Spin and charge structure of the surface states in topological insulators

P. G. Silvestrov,[1] P. W. Brouwer,[1] and E. G. Mishchenko[2]

[1]*Physics Department and Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany*
[2]*Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah 84112, USA*

We investigate the spin and charge densities of surface states of the three-dimensional topological insulator Bi_2Se_3, starting from the continuum description of the material [Zhang *et al*., Nat. Phys. **5**, 438 (2009)]. The spin structure on surfaces other than the (111) surface has additional complexity because of a misalignment of the contributions coming from the two sublattices of the crystal. For these surfaces we expect new features to be seen in the spin-resolved angular resolved photoemission spectroscopy (ARPES) experiments, caused by a nonhelical spin polarization of electrons at the individual sublattices as well as by the interference of the electron waves emitted coherently from two sublattices. We also show that the position of the Dirac crossing in the spectrum of the surface states depends on the orientation of the interface. This leads to contact potentials and surface charge redistribution at edges between different facets of the crystal.

I. INTRODUCTION

A new class of materials, topological insulators, have attracted a great deal of attention in the last two years (see Refs. 1 and 2 for reviews). Typically these are band insulators for which strong spin-orbit coupling leads to an inversion of the bulk band gap. The band inversion[3] causes the emergence of protected two-dimensional states on the surface of this, in other respects, conventional bulk insulator.

The surface states of a topological insulator have a conical energy spectrum, characteristic of eigenmodes of the massless Dirac equation. They are routinely described by the effective two-dimensional Hamiltonian[4]

$$H_S = v_F(\sigma_x p_y - \sigma_y p_x),\qquad(1)$$

where v_F is the Fermi velocity and σ_x and σ_y are Pauli matrices. The operator $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is usually identified with the true electron spin.[5] The spin polarization is then perpendicular to the electron's momentum, which leads to the surface states being referred to as "helical."[6] Yet, we are not aware of a detailed investigation of the relation between $\boldsymbol{\sigma}$ and the true spin. The nature of $\boldsymbol{\sigma}$ is not important for a number of phenomena involving the surface states, in particular to those that only depend on the nontrivial Berry phase acquired by surface-state electrons.[7,8] However, for the properties that explicitly probe the electron's spin, the nature of operators in the effective Hamiltonian becomes of principal importance.[9,10] In this paper we elucidate how the spin structure of the surface electrons follows from the bulk Hamiltonian of the topological insulator and shows that the relation between spin and momentum is richer and more delicate than suggested by a naive interpretation of Eq. (1).

As a specific material we will consider the compound Bi_2Se_3, presently probably the most popular example of a topological insulator. This material was considered both theoretically and experimentally in Refs. 11 and 12, where it was shown that Bi_2Se_3 has a large topologically nontrivial band gap ∼0.3 eV, leading to surface states with a single Dirac cone. Zhang and co-workers[12,13] also suggested a simple three-dimensional Hamiltonian describing the long-wavelength electronic dynamics in Bi_2Se_3. This low-energy

Hamiltonian contains both the true electron spin and a pseudospin as its primary degrees of freedom, where the pseudospin refers to states with support on the Bi and Se sublattices.

Bi_2Se_3 is a strongly anisotropic material, with a layered structure involving quintuple layers of Bi and Se atoms. (One quintuple layer consists of three Se layers strongly bonded to two Bi layers in between.) Yet, by virtue of their topological protection, the surface states exist for arbitrarily oriented crystal surfaces, not only at the (111) surface parallel to the quintuple layer. Based on the low-energy continuum Hamiltonian of Zhang *et al*.,[12] here we show that the σ operator entering the surface Hamiltonian Eq. (1) coincides with the electron's spin for the (111) surface only. For any other surface orientation, the spin content of the electron wave function components is different on different sublattices, see Fig. 1. Whereas most experiments are carried out for the (111) surface of Bi_2Se_3 because the material cleaves well in this direction, other surfaces are also realized, for example, in TI nanoribbons in the experiment of Ref. 7. Our findings are relevant for those nonstandard surfaces.

Experimentally, the dispersion of surface states is found through angular resolved photoemission spectroscopy (ARPES).[14] Spin-resolved ARPES provides momentum-resolved information about the spin polarization of the surface states in topological insulators.[15–17] In Sec. III we consider the electron's out-of-plane polarization measured via ARPES, which may be caused by a spin structure of the surface states that is more complicated than the spin structure suggested by Eq. (1). The out-of-plane polarization may arise from the interference effects of photoelectrons emitted from the two sublattices for surfaces other than the standard (111) surface, or from photoelectrons emitted from Se atoms for surfaces that are neither parallel nor perpendicular to the plane of the quintuple layers.

Another feature of realistic three-dimensional topological insulators, that is easily overlooked in the effective surface Hamiltonian (1) (but present in the effective low-energy Hamiltonian of Ref. 12), is that the combination of broken particle-hole symmetry and angular anisotropy of the bulk Hamiltonian leads to different energies of the Dirac points at
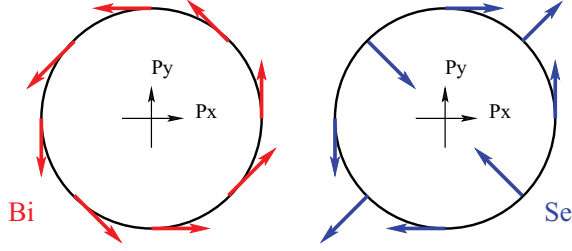
FIG. 1. (Color online) Polarization of surface-state electrons on the Bi and Se sublattices for different orientations of the in-plane momentum p for a surface orthogonal to the quintuple layer. Spin textures for a symmetric model [Eq. (2)] are shown to demonstrate the inevitable strong differences between the two sublattices spin for surfaces other than (111). Details of the spin behavior for the asymmetric case are given in Sec. IV.

different facets of the topological insulator crystal. Different positions of the Dirac crossing (or the neutrality point) with respect to Fermi energy would result in different surface electrons densities at different surfaces of the crystal. However, an uncompensated charge density at one or more crystal surfaces comes at a large electrostatic energetic cost, and one expects a charge redistribution in order to reduce the Coulomb interaction energy. Below we describe the solution of the corresponding electrostatic problem at an edge of a crystal, at the intersection of two differently oriented surfaces. For short facets, where the Coulomb interaction energies are not large enough to induce full screening, a finite carrier density may remain, illustrating the impossibility to tune the full topological insulator surface to the Dirac point by means of doping or gating.

II. SURFACE STATES

We first discuss a simplified version of the effective low-energy continuum Hamiltonian of Ref. 12, which consists of a three-dimensional Dirac-like Hamiltonian with momentum dependent mass,

$$H = \tau_z \mathcal{M}(\mathbf{K}) + \tau_x (\sigma \cdot \mathbf{K}) A, \ \ \mathcal{M}(\mathbf{K}) = M - BK^2, \quad (2)$$

where we use the capital $\mathbf{K} = (K_x, K_y, K_z)$ for the three-dimensional momentum. The z axis is chosen in the (111) direction, that is, perpendicular to the plane of the quintuple layers. The vectors $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ and $\tau = (\tau_x, \tau_y, \tau_z)$ contain two sets of Pauli matrices operating in different pseudospin spaces. The matrices τ_x, τ_y, and τ_z refer to the $P1_z^+$ and $P2_z^-$ orbitals of Refs. 12 and 13, which are such that states with $\tau_z = 1$ have support mostly on the Bi sublattice and states with $\tau_z = -1$ have support mostly on the Se sublattice. Below we will refer to these $\tau_z = 1$ and $\tau_z = -1$ states simply as states on the Bi and Se sublattices, respectively. The electron's spin is expressed in terms of the Pauli matrices as[13,18]

$$s_z = \frac{\sigma_z}{2}, \quad s_x = \frac{\sigma_x \tau_z}{2}, \quad s_y = \frac{\sigma_y \tau_z}{2}. \quad (3)$$

(The correct form of the spin operator for Bi$_2$Se$_3$ is missing in Ref. 12.) Equation (3) implies that the operators $\sigma/2$ and $\mathbf{s}$ coincide for the electron residing on the Bi sublattice, whereas

the relation between $\mathbf{s}$ and $\sigma/2$ involves a rotation by a π angle around the z axis on the Se sublattice.

The Hamiltonian of Eq. (2) has a higher symmetry than the true low-energy Hamiltonian of Bi$_2$Se$_3$: It is not only valence-band–conduction-band (particle-hole) symmetric, but also symmetric under spatial rotations, provided the Pauli matrices σ are rotated together with the coordinates, whereas the Pauli matrices τ are kept fixed. The rotation symmetry is a mathematical artifact of the pseudospin basis used in Eq. (2). The strong asymmetry of the real crystal of Bi$_2$Se$_3$, in which the z axis plays a special role, is reflected in the anisotropic assignment of the spin operators s_x, s_y, and s_z in Eq. (3). Corrections to this simplified model Hamiltonian will be discussed in Sec. IV. Note that a direct association of the spin with σ in Eq. (2) would violate the parity of the true system. That is why one needs to introduce the asymmetric spin operator with the special axis z [Eq. (3)] even in case of the most symmetric model Hamiltonian.[13,18]

The Hamiltonian Eq. (2) is of the second order in spatial derivatives (momentum). Following Ref. 12, as well as a number of subsequent articles,[9,10,19,20] we require that all four components of the wave function vanish at the crystal surface. This choice of the boundary conditions guarantees the existence of a branch of Dirac-like surface states with Dirac crossing at the Γ point (the point of vanishing in-plane momentum), even in the strongly asymmetric case of Sec. IV. We should point out, however, that this boundary condition is not unique, and that other choices of boundary conditions have been advocated in the literature. References 18, 21, and 22 suggest using an effective Hamiltonian that is linear in momentum, with boundary conditions for which only one pseudospin component of the wave function vanishes at the surface. The two choices of boundary conditions agree qualitatively for the standard termination at the (111) surface (both give a Dirac-like branch of surface states near the Γ point), but it is not obvious how to extend the boundary condition of Refs. 18, 21, and 22 to surfaces of arbitrary orientation, and the two boundary conditions may give different predictions for other surfaces or nonstandard surface terminations.[22]

For the explicit calculation of the surface states from the model Hamiltonian (2) it is convenient to rotate the coordinate system, such that the TI fills the half-space $z < 0$. This rotation does not change Eq. (2), but it changes the relation between the spin operators and the σ matrices: In the rotated coordinate system, Eq. (3) takes the form

$$\mathbf{s} = \frac{(\sigma \cdot \mathbf{n})\mathbf{n}}{2} + \frac{\sigma - (\sigma \cdot \mathbf{n})\mathbf{n}}{2} \tau_z, \quad (4)$$

where $\mathbf{n}$ is the unit vector pointing in the direction perpendicular to the quintuple layer plane in the rotated coordinate frame.

Let us introduce two two-component spinor functions ψ_B and ψ_S that describe the pseudospin $\tau_z = 1$ and $\tau_z = -1$ components of the four-component wave function. Eigenfunctions of the Hamiltonian Eq. (2) at energy ε that correspond to surface states with momentum $\mathbf{p}$ parallel to the surface should be found as linear combinations of the functions

$$\psi_B = u_B e^{i(p_x x + p_y y) + \lambda z}, \quad \psi_S = u_S e^{i(p_x x + p_y y) + \lambda z}, \quad (5)$$

with different values of λ (Re $\lambda > 0$). Here and below we use the lower-case symbol $\mathbf{p} = (p_x, p_y, 0)$ to denote the in-plane momentum of the surface states. The spinor amplitudes u_B and u_S satisfy the coupled system of equations

$$(\varepsilon - \mathcal{M}(p, \lambda))u_B = A(\sigma_x p_x + \sigma_y p_y - i\lambda\sigma_z)u_S,$$
$$(\varepsilon + \mathcal{M}(p, \lambda))u_S = A(\sigma_x p_x + \sigma_y p_y - i\lambda\sigma_z)u_B, \quad (6)$$

where $\mathcal{M}(p, \lambda) = M - Bp^2 + B\lambda^2$ and $p^2 = p_x^2 + p_y^2$. The inverse decay length can take two values $\lambda_{1,2}$ which are the solutions of the equation

$$\varepsilon^2 = (M - Bp^2 + B\lambda^2)^2 + A^2 p^2 - A^2 \lambda^2, \quad (7)$$

with Re $\lambda > 0$. In addition, the surface state satisfies the boundary condition $\psi_B(z = 0) = \psi_S(z = 0) = 0$, which provides one additional constraint from which the dispersion $\varepsilon(p)$ can be calculated.

To find the solution of Eqs. (6) and (7) one may first guess the (correct) result for the spectrum,

$$\varepsilon(p) = \pm Ap. \quad (8)$$

Then Eq. (7) yields a quadratic equation for λ,

$$A\lambda = \mathcal{M}(p, \lambda) = (M - Bp^2 + B\lambda^2), \quad (9)$$

where the parameters A and B are taken from the original Hamiltonian Eq. (2). The first equality here is the energy of the surface state consistent with the surface Hamiltonian Eq. (1) with $A = v_F$. The second equality gives the values of

$$\lambda_{1,2} = (A \pm \sqrt{A^2 + 4B^2 p^2 - 4MB})/2B, \quad (10)$$

consistent with this value of energy. Substitution of the ansatz (8) into Eq. (6) then gives identical spinors u_B and u_S for the two roots $\lambda_{1,2}$, which guarantees that an appropriate linear combination can be found that satisfies the boundary condition $\psi_B(z = 0) = \psi_S(z = 0) = 0$ for all four components of the spinor wave functions ψ_B and ψ_S simultaneously. The spinor structure corresponding to the cases $\varepsilon = \pm Ap$ is

$$u_B = \frac{1}{2}\begin{pmatrix} \pm i \\ \gamma \end{pmatrix}, \quad u_S = \frac{i}{2}\begin{pmatrix} \mp i \\ \gamma \end{pmatrix}, \quad (11)$$

with $\gamma = (p_x + ip_y)/p$. The full surface state then has the form

$$\psi_{B,S}(\mathbf{r}) = u_{B,S} e^{i(p_x x + p_y y)}(e^{\lambda_1 z} - e^{\lambda_2 z}). \quad (12)$$

For a particle-hole symmetric Hamiltonian Eq. (2) an electron in a surface state is found with equal probability on either Bi or Se sublattice. In order to clarify the spin content of the surface states, we first calculate the expectation value of the in-plane components of a σ operator on each sublattice. For the in-plane components one finds

$$\langle\sigma_i\rangle_B = u_B^\dagger \sigma_i u_B = \pm\frac{\varepsilon_{ijz} p_j}{2p},$$
$$\langle\sigma_i\rangle_S = u_S^\dagger \sigma_i u_S = -\langle\sigma_i\rangle_B, \quad (13)$$

where $i, j = x, y$, whereas

$$\langle\sigma_z\rangle_B = \langle\sigma_z\rangle_S = 0. \quad (14)$$

As in Eq. (11), the upper/lower sign corresponds to the energies above/below the Dirac crossing. We see that for

both sublattices the expectation value of σ is an in-surface-plane vector perpendicular to the momentum p. However the direction of $\langle\sigma\rangle$ is opposite for two sublattices. As we see from Eqs. (3) and (4), the direction of $\langle\sigma\rangle_B$ for the Bi sublattice always coincides with the direction of the true spin. Thus the spin of this component is always described by the surface Hamiltonian Eq. (1). On the contrary, the spin of Se sublattice electron component of the wave function differs from $\langle\sigma_i\rangle_S$ by a 180-deg rotation around $\mathbf{n}$, the axis normal to the layer plane.

We conclude that the spin orientation of two sublattices coincide with each other and with the prediction of Eq. (1) if and only if the surface of the crystal coincides with the quintuple layer plane. For any other crystal surface the two sublattices spins differ, and the difference is typically large. In Fig. 1 we show the spin directions for two sublattices for a surface normal to the layer plane.

For certain applications, such as the description of spin and angle resolved photoemission (see next section), it is preferable to write the Hamiltonian (2) in a form in which the electron spin operator $\mathbf{s}$ is directly proportional to σ. (Photoelectrons do not carry a sublattice index, so that a spin operator that contains τ is problematic in that context.) Here, following Ref. 13, we perform the unitary transformation

$$\tilde{H} = UHU^*, \quad U = \frac{1 + \tau_z}{2} + i(\sigma \cdot \mathbf{n})\frac{1 - \tau_z}{2}, \quad (15)$$

where $\mathbf{n}$ is the vector normal to the quintuple layer. The Hamiltonian $\tilde{H}$ no longer has the manifest rotational symmetry of Eq. (2), but the relation between spin and Pauli matrices now takes the standard form

$$\mathbf{s} = \sigma/2. \quad (16)$$

After the unitary transformation the two spinor amplitudes take the form

$$\tilde{u}_B = u_B, \quad \tilde{u}_S = i(\sigma \cdot \mathbf{n})u_S, \quad (17)$$

where u_B and u_S are given in Eq. (11).

III. PHOTOEMISSION

Spin and angle resolved photoemission spectroscopy (spin-resolved ARPES)[23] has been the tool of choice for the experimental investigation of the helical nature of the surface states in three-dimensional topological insulators.[15,16] What does the coexistence of two different spin polarizations of the surface electron imply for the polarization of photoelectrons?

In general, the probability P of electron photoemission is proportional to the squared matrix element of the interaction H_{int} with the photon field,

$$P \propto |\langle f|H_{\text{int}}|i\rangle|^2, \quad (18)$$

where $|f\rangle$ and $|i\rangle$ are the spinor states corresponding to the free final and the initial surface electron states, respectively. [One has $H_{\text{int}} = (e/c)\mathbf{A} \cdot \hat{\mathbf{j}}$, where $\mathbf{A}$ is the vector potential of the photon field and $\hat{\mathbf{j}}$ is the nonrelativistic current density operator.] In our case, the initial state is described by separate amplitudes $\tilde{u}_B$ and $\tilde{u}_S$ for electrons from the Bi and Se

sublattices, which are given in Eqs. (11) and (17) above. Since the interaction H_{int} is spin conserving and local, the spinor structure u_f of the photoelectron becomes a linear combination

$$u_f = \alpha \tilde{u}_B + \beta \tilde{u}_S, \tag{19}$$

where the coefficients α and β contain contributions from matrix elements of H_{int} on the Bi and Se atoms in the crystal, respectively. Their precise value depends on the details of the atomic structure and the frequency of the incident light.

Since the surface Hamiltonian Eq. (1) predicts a spin polarization parallel to the surface, direct observation of an out-of-plane spin would be of most interest. A simple calculation gives for the z component

$$u_f^+ \sigma_z u_f = \frac{(n_z |\beta|^2 + \text{Re}\,\alpha\beta^*)(n_x p_y - n_y p_x)}{p}. \tag{20}$$

We see that in case of arbitrary surface orientation the electron emitted from the Se sublattice acquires an out-of-plane polarization even without interference. The nature of this out-of-plane polarization is a direct consequence of the special form of the spin operators in Eq. (4): The surface electrons described by the symmetric Hamiltonian Eq. (2) for each sublattice have the expectation value of the σ operator parallel to the surface. However, since the spin of one sublattice is obtained from $\sigma/2$ by a π rotation around the axis normal to the quintuple layer $\mathbf{n}$, the spins of this sublattice all lie in a plane different from the crystal surface plane. (Similarly for the asymmetric Hamiltonian of the following section, for the general surface orientation, spins of both sublattices at different values of momentum lie in two planes, different from each other and from the surface plane.)

As we saw in Sec. II, only for the most studied (111) surface (the surface parallel to the quintuple layer), polarizations of the electron on both sublattices coincide. Consequently Eq. (20) predicts no out-of-plane polarization here.

Most interesting is the photoemission from the surface normal to the quintuple layers ($n_z = 0$), where the out-of-plane polarization appears due to the interference of contributions from the two sublattices only, $u_f^+ \sigma_z u_f \sim \text{Re}\,\alpha\beta^*$ (20). The unit cell of Bi_2Se_3 consists of five atoms, each hosting a p_z electron. The Hamiltonian Eq. (2) operates in a reduced 2×2 pseudospin subspace built from an even (with respect to reflection through the middle plane of the quintuple layer) state from the conduction band and an odd state from the valence band.[12] In our case of a surface normal to the quintuple layer, in the odd state electron waves emitted from the different atoms would cancel each other, unless there is a finite phase difference due to the momentum of outgoing electron. Thus we expect the photoionization amplitude from the odd state to be $\beta \sim (\mathbf{p} \cdot \mathbf{n})$. Consequently the out-of-plane spin component Eq. (20) should have a node and change sign at $(\mathbf{p} \cdot \mathbf{n}) = 0$.

IV. ANISOTROPIC TOPOLOGICAL INSULATORS

In the simplified model of Sec. II, the origin of the different spin content of the surface states at different facets was the special form of the spin operator Eq. (4). Whereas the special symmetries of the Hamiltonian Eq. (2), rotational and particle-hole symmetry are broken in real crystal, the

mechanism leading to the different spin content of surface states at different facets continues to operate, as we now show.

The general anisotropic and particle-hole asymmetric generalization of the Hamiltonian (2) is[24]

$$H = \epsilon(\mathbf{K}) + \tau_z \mathcal{M} + \tau_x \sum_{x,y,z} A_i \sigma_i \mathbf{K}_i, \tag{21}$$

where now

$$\mathcal{M}(\mathbf{K}) = M - \sum \mathbf{K}_i B_{ij} \mathbf{K}_j \tag{22}$$

and

$$\epsilon(\mathbf{K}) = \sum \mathbf{K}_i D_{ij} \mathbf{K}_j, \tag{23}$$

with a positive-definite symmetric matrix B_{ij} and a symmetric matrix D_{ij}. The first term $\epsilon(\mathbf{K})$ in Eq. (21) is explicitly particle-hole (valence-band–conduction-band) asymmetric. Choosing the z axis to be orthogonal to the plane of the quintuple layer of the Bi_2Se_3 crystal, the matrices B_{ij} and D_{ij} are found to be diagonal[12]

$$B_{ij} = B_i \delta_{ij}, \qquad D_{ij} = D_i \delta_{ij}. \tag{24}$$

The microscopic calculations of Ref. 12 give the following values for their elements:

$$\begin{aligned} A_z &= A_1 = 2.2\,\text{eV\,Å}, \\ B_z &= B_1 = 10\,\text{eV\,Å}^2, \\ D_z &= D_1 = 1.3\,\text{eV\,Å}^2 \end{aligned} \tag{25}$$

for the direction normal to the layer,

$$\begin{aligned} A_x &= A_y = A_2 = 4.1\,\text{eV\,Å}, \\ B_x &= B_y = B_2 = 56.6\,\text{eV\,Å}^2, \\ D_x &= D_y = D_2 = 19.6\,\text{eV\,Å}^2 \end{aligned} \tag{26}$$

for the two axes in the layer plane, and

$$M = 0.28\,\text{eV}. \tag{27}$$

Finding the surface states for the general Hamiltonian Eq. (21) become a complicated algebraic problem. Here we discuss the solutions only in case that B_{ij} and D_{ij} remain diagonal after a rotation that brings the surface normal to the z axis. This includes the case of a surface parallel to the quintuple layer and a surface perpendicular to the layers in Bi_2Se_3, as in the nanoribbons of Ref. 7.

As in Sec. II, we search for eigenfunctions of the Hamiltonian (21) of the form (we use again $\mathbf{K}$ for the three-dimensional momentum and $\mathbf{p}$ for the two-dimensional surface states momentum)

$$\psi_B = u_B e^{i(p_x x + p_y y) + \lambda z}, \qquad \psi_S = u_S e^{i(p_x x + p_y y) + \lambda z}, \tag{28}$$

where $\text{Re}\,\lambda > 0$ and we rotated the coordinate system, such that the topological insulator occupies the half space $z < 0$. With the particle-hole asymmetric term $\epsilon(K)$, simple guessing of the appropriate solution, like we did in Sec. II, does not work and one has to pursue an explicit derivation of the result. To this end it is convenient to replace the sublattice spinors $u_S = (u_{B_\uparrow}, u_{B_\downarrow})$ and $u_S = (u_{S_\uparrow}, u_{S_\downarrow})$ by spin-up and spin-down

pseudospinors

$$u = (u_{B_\uparrow}, u_{S_\uparrow}), \quad v = (u_{B_\downarrow}, u_{S_\downarrow}). \qquad (29)$$

These pseudospinor amplitudes now satisfy the system of equations

$$(\varepsilon - \epsilon - \mathcal{M}\tau_z + i\lambda A_z \tau_x)u = (A_x p_x - i A_y p_y)\tau_x v, \qquad (30)$$

$$(\varepsilon - \epsilon - \mathcal{M}\tau_z - i\lambda A_z \tau_x)v = (A_x p_x + i A_y p_y)\tau_x u, \qquad (31)$$

where

$$\mathcal{M} = \mathcal{M}(p,\lambda) = M - B_x p_x^2 - B_y p_y^2 + B_z \lambda^2,$$
$$\epsilon = \epsilon(p,\lambda) = + D_x p_x^2 + D_y p_y^2 - D_z \lambda^2. \qquad (32)$$

From the first equation we find

$$v = \Omega u, \qquad (33)$$

with

$$\Omega = \frac{1}{A_x p_x - i A_y p_y}[(\varepsilon - \epsilon)\tau_x + i\mathcal{M}\tau_y - i\lambda A_z]. \qquad (34)$$

Substituting this expression for v into Eq. (31) one finds

$$[\varepsilon - \epsilon(p,\lambda)]^2 = \mathcal{M}(p,\lambda)^2 + A_x^2 p_x^2 + A_y^2 p_y^2 - A_z^2 \lambda^2. \qquad (35)$$

This equation is biquadratic in λ and allows us to find two values of the squared inverse decay length $\lambda_{1,2}^2$ for each energy ε, and consequently two decaying solutions with Re $\lambda_{1,2} > 0$.

The open boundary conditions at the surface can be satisfied if there exists a choice of pseudospinors corresponding to two solutions λ_1 and λ_2 such that

$$u_1 = u_2, \quad v_1 = v_2. \qquad (36)$$

The first equality here is easily satisfied, because Eq. (35) imposes no restrictions on the pseudospinor u. The second equality then involves the $\Omega = \Omega(\varepsilon,\lambda)$ of Eq. (34), which depends on energy and on the specific root of the biquadratic equation (35),

$$0 = v_1 - v_2 = [\Omega(\varepsilon,\lambda_1) - \Omega(\varepsilon,\lambda_2)]u. \qquad (37)$$

This equation has a solution only if the difference matrix $\Omega(\varepsilon,\lambda_1) - \Omega(\varepsilon,\lambda_2)$ has a zero eigenvalue, and consequently $\det[\Omega(\varepsilon,\lambda_1) - \Omega(\varepsilon,\lambda_2)] = 0$. Thus we find the condition

$$\det[(\epsilon_1 - \epsilon_2)\tau_x + i(\mathcal{M}_1 - \mathcal{M}_2)\tau_2 - i A_z(\lambda_1 - \lambda_2)] = 0, \qquad (38)$$

where λ_1 and λ_2 are the two roots of Eq. (35). This immediately gives a simple formula

$$A_z^2 = (B_z^2 - D_z^2)(\lambda_1 + \lambda_2)^2. \qquad (39)$$

After straightforward calculation we can now find the energy of the surface states

$$\varepsilon(p) = \frac{D_z}{B_z}M \pm \sqrt{A_x^2 p_x^2 + A_y^2 p_y^2}\sqrt{1 - \frac{D_z^2}{B_z^2}}$$
$$+ \frac{D_x B_z - D_z B_x}{B_z}p_x^2 + \frac{D_y B_z - D_z B_y}{B_z}p_y^2, \qquad (40)$$

and the two spinor amplitudes (here we use again the more meaningful spinor functions for two sublattices instead of the

pseudospinors u and v)

$$u_B = \frac{1}{2}\sqrt{1 + \frac{D_z}{B_z}}\begin{pmatrix} \pm i \\ \gamma_A \end{pmatrix},$$
$$u_S = \frac{i}{2}\sqrt{1 - \frac{D_z}{B_z}}\begin{pmatrix} \mp i \\ \gamma_A \end{pmatrix}, \qquad (41)$$

where now

$$\gamma_A = \sqrt{(A_x p_x + i A_y p_y)/(A_x p_x - i A_y p_y)}. \qquad (42)$$

The main differences between the solution of the simplified Hamiltonian of Sec. II and the solution given in Eqs. (40) and (41) are:

(1) The electron no longer spends equal time on the two sublattices. The difference in probabilities is governed by the ratio D_z/B_z, which can actually be small. (Reference 12 finds $D_1/B_1 \approx 0.13$.)

(2) For the anisotropic Hamiltonian the expectation value of the σ operator, which we discussed in Sec. II, is no longer perpendicular to the momentum. However, the expectation values $\langle \sigma \rangle_{B,S}$ on the Bi and Se sublattices are still perpendicular to the rescaled momentum $p_x \to A_x p_x$, $p_y \to A_y p_y$, and Fig. 1 remains valid if the rescaled momentum is used. We observe that such a rescaling can also be used to account for the difference between the surface-state spin on the Bi and Se sublattices that is found here and in Sec. II: Rescaling with one inverted sign $p_x \to -A_x p_x$, $p_y \to A_y p_y$ is sufficient to permute the two panels in Fig. 1.

(3) The third important feature of the solution Eq. (40) is also caused by the particle-hole asymmetric term $\epsilon_0(\mathbf{K})$ in Eq. (21). As long as one considers the bulk states, this term leads to a trivial bending of both conduction and valence bands. However, the inclusion of the energy $\epsilon_0(\mathbf{K})$ into the Hamiltonian (21) leads to a nonzero value of the Dirac crossing energy, as can be seen from the explicit solution (40). In the most general case of matrices D_{ij} and B_{ij} having no vanishing elements we found the position of the Dirac crossing, which is the (unique) energy eigenvalue of a surface state with vanishing momentum parallel to the surface, to be simply $\varepsilon = \varepsilon_{\text{Dirac}} = M D_{zz}/B_{zz}$. Thus, returning to the original coordinate system of Eqs. (21)–(27), we may introduce a vector $\mathbf{e}$ normal to the surface at a particular point and write a formula for the energy of the Dirac crossing valid for the entire crystal

$$\varepsilon_{\text{Dirac}} = M \left(\sum \mathbf{e}_i D_{ij} \mathbf{e}_j \right) \Big/ \left(\sum \mathbf{e}_i B_{ij} \mathbf{e}_j \right). \qquad (43)$$

Substituting the actual values of the matrices D_{ij} and B_{ij}, we find that for Bi_2Se_3 the position of the Dirac crossing may vary by as much as

$$\Delta\varepsilon_{\text{Dirac}} = \left(\frac{D_2}{B_2} - \frac{D_1}{B_1}\right) M \approx 0.06\,\text{eV} \approx \frac{A_2}{6.8\,\text{nm}}, \qquad (44)$$

upon varying the orientation of the surface.

V. ELECTROSTATICS OF A TOPOLOGICAL INSULATOR EDGE

Being built of neutral atoms, the topological insulator crystal is obviously a charge-neutral system. Because of the
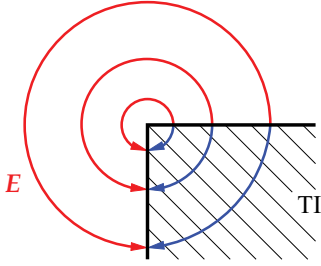
FIG. 2. (Color online) Electric field near the edge of an anisotropic TI.

bulk band gap, the bulk charge density is insensitive to the precise choice of the Fermi energy. However, the requirement of neutrality should also include the surface charges. Even in the absence of particle-hole symmetry, the structure of the surface-state spectrum is such that generically charge neutrality is achieved if the Fermi energy equals the energy of the Dirac crossing $\varepsilon_{\text{Dirac}}$. Since the latter energy depends on the orientation of the surface, see Eqs. (43) and (44), surface-charge neutrality cannot be satisfied by a uniform choice of the Fermi energy for the entire crystal. [The variation of the surface charge density for a hypothetical uniform choice of the Fermi energy is estimated as $\Delta n \sim (\Delta \varepsilon_{\text{Dirac}}/A_2)^2/4\pi \approx 1.7 \times 10^{11}$ cm^{-2}.] Such a choice of the Fermi energy would lead to a strong electric field along the surface, which should not happen for a metallic surface. Instead, the surface charge is redistributed near the edges between different facets, so that the electric field parallel to the metallic surface vanishes. Far away from the edge, surface neutrality will then be achieved simultaneously for both surfaces, while the electrostatic potential difference between two facets will compensate the difference of the Dirac crossing energies $\Delta \varepsilon_{\text{Dirac}} = \Delta V$. In the limit that the surfaces can be approximated as perfect metals, the solution of such electrostatic problem is straightforward.[25] For two surfaces at a 90 deg angle, as shown in Fig. 2, the potential Φ_{in} and Φ_{out} inside (polar angle $-\pi/2 < \phi < 0$) or outside (polar angle $0 < \phi < 3\pi/2$) the topological insulator is

$$\Phi_{\text{out}}(r,\phi) = \frac{2\Delta V}{3\pi e}\phi, \quad \Phi_{\text{in}}(r,\phi) = -\frac{2\Delta V}{\pi e}\phi, \quad (45)$$

corresponding to a surface charge density

$$|\sigma(r)| = \left(\kappa + \frac{1}{3}\right)\frac{\Delta V}{2\pi^2 e r} \quad (46)$$

at a distance r away from the edge. Substituting the dielectric constant of Bi$_2$Se$_3$, $\kappa \approx 110$, and taking $\Delta V = 0.06$ eV from Eq. (44), we arrive at the estimate $|\sigma(r)| \sim (e/r) \times 2.3 \times 10^6$ cm^{-1}. Doping of the surface electron gas by electrostatic external gates would lead to an additional edge/corner charge accumulation, similar to the edge charge accumulation in graphene ribbons.[26]

The electrostatic calculation of Eqs. (45) and (46) ignores completely the kinetic energy of the surface electrons. However, due to the large dielectric constant in Bi$_2$Se$_3$ there exists a regime where the semiclassical electron dynamics is already described by a smooth coordinate dependent Fermi energy $E_F(r) = Ak_F(r)$, but the electrostatic energy is not yet dominant, $e\Phi \sim E_F$ (Thomas Fermi approximation). In

this case one requires the sum $E_F(r) + e\Phi$ to be a constant over the metallic surface, while electrostatics is recovered in the large-sample limit. The range of validity of Eq. (45) is now found as $\Delta V > A_2 k_F(r) = A_2 2\sqrt{\pi\sigma(r)/e}$, leading to $r > 135$ nm. For much smaller samples one should ignore the Coulomb interaction and describe the surface electrons by the Dirac equation with shifted crossing energy [Eqs. (43) and (44)].

VI. DISCUSSION

In this article we have analyzed the spin structure of surface states in the three-dimensional topological insulator Bi$_2$Se$_3$, as it follows from the low-energy continuum description proposed in Ref. 12. The inequivalence of the spin variables on the Bi and Se sublattices in this continuum description leads to a nontrivial spin structure of the surface states at surfaces other than the (111) surface. We also found that the energy of the Dirac crossing in the surface-state dispersion depends on the surface orientation. Although the precise boundary conditions of the effective low-energy description—open boundary conditions, in which all components of the spinor wave function vanish at the surface of the topological insulator, as in Ref. 12—are key to our quantitative analysis, we expect that the effects we predict persist if the boundary conditions are changed. Different boundary conditions have appeared in the literature for the (111) surface,[18,21,22] but not for other surfaces of a Bi$_2$Se$_3$ crystal.

The theoretical findings of this paper may be applied to the interpretation of several experiments involving the surface states of three-dimensional topological insulators. Here we mention the STM measurements of the surface electrons charge near an artificial step on Bi$_2$Te$_3$ surface[27,28] and the measurement of Aharonov-Bohm oscillations in the surface-state-mediated transport in Bi$_2$Se$_3$ nanoribbons.[7] In the latter case a theoretical analysis predicts that the magnetoconductance should depend strongly on the position of the Dirac crossing for the electrons on the surface of the nanoribbon.[8] Obviously a nonuniformity of this Dirac crossing of the type discussed above will affect the minimal conductivity and interference pattern for the thick ($\sim$100 nm) rectangular shaped nanoribbons of Ref. 7 and should be taken into account in a quantitative modeling of the device.

Our most remarkable result is the prediction of the possibility of out-of-plane momentum polarizations of photoemitted electrons [Eq. (20)] depending on the exact orientation of the surface. Measuring these spin components would confirm the validity of the microscopic Hamiltonian of Ref. 12 and its boundary conditions. It will also explicitly demonstrate the signatures of interference of photoemission contributions from two sublattices of Bi$_2$Se$_3$. Such experiments require, however, the preparation of Bi$_2$Se$_3$ crystals with sufficient quality surfaces other than in the (111) direction.

Although the low-energy model of Ref. 12, which we used as the basis for our calculations, predicts a nontrivial spin structure for surfaces other than the (111) surface, there may be other mechanisms that lead to an effectively reduced spin that are not included in this model. In this context, we mention recent first-principles calculations of Yazyev and co-workers, who find that the net spin polarization of surface states on the

(111) surface is reduced by an amount of order 50%,[29] which they attribute to the effect of strong spin-orbit interaction on the heavy Bi atoms. Since the magnitude of the spin of a photoemitted electron has to be $s = 1/2$ (a free electron always has its spin pointed in some direction), the reduced net spin found in Ref. [29] implies a likely out-of-plane component of photoemitted electrons, even for the (111) surface. This effect may be suppressed since it involves the photoionization from the inner row of Bi atoms, less accessible for ARPES. On the other hand, as the spin-orbit interaction admixes $p_x \pm i p_y$ electron states to p_z ones, one may play with the photon polarization to selectively enhance the electron's ionization from Bi. Further investigation in this direction is obviously desired.

Note added: Upon completing this version of the article, we learned of Ref. [30], which also addresses the spin structure and Dirac crossing of the surface states.

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

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

[3] B. A. Volkov and O. A. Pankratov, JETP Lett. **42**, 178 (1985).

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

[5] Proposals employing a net spin polarization of the surface states governed by Eq. (1) can be found in S. Raghu, S. B. Chung, X.-L. Qi, and S.-C. Zhang, Phys. Rev. Lett. **104**, 116401 (2010); I. Appelbaum, H. D. Drew, and M. S. Fuhrer, Appl. Phys. Lett. **98**, 023103 (2011).

[6] This differs from the textbook definition of the helical states as states with spin parallel to momentum.

[7] H. Peng, K. Lai, D. Kong, S. Meister, Y. Chen, X.-L. Qi, S.-C. Zhang, Z.-X. Shen, and Y. Cui, Nat. Mater. **9**, 225 (2010).

[8] J. H. Bardarson, P. W. Brouwer, and J. E. Moore, Phys. Rev. Lett. **105**, 156803 (2010).

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

[10] M. Lasia and L. Brey, Phys. Rev. B **86**, 045317 (2012).

[11] Y. Xia, D. Qian, D. Hsieh1, 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).

[12] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Nat. Phys. **5**, 438 (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] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **452**, 970 (2008).

[15] D. Hsieh, Y. Xia, D. Qian, L. Wray, J. H. Dil, F. Meier, J. Osterwalder, L. Patthey, J. G. Checkelsky, N. P. Ong, A. V. Fedorov, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **460**, 1101 (2009).

[16] 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).

[17] M. R. Scholz, D. Marchenko, A. Varykhalov, A. Volykhov, L. V. Yashina, and O. Rader, arXiv:1108.1053v1.

[18] L. Fu and E. Berg, Phys. Rev. Lett. **105**, 097001 (2010).

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

[20] W.-Y. Shan, H.-Z. Lu, and S.-Q. Shen, New J. Phys. **12**, 043048 (2010).

[21] T. H. Hsieh and L. Fu, Phys. Rev. Lett. **108**, 107005 (2012).

[22] S. S. Pershoguba and V. M. Yakovenko, arXiv:1202.5526v1.

[23] For a review see J. H. Dil, J. Phys.: Condens. Matt. **21**, 403001 (2009).

[24] To consider the most important surface-orientation-dependent features of the surface states, we are keeping only the linear and quadratic in momentum terms in the effective Hamiltonian. Thus we disregard more delicate effects seen already for the (111) surface, like the Fermi surface wrapping,[4] or the polarization of electrons on the same sublattice variations within the first quintuple layer, seen, e.g., in S. V. Eremeev *et al.*, Nat. Commun. **3**, 635 (2012).

[25] L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamics of Continuous Media* (Butterworth-Heinemann, London, 1984), Vol. 8.

[26] P. G. Silvestrov and K. B. Efetov, Phys. Rev. B **77**, 155436 (2008).

[27] Z. Alpichshev, J. G. Analytis, J.-H. Chu, I. R. Fisher, and A. Kapitulnik, Phys. Rev. B **84**, 041104(R) (2011).

[28] 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).

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

[30] F. Zhang, C. L. Kane, and E. J. Mele, arXiv:1203.6382v1.