity. Equation (9) is an exact result at the Γ point while m gets reduced by the $m_1 k_1^2$ and $m_\parallel k_y^2$ terms away from $k = 0$.

[22]D. Kong, Y. Chen, J. Cha, Q. Zhang, J. G. Analytis, K. Lai, Z. Liu, S. Hong, K. J. Koski, S. Mo, Z. Hussain, I. Fisher, Z. Shen, and Y. Cui, Nat. Nanotechnol. **6**, 705 (2011).

[23]J. Zhang, C. Chang, Z. Zhang, J. Wen, X. Feng, K. Li, M. Liu, K. He, L. Wang, X. Chen, Q. Xue, X. Ma, and Y. Wang, Nat. Commun. **2**, 574 (2011).

[24]T. Arakane, T. Sato, S. Souma, K. Kosaka, K. Nakayama, M. Komatsu, T. Takahashi, Z. Ren, K. Segawa, and Y. Ando, Nat. Commun. **3**, 636 (2012).

[25]$(\psi_1,\psi_2)'$ is the positive eigenstate of S_1^x in the $\{S_1, S_1^z\}$ representation.

[26]X. Wang, G. Bian, T. Miller, and T. Chiang, Phys. Rev. Lett. **108**, 096404 (2012).

[27]M. Bianchi, R. C. Hatch, J. Mi, B. B. Iversen, and Ph. Hofmann, Phys. Rev. Lett. **107**, 086802 (2011).

[28]N. Fukui, T. Hirahara, T. Shirasawa, T. Takahashi, K. Kobayashi, and S. Hasegawa, Phys. Rev. B **85**, 115426 (2012).

[29]O. V. Yazyev, E. Kioupakis, J. E. Moore, and S. G. Louie, Phys. Rev. B **85**, 161101(R) (2012).

[30]H. Steinberg, D. R. Gardner, Y. Lee, and P. Jarillo-Herrero, Nano Lett. **10**, 5032 (2010).

[31]J. G. Checkelsky, Y. S. Hor, R. J. Cava, and N. P. Ong, Phys. Rev. Lett. **106**, 196801 (2011).

[32]T. H. Hsieh, H. Lin, J. Liu, W. Duan, A. Bansil, and L. Fu, Nature Communications **3**, 982 (2012).