PHYSICAL REVIEW B **82**, 045122 (2010)

Model Hamiltonian for topological insulators

Chao-Xing Liu,[1] Xiao-Liang Qi,[2] HaiJun Zhang,[3] Xi Dai,[3] Zhong Fang,[3] and Shou-Cheng Zhang[2]
[1]*Physikalisches Institut (EP3) and Institute for Theoretical Physics and Astrophysics, University of Würzburg, 97074 Würzburg, Germany*
[2]*Department of Physics, McCullough Building, Stanford University, Stanford, California 94305-4045, USA*
[3]*Beijing National Laboratory for Condensed Matter Physics, and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*

(Received 11 May 2010; published 26 July 2010)

In this paper we give the full microscopic derivation of the model Hamiltonian for the three-dimensional topological insulators in the Bi_2Se_3 family of materials (Bi_2Se_3, Bi_2Te_3 and Sb_2Te_3). We first give a physical picture to understand the electronic structure by analyzing atomic orbitals and applying symmetry principles. Subsequently, we give the full microscopic derivation of the model Hamiltonian introduced by Zhang *et al.* [Nat. Phys. **5**, 438 (2009)] based both on symmetry principles and the $\mathbf{k}\cdot\mathbf{p}$ perturbation theory. Two different types of k^3 terms, which break the in-plane full rotation symmetry down to threefold rotation symmetry, are taken into account. An effective Hamiltonian is derived for the topological surface states. Both bulk and surface models are investigated in the presence of an external magnetic field, and the associated Landau level structure is presented. For a more quantitative fitting to the first principle calculations, we also present a model Hamiltonian including eight energy bands.

DOI: 10.1103/PhysRevB.82.045122 PACS number(s): 71.15.−m, 71.18.+y, 73.20.−r, 73.61.Le

I. INTRODUCTION

Recently, topological insulators (TIs) have been investigated intensively both theoretically and experimentally.[2–4] These insulators are fully gapped in the bulk but have gapless edge or surface states which are topologically protected by the time reversal symmetry. The topological insulator was first theoretically predicted[5] and experimentally observed[6] in HgTe quantum wells. Transport measurements[6,7] show the existence of gapless edge channels, which demonstrates that the HgTe/CdTe quantum well is a two-dimensional (2D) TI with a quantum spin Hall effect. Later, Bi_xSb_{1-x} was suggested to be a three-dimensional (3D) TI (Ref. 8) with topologically nontrivial surface states, which were observed by angle-resolved photoemission spectroscopy (ARPES).[9] However, Bi_xSb_{1-x} has a small energy gap, is subjected to alloy disorder and possesses rather complicated surface states. More recently, TIs with large bulk gaps of ~0.3 eV and single Dirac cone surface states have been theoretically predicted for Bi_2Te_3,[1] Sb_2Te_3[1] and Bi_2Se_3.[1,10] ARPES measurements[10,11] indeed show a single Dirac cone with linear dispersion around the Γ point in both Bi_2Se_3 and Bi_2Te_3. Current research on these materials is developing rapidly.[12–30]

For a deeper understanding and quantitative predictions of the phenomena associated with the TIs, it is highly desirable to construct standard models for both 2D and 3D TIs. Bernevig, Hughes and Zhang (BHZ) (Ref. 5) constructed a model Hamiltonian for the 2D TI in HgTe quantum wells. This model Hamiltonian demonstrates the basic mechanism of TI behavior through band inversion induced by spin-orbit coupling (SOC). It has been applied successfully for quantitative predictions of the helical edge states and properties under magnetic fields.[31] Zhang *et al.*[1] derived a model Hamiltonian for the 3D TI Bi_2Se_3, Bi_2Te_3, and Sb_2Te_3 and obtained topological surface states consisting of a single Dirac cone. Interestingly, in the thin film limit, the 3D TI model reduces exactly to the 2D TI model by BHZ.[32–34] In this paper, we give the full microscopic derivation of our model Hamiltonian, first by constraining its form by symmetry principles and a careful analysis of the relevant atomic orbitals. Subsequently, we determine the parameters of our model Hamiltonian by a systematic $\mathbf{k}\cdot\mathbf{p}$ expansion near the Γ point and comparison with the *ab initio* calculations.[1] Furthermore the higher order k^3 terms neglected in Ref. 1 are also included in the derivation in order to recover the crystal C_3 rotation symmetry.[35] Compared to the symmetry arguments given in Ref. 1, the derivation given in this paper determines all the parameters of our model Hamiltonian by the wave functions from *ab initio* calculations so that no fitting is required and no ambiguity is introduced. As an application of our model Hamiltonian, we study the bulk and surface Landau level spectra in a magnetic field. The surface Landau levels have a $\sqrt{B}$ field dependence, as is expected from the Dirac-type dispersion of the surface states. The gap between zeroth and first Landau levels can be as large as 50 meV for a 10 T magnetic field, which suggests that the topological magnetoelectric effect[36,37] can be observable at such energy scales. Furthermore, we propose a more quantitative description of the Bi_2Se_3 family of TIs by going beyond four bands and present a model Hamiltonian with eight bands. Recently, our model Hamiltonian has been applied successfully to understand a number of experiments, including the scanning tunneling microscopy (STM) study of the topological surface states,[14,20] STM study of the surface bound states,[30] STM study of the quasiparticle interference,[14,20,38] the crossover from 3D to 2D topological insulators,[21,32–34] and the Landau levels of the topological surface states.[22,23]

The paper is organized as follows. In Sec. II we first present the lattice structure and the symmetry properties of the Bi_2Se_3 crystal. Then we turn to the electronic band structure and discuss the atomic orbital picture, which is helpful to capture the essential physics in the long wave length limit.

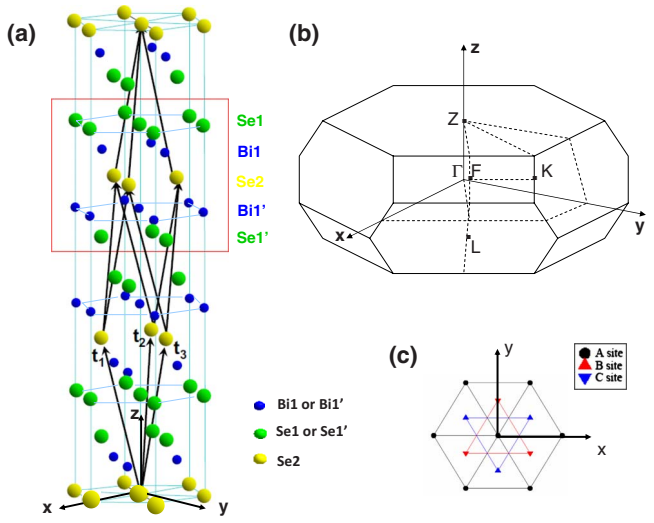

FIG. 1. (Color online) (a) The crystal structure of Bi_2Se_3. $\vec{t}_{1,2,3}$ is the primitive lattice vector, given by $\vec{t}_1=(\sqrt{3}a/3,0,c/3)$, $\vec{t}_2=(-\sqrt{3}a/6,a/2,c/3)$, and $\vec{t}_3=(-\sqrt{3}a/6,-a/2,c/3)$, where a is the lattice constant in the x-y plane and c is the lattice constant along the z direction. The quintuple layer is shown in the red box with Se1-Bi1-Se2-Bi1′-Se1′. Se1 Bi1 and Se1′ (Bi1′) are equivalent. (b) Brillioun zone of Bi_2Se_3. (c) The in-plane triangle lattice has three possible positions A, B and C.

Keeping this atomic orbital picture in mind, we investigate in detail the properties of the bands near the Fermi surface based on the symmetry argument in Sec. III. Furthermore, our model Hamiltonian for the conduction and valence bands is derived from the theory of invariants. In Sec. IV, we rederive our model Hamiltonian from the $\mathbf{k}\cdot\mathbf{p}$ theory and determine its parameters by more fundamental matrix elements of the momentum operator in the $\mathbf{k}\cdot\mathbf{p}$ theory. As an application of our model Hamiltonian, the surface state Hamiltonian and the Landau levels for both bulk and surface states are calculated in Sec. V and Sec. VI, respectively. In Sec. VII, the quantitative limitation of our model Hamiltonian with four-bands is discussed and a model Hamiltonian is proposed to describe the Bi_2Se_3 type of materials more quantitatively. In Sec. VIII we provide a brief discussion and conclusion.

II. CRYSTAL STRUCTURE, ATOMIC ORBITALS, AND SYMMETRY

In this section we will describe the crystal structure of the Bi_2Se_3 family of materials and discuss the relevant atomic orbitals and the discrete symmetries. A large portion of the content of this section is already discussed in Ref. 1, but we feel that it is helpful to present the more complete version of this discussion here to make this paper self-contained. The crystal structure of Bi_2Se_3 is rhombohedral with the space group D_{3d}^5 ($R\bar{3}m$). As shown in Fig. 1(a), the crystal has a layered structure stacked along z direction with five atoms (two Bi atoms and three Se atoms) in one unit cell, including two equivalent Se atoms (Se1 and Se1′), two equivalent Bi atoms (Bi1 and Bi1′), and one Se atom (Se2) which is inequivalent to the Se1 and Se1′ atoms. Therefore five atomic layers can be viewed as one unit, which is usually called a quintuple layer. Each atomic layer forms a triangle lattice, which has three possible positions, denoted as A, B and C, as shown in Fig. 1(c). Along the z direction, the triangle layers are stacked in the order A-B-C-A-B-C-$\cdots$. We note that the primitive lattice vector t_i ($i=1,2,3$) is not directed along the z direction. For example, in one quintuple layer, the Se2 atoms occupy the A sites; in the next quintuple layer, the Se2 atoms do not occupy the A sites but rather the C or B sites. Our coordinate is set as the following: the origin is set at the Se2 site; the z direction is set perpendicular to the atomic layer, the x direction is taken along the binary axis with the twofold rotation symmetry, and the y direction is taken along the bisectrix axis, which is the crossing line of the reflection plane and the Se2 atomic layer plane. The Brillouin zone (BZ) of this lattice structure is shown in Fig. 1(b). This crystal structure has the following discrete symmetries:

(1) *Threefold rotation R_3 along the z direction*. R_3 can be generated by the following transformation: $x\rightarrow x\cos\theta - y\sin\theta$, $y\rightarrow x\sin\theta+y\cos\theta$, and $z\rightarrow z$, where $\theta=\frac{2\pi}{3}$.

(2) *Twofold rotation R_2 along the x direction*. R_2 corresponds to the following transformation: Se2$\rightarrow$Se2, Bi1$\rightarrow$Bi1′, and Se1$\rightarrow$Se1′; $z\rightarrow -z$, $x\rightarrow x$, and $y\rightarrow -y$. For this symmetry operation, we find that Bi1 (Se1) and Bi1′ (Se1′) layers interchange their positions.

(3) *Inversion P*. P: Se2$\rightarrow$Se2, Bi1$\rightarrow$Bi1′, and Se1$\rightarrow$Se1′; $z\rightarrow -z$, $x\rightarrow -x$, and $y\rightarrow -y$. The Se2 site is the inversion center of this lattice structure; hence we set Se2 as the origin point. Under inversion operation, Bi1 (Se1) is changed to Bi1′ (Se1′).

(4) *Time reversal T*. Time reversal operation is given by $T=\Theta K$, where $\Theta=i\sigma_2$ and K is the complex conjugate operator. Here $\sigma_{1,2,3}$ are the Pauli matrice for spin.

In order to get a physical picture of the band structure of Bi_2Se_3, we start from the atomic orbitals of Bi and Se. The electron configuration of Bi is $6s^26p^3$ and that of Se is $4s^24p^4$. The outmost shells for both Bi and Se are p orbitals; therefore it is natural to consider only the p orbitals of Bi and Se and neglect other orbitals. As discussed above, Bi_2Se_3 has a layered structure. The chemical bonding is very strong within one quintuple layer but the two neighboring quintuple layers are only coupled by the van der Waals force. Therefore it is reasonable for us to first focus on one quintuple layer. Within one quintuple layer there are five atoms in one unit cell and each atom has three orbitals (p_x, p_y, and p_z); therefore in total there are 15 orbitals. The spin is neglected first and will be discussed later when we introduce SOC into the system. We denote these orbitals as $|\Lambda,\alpha\rangle$ with Λ=Bi1, Bi1′, Se1, Se2, and Se1′ and $\alpha=p_x$, p_y, and p_z. As shown in Fig. 1(a), the Se2 atomic layer stays in the middle of the quintuple layer and is sandwiched by two Bi layers (Bi1 and Bi1′), while two Se layers (Se1 and Se1′) are located at the outermost. Since all the Se layers are separated by Bi layers, the strongest coupling in this system is the coupling between Bi layers and Se layers. Such coupling causes level repulsion, so that the Bi energy levels are pushed up and form new hybridized states $|B_\alpha\rangle$ and $|B'_\alpha\rangle$ while the Se energy levels are pushed down and yield three states $|S_\alpha\rangle$, $|S'_\alpha\rangle$, and $|S0_\alpha\rangle$, as shown in Fig. 2(I). Since the system has inversion symmetry, it is convenient to combine these orbitals to form

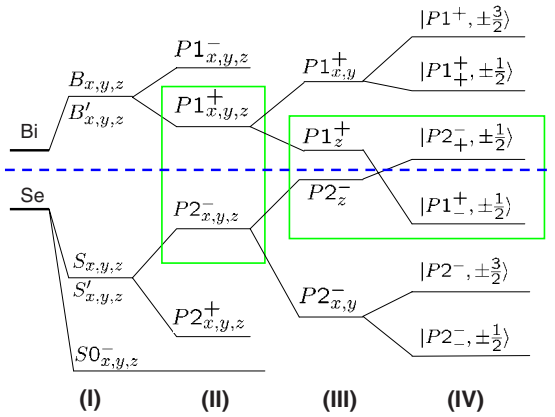

FIG. 2. (Color online) Schematic picture of the origin of the band structure of Bi_2Se_3. Starting from the atomic orbitals of Bi and Se, the following four steps are required to understand the band structure: (I) the hybridization of Bi orbitals and Se orbitals, (II) the formation of the bonding and antibonding states due to the inversion symmetry, (III) the crystal field splitting, and (IV) the influence of the SOC.

the bonding and anti-bonding states with the definite parity, which are given by

$$|P1^{\pm},\alpha\rangle = \frac{1}{\sqrt{2}}(|B_\alpha\rangle \mp |B'_\alpha\rangle),$$

$$|P2^{\pm},\alpha\rangle = \frac{1}{\sqrt{2}}(|S_\alpha\rangle \mp |S'_\alpha\rangle), \tag{1}$$

with the upper index denoting the parity and $\alpha = p_x, p_y, p_z$. When the coupling between $|B_\alpha(S_\alpha)\rangle$ and $|B'_\alpha(S'_\alpha)\rangle$ is taken into account, the bonding and antibonding states are split, with the antibonding state having higher energy than the bonding state. Therefore as shown in Fig. 2(II), the states $|P1^+,\alpha\rangle$ and $|P2^-,\alpha\rangle$ are found to be near the Fermi surface; hence we focus on $|P1^+,\alpha\rangle$ and $|P2^-,\alpha\rangle$ ($\alpha = p_x, p_y, p_z$) and neglect the other states. Furthermore the crystal has a layered structure, so the z direction is different from the x or y directions in the atomic plane. Thus there is an energy splitting between p_z and $p_{x,y}$ orbitals for both $P1^+$ and $P2^-$ states. We find that $|P1^+,p_{x,y}\rangle$ orbitals have higher energy than $|P1^+,p_z\rangle$, while $|P2^-,p_{x,y}\rangle$ orbitals have lower energy than $|P2^-,p_z\rangle$. Consequently, the conduction band mainly consists of $|P1^+,p_z\rangle$ while the valence band is dominated by the $|P2^-,p_z\rangle$ orbital before SOC is considered, as shown in Fig. 2(III).

Next we include SOC effect in the above atomic picture. The states $|P1^+,\alpha,\sigma\rangle$ and $|P2^-,\alpha,\sigma\rangle$ are all doubly degenerate, with one more index $\sigma = \uparrow,\downarrow$ to denote spin. The atomic SOC Hamiltonian is given by $\hat{H}_{so} = \lambda \mathbf{s}\cdot\mathbf{L}$ with $\lambda = \frac{1}{2m_0^2c^2}\frac{1}{r}\frac{\partial U}{\partial r}$ depending on the detailed potential U of atoms, which couples orbital angular momentum to spin. It is convenient to transform the p_x and p_y orbitals to $p_\pm$ with definite orbital angular momentum

$$|\Lambda,p_+,\sigma\rangle = -\frac{1}{\sqrt{2}}(|\Lambda,p_x,\sigma\rangle + i|\Lambda,p_y,\sigma\rangle), \tag{2}$$

$$|\Lambda,p_-,\sigma\rangle = \frac{1}{\sqrt{2}}(|\Lambda,p_x,\sigma\rangle - i|\Lambda,p_y,\sigma\rangle), \tag{3}$$

where $\Lambda = P1^+, P2^-$. Within this basis, the atomic SOC Hamiltonian is given by

$$\left\langle \Lambda,p_+,\uparrow \middle| H_{so} \middle| \Lambda,p_+,\uparrow \right\rangle = \left\langle \Lambda,p_-,\downarrow \middle| H_{so} \middle| \Lambda,p_-,\downarrow \right\rangle \equiv \frac{\lambda_\Lambda}{2},$$

$$\left\langle \Lambda,p_+,\downarrow \middle| H_{so} \middle| \Lambda,p_+,\downarrow \right\rangle = \left\langle \Lambda,p_-,\uparrow \middle| H_{so} \middle| \Lambda,p_-,\uparrow \right\rangle \equiv -\frac{\lambda_\Lambda}{2},$$

$$\left\langle \Lambda,p_+,\downarrow \middle| H_{so} \middle| \Lambda,p_z,\uparrow \right\rangle = \left\langle \Lambda,p_-,\uparrow \middle| H_{so} \middle| \Lambda,p_z,\downarrow \right\rangle \equiv \frac{\lambda_\Lambda}{\sqrt{2}},$$

$$\langle \Lambda,p_z,\uparrow(\downarrow)|H_{so}|\Lambda,p_z,\uparrow(\downarrow)\rangle = 0. \tag{4}$$

Here the value of λ_Λ is a linear combination of the SOC coefficient for Bi and Se, depending on how much the orbitals of Bi and Se are mixed into the state $|\Lambda\rangle$. The sign of λ_Λ is always positive for $\Lambda = P1^+, P2^-$ since the potential is always attractive for atoms. As we see, since the total angular momentum along the z direction is still conserved, hybridization only occurs between $|\Lambda,p_z,\uparrow\rangle(|\Lambda,p_z,\downarrow\rangle)$ and $|\Lambda,p_+,\downarrow\rangle(|\Lambda,p_-,\uparrow\rangle)$. After taking into account SOC, the new eigenstates are given by

$$\left|\Lambda,\frac{3}{2}\right\rangle = |\Lambda,p_+,\uparrow\rangle, \tag{5}$$

$$\left|\Lambda,-\frac{3}{2}\right\rangle = |\Lambda,p_-,\downarrow\rangle, \tag{6}$$

$$\left|\Lambda_+,\frac{1}{2}\right\rangle = u_+^\Lambda|\Lambda,p_z,\uparrow\rangle + v_+^\Lambda|\Lambda,p_+,\downarrow\rangle, \tag{7}$$

$$\left|\Lambda_-,\frac{1}{2}\right\rangle = u_-^\Lambda|\Lambda,p_z,\uparrow\rangle + v_-^\Lambda|\Lambda,p_+,\downarrow\rangle, \tag{8}$$

$$\left|\Lambda_+,-\frac{1}{2}\right\rangle = (u_+^\Lambda)^*|\Lambda,p_z,\downarrow\rangle + (v_+^\Lambda)^*|\Lambda,p_-,\uparrow\rangle, \tag{9}$$

$$\left|\Lambda_-,-\frac{1}{2}\right\rangle = (u_-^\Lambda)^*|\Lambda,p_z,\downarrow\rangle + (v_-^\Lambda)^*|\Lambda,p_-,\uparrow\rangle, \tag{10}$$

with the eigenenergies $E_{3/2}^\Lambda$ and $E_{1/2}^{\Lambda\pm}$ (each is doubly degenerate) and u and v obtained by solving the following 2×2 Hamiltonian,

$$\hat{H} = \begin{pmatrix} E_{\Lambda,x} - \lambda_\Lambda/2 & \lambda_\Lambda/\sqrt{2} \\ \lambda_\Lambda/\sqrt{2} & E_{\Lambda,z} \end{pmatrix}. \tag{11}$$

For the above eigenstates (5)–(10), all the information about SOC is included in the coefficients u and v, which are given by

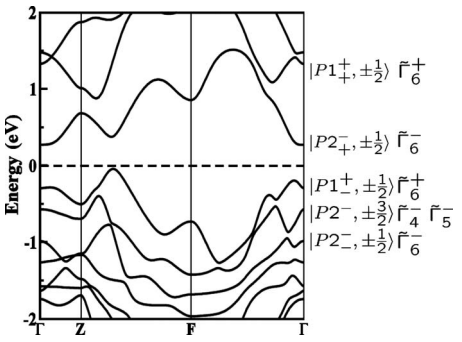

FIG. 3. The band structure of Bi_2Se_3 is obtained from *ab initio* calculation, and the bands near Fermi surface are identified with $|\Lambda^{\pm},\alpha\rangle$. $\Lambda=P1_{\pm},P2_{\pm}$ and $\alpha=\pm\frac{1}{2},\pm\frac{3}{2}$. The corresponding irreducible representation is also given.

$$\begin{pmatrix} u^{\Lambda}_{\pm} \\ v^{\Lambda}_{\pm} \end{pmatrix} = \frac{1}{N_{\pm}} \begin{pmatrix} \Delta E_{\Lambda} \pm \sqrt{(\Delta E_{\Lambda})^2 + \frac{\lambda_{\Lambda}^2}{2}} \\ \lambda_{\Lambda}/\sqrt{2} \end{pmatrix} \tag{12}$$

explicitly, where $N_{\pm}=\lambda_{\Lambda}^2+2\Delta E_{\Lambda}^2\pm 2\Delta E_{\Lambda}\sqrt{\Delta E_{\Lambda}^2+\lambda_{\Lambda}^2/2}$ and $\Delta E_{\Lambda}=\frac{E_{\Lambda,x}-E_{\Lambda,z}-\lambda_{\Lambda}/2}{2}$. The energy splitting between the $p_{x(y)}$ orbital and the p_z orbital due to the crystal field is larger than the energy scale of SOC and ΔE_{Λ} is dominated by $E_{\Lambda,x}-E_{\Lambda,z}$. Now as we see, the SOC couples $|\Lambda,p_z,\uparrow\rangle(|\Lambda,p_z,\downarrow\rangle)$ to $|\Lambda,p_+,\downarrow\rangle(|\Lambda,p_-,\uparrow\rangle)$ so that it induces a level repulsion between these two states. Consequently, $|P1^+_-,\pm\frac{1}{2}\rangle$ is pushed down while $|P2^-_+,\pm\frac{1}{2}\rangle$ is pushed up, which yields a level crossing between these two pairs of states, when the SOC is strong enough, as shown in Fig. 2(IV). Since these two pairs of states have opposite parity, their crossing leads to a *band inversion*, similar to the case of the HgTe quantum wells.[5] This is the key signature of the topological insulator phase in the Bi_2Se_3 family of materials.[1] Therefore in the following we will focus on these four states and regard the other states as a perturbation.

III. MODEL HAMILTONIAN DERIVED FROM SYMMETRY PRINCIPLES

From the discussion of the atomic orbitals in the last section, we obtain an intuitive physical picture of the band structure of Bi_2Se_3. Compared with the *ab initio* calculation, we can denote the bands near the Fermi surface by $|\Lambda^{\pm},\alpha\rangle$ where $\Lambda=P1_{\pm},P2_{\pm}$ and $\alpha=\pm\frac{1}{2},\pm\frac{3}{2}$, as shown in Fig 3. Roughly, these states mainly consist of the bonding or antibonding states of the p orbitals of Bi or Se atoms. However, other orbitals such as the s orbitals of Bi and Se will also mix into these states. To identify each band without any ambiguity, it is necessary to relate each band with the representation of the crystal symmetry. At Γ point, each state should belong to an irreducible representation of the crystal symmetry group and the hybridization between orbitals preserve the symmetry properties. Therefore, a suitable method to identify each band is to use the symmetry of the crystal. In this section, we will first identify each band according to the irreducible representation of the crystal group D^5_{3d} and then try to derive our model Hamiltonian just from symmetry principles.

First let us consider the states without spin, which are denoted as $|\Lambda^{\pm},\alpha\rangle$ with $\Lambda=P1,P2$ and $\alpha=p_x,p_y,p_z$. The crystal of Bi_2Se_3 belongs to the group D^5_{3d} with the character table given in Table I (Appendix A).[39] Since the crystal is inversion symmetric, each representation has a definite parity eigenvalue. For each parity, there are two one-dimensional representations $\tilde{\Gamma}^{\pm}_1$ and $\tilde{\Gamma}^{\pm}_2$ and one two-dimensional representation $\tilde{\Gamma}^{\pm}_3$, where the upper index denotes the parity (+ for even and − for odd). According to the wave functions constructed from the simple atomic orbital picture, we can determine the transformation property of the wave functions under the generators R_3, R_2, and P of the point group. For example, let us look at the operation R_2 on the state $|P1^+,p_x\rangle=\frac{1}{\sqrt{2}}(|B_x\rangle-|B'_x\rangle)$. The R_2 rotation does not change the p_x orbital; however it changes the position of Bi1 (Se1) and Bi1′ (Se1′) and correspondingly changes $|B\rangle$ to $|B'\rangle$; thus we should have $R_2|P1^+,p_x\rangle=-|P1^+,p_x\rangle$. A similar argument can be applied to other states and finally the transformation of the states under the crystal symmetry operations is listed as follows.

(1) *Threefold rotation* R_3: $|\Lambda^{\pm},p_x\rangle\rightarrow\cos\theta|\Lambda^{\pm},p_x\rangle-\sin\theta|\Lambda^{\pm},p_y\rangle$, $|\Lambda^{\pm},p_y\rangle\rightarrow\sin\theta|\Lambda^{\pm},p_x\rangle+\cos\theta|\Lambda^{\pm},p_y\rangle$, and $|\Lambda^{\pm},p_z\rangle\rightarrow|\Lambda^{\pm},p_z\rangle$, with $\theta=\frac{2\pi}{3}$.

(2) *Twofold rotation* R_2: $|\Lambda^{\pm},p_x\rangle\rightarrow\mp|\Lambda^{\pm},p_x\rangle$, $|\Lambda^{\pm},p_y\rangle\rightarrow\pm|\Lambda^{\pm},p_y\rangle$, and $|\Lambda^{\pm},p_z\rangle\rightarrow\pm|\Lambda^{\pm},p_z\rangle$.

(3) *Inversion* P: $|\Lambda^{\pm},\alpha\rangle\rightarrow\pm|\Lambda^{\pm},\alpha\rangle$, $\alpha=p_x,p_y,p_z$.

Here $\Lambda=P1_{\pm},P2_{\pm}$. According to the above transformation, we find that $|\Lambda^{+(-)},p_x\rangle$ and $|\Lambda^{+(-)},p_y\rangle$ belong to the $\tilde{\Gamma}^{+(-)}_3$ representation. $|\Lambda^{+},p_z\rangle$ belongs to the $\tilde{\Gamma}^{+}_1$ representation and $|\Lambda^{-},p_z\rangle$ belongs to the $\tilde{\Gamma}^{-}_2$ representation.

TABLE I. The character table for $D^5_{3d}(R\bar{3}m)$.

$D_{3d}(\bar{3}m)$	E	$2R_3$	$3R_2$	P	$2PR_3$	$3PR_2$
$\tilde{\Gamma}^{+}_1$	1	1	1	1	1	1
$\tilde{\Gamma}^{+}_2$	1	1	−1	1	1	−1
$\tilde{\Gamma}^{+}_3$	2	−1	0	2	−1	0
$\tilde{\Gamma}^{-}_1$	1	1	1	−1	−1	−1
$\tilde{\Gamma}^{-}_2$	1	1	−1	−1	−1	1
$\tilde{\Gamma}^{-}_3$	2	−1	0	−2	1	0

TABLE II. The character table for the double group of $D_{3d}^5(R\bar{3}m)$.

$D_{3d}(\bar{3}m)$	E	$2R_3$	$3R_2$	P	$2PR_3$	$3PR_2$	$\mathcal{C}$	$2\mathcal{C}R_3$	$3\mathcal{C}R_2$	$\mathcal{C}P$	$2\mathcal{C}PR_3$	$3\mathcal{C}PR_2$
$\tilde{\Gamma}_1^+$	1	1	1	1	1	1	1	1	1	1	1	1
$\tilde{\Gamma}_2^+$	1	1	−1	1	1	−1	1	1	−1	1	1	−1
$\tilde{\Gamma}_3^+$	2	−1	0	2	−1	0	2	−1	0	2	−1	0
$\tilde{\Gamma}_4^+$	1	−1	i	1	−1	i	−1	1	$-i$	−1	1	$-i$
$\tilde{\Gamma}_5^+$	1	−1	$-i$	1	−1	$-i$	−1	1	i	−1	1	i
Γ_6^+	2	1	0	2	1	0	−2	−1	0	−2	−1	0
$\tilde{\Gamma}_1^-$	1	1	1	−1	−1	−1	1	1	1	−1	−1	−1
$\tilde{\Gamma}_2^-$	1	1	−1	−1	−1	1	1	1	−1	−1	−1	1
Γ_3^-	2	−1	0	−2	1	0	2	−1	0	−2	1	0
$\tilde{\Gamma}_4^-$	1	−1	i	−1	1	$-i$	−1	1	$-i$	1	−1	i
Γ_5^-	1	−1	$-i$	−1	1	i	−1	1	i	1	−1	$-i$
$\tilde{\Gamma}_6^-$	2	1	0	−2	−1	0	−2	−1	0	2	1	0

To take into account spin, we introduce the spinor representation $\tilde{\Gamma}_6^+$, which changes its sign under the rotation $\mathcal{C}=2\pi$. The double group of D_{3d}^5 can be constructed by the direct product of $\tilde{\Gamma}_{1,2,3}^{\pm}$ and $\tilde{\Gamma}_6^+$. As shown in Eqs. (A1)–(A3), we find that $\tilde{\Gamma}_3^{\pm}\otimes\tilde{\Gamma}_6^+$ will give two new one-dimensional representations $\tilde{\Gamma}_4^{\pm}$ and $\tilde{\Gamma}_5^{\pm}$, which are conjugate to each other. The character table of the double group for D_{3d}^5 is given in Table II (Appendix A).[39] With SOC, the eigenstates in Eqs. (5)–(10) can also be analyzed by the decomposition of direct products. From Eqs. (A2) and (A3), the direct product of $\tilde{\Gamma}_6^{\pm}$ and $\tilde{\Gamma}_{1,2}^{\pm}$ always gives the $\tilde{\Gamma}_6^{\pm}$ representation; therefore $|\Lambda^+,\pm\frac{1}{2}\rangle$ with $\Lambda=P1,P2$ should belong to the $\tilde{\Gamma}_6^+$ representation while $|\Lambda^-,\pm\frac{1}{2}\rangle$ should belong to the $\tilde{\Gamma}_6^-$ representation. The states $|\Lambda^{\pm},\pm 3/2\rangle$ originate from the combination of $|\Lambda,p_{x,y}\rangle$ and spin. According to Eq. (A1), it is expected that $|\Lambda^{\pm},\pm 3/2\rangle$ should be a combination of $\tilde{\Gamma}_4^{\pm}$ and $\tilde{\Gamma}_5^{\pm}$ representations. Indeed by carefully inspecting the transformation behavior under the operations R_2 and R_3, we find that

$$|\Lambda^{\pm},\tilde{\Gamma}_4\rangle=\frac{1}{\sqrt{2}}(|\Lambda^{\pm},3/2\rangle+|\Lambda^{\pm},-3/2\rangle) \tag{13}$$

belongs to the $\tilde{\Gamma}_4^{\pm}$ representation, while

$$|\Lambda^{\pm},\tilde{\Gamma}_5\rangle=\frac{1}{\sqrt{2}}(|\Lambda^{\pm},3/2\rangle-|\Lambda^{\pm},-3/2\rangle) \tag{14}$$

belongs to the $\tilde{\Gamma}_5^{\pm}$ representation. The above results can also be worked out by considering the forms of the transformation for the states (5)–(10), which are given by

(1) *Threefold rotation* R_3: $|\Lambda,\pm\frac{1}{2}\rangle\rightarrow e^{\pm i\pi/3}|\Lambda,\pm\frac{1}{2}\rangle$ and $|\Lambda,\pm\frac{3}{2}\rangle\rightarrow-|\Lambda,\pm\frac{3}{2}\rangle$, where $\Lambda=P1_{\pm}^{\pm},P2_{\pm}^{\pm}$.

(2) *Twofold rotation* R_2: $|\Lambda^+,\pm\frac{1}{2}\rangle\rightarrow i|\Lambda^+,\mp\frac{1}{2}\rangle$, $|\Lambda^-,\pm\frac{1}{2}\rangle\rightarrow-i|\Lambda^-,\mp\frac{1}{2}\rangle$, $|\Lambda^+,\pm\frac{3}{2}\rangle\rightarrow i|\Lambda^+,\mp\frac{3}{2}\rangle$, and $|\Lambda^-,\pm\frac{3}{2}\rangle\rightarrow-i|\Lambda^-,\mp\frac{3}{2}\rangle$, with $\Lambda=P1_{\pm},P2_{\pm}$.

(3) *Inversion* P: $|\Lambda^{\pm},\alpha\rangle\rightarrow\pm|\Lambda^{\pm},\alpha\rangle$, with $\Lambda=P1_{\pm},P2_{\pm}$ and $\alpha=\pm\frac{3}{2},\pm\frac{1}{2}$.

It is instructive to compare the present case with the more common semiconductor crystal structures, such as diamond or zinc-blende structure. In that case, the coupling between p orbitals and the spin usually gives the four-dimensional $\tilde{\Gamma}_8$ and the two-dimensional $\tilde{\Gamma}_7$ representations. In the present case, due to the lower symmetry of the crystal structure, the $\tilde{\Gamma}_7$ representation is the same as the $\tilde{\Gamma}_6$ representation while the $\tilde{\Gamma}_8$ representation is reduced to two one-dimensional representations $\tilde{\Gamma}_4$ and $\tilde{\Gamma}_5$ and one two-dimensional representation $\tilde{\Gamma}_6$. In Fig. 3, the representation of the bands near the Fermi surface is given.

Next we derive our model Hamiltonian to describe the low energy physics of Bi_2Se_3 just based on the symmetry of the wave functions at Γ point. As described above, near the Fermi surface the conduction and valence band are determined by the four states $|P1_-^+,\pm\frac{1}{2}\rangle$ and $|P2_+^-,\pm\frac{1}{2}\rangle$, belonging to the $\tilde{\Gamma}_6^+$ and $\tilde{\Gamma}_6^-$ representations. Therefore the minimum model Hamiltonian for Bi_2Se_3 should be written with these four states as the basis. Generally, any 4×4 Hamiltonian can be expanded with Dirac Γ matrices as

$$\hat{H}_{eff}=\epsilon(\mathbf{k})\mathbf{I}+\sum_i d_i(\mathbf{k})\Gamma_i+\sum_{ij}d_{ij}(\mathbf{k})\Gamma_{ij}, \tag{15}$$

where $\mathbf{I}$ is the 4×4 identity matrix, $\Gamma_i(i=1\cdots5)$ denote the five Dirac Γ matrices satisfying $\{\Gamma_i,\Gamma_j\}=2\delta_{ij}$, and the ten commutators of Γ matrices are given by $\Gamma_{ij}=[\Gamma_i,\Gamma_j]/2i$. $\epsilon(\mathbf{k})$, $d_i(\mathbf{k})$, and $d_{ij}(\mathbf{k})$ can be expanded in the powers of the momentum $\mathbf{k}$. The construction of Γ is given in Appendix B. Now let us assume that the above Hamiltonian is written in the basis $|P1_-^+,\frac{1}{2}\rangle$, $|P2_+^-,\frac{1}{2}\rangle$, $|P1_-^+,-\frac{1}{2}\rangle$, and $|P2_+^-,-\frac{1}{2}\rangle$. Then according to the transformation of the states under the symmetry operation discussed above, we can construct the transformation matrices as follows.

TABLE III. The character table of Γ matrices and the polynomials of the momentum **k**.

	Representation	T
$\{\Gamma_1,\Gamma_2\}$	$\tilde{\Gamma}_3^-$	−
Γ_3	$\tilde{\Gamma}_1^-$	−
Γ_4	$\tilde{\Gamma}_2^-$	−
Γ_5	$\tilde{\Gamma}_1^+$	+
Γ_{12}	$\tilde{\Gamma}_2^+$	−
$\{\Gamma_{23},\Gamma_{31}\}$	$\tilde{\Gamma}_3^+$	−
$\{\Gamma_{14},\Gamma_{24}\}$	$\tilde{\Gamma}_3^+$	−
$\{\Gamma_{15},\Gamma_{25}\}$	$\tilde{\Gamma}_3^-$	+
Γ_{34}	$\tilde{\Gamma}_2^+$	−
Γ_{35}	$\tilde{\Gamma}_1^-$	+
Γ_{45}	$\tilde{\Gamma}_2^-$	+
$\{k_x,k_y\}$	$\tilde{\Gamma}_3^-$	−
k_z,k_z^3	$\tilde{\Gamma}_2^-$	−
$1,k_x^2+k_y^2,k_z^2$	$\tilde{\Gamma}_1^+$	+
$\{k_x^2-k_y^2,2k_xk_y\}$	$\tilde{\Gamma}_3^+$	+
$k_x^3-3k_xk_y^2$	$\tilde{\Gamma}_1^-$	−
$3k_x^2k_y-k_y^3$	$\tilde{\Gamma}_2^-$	−
$\{k_x^3+k_xk_y^2,k_x^2k_y+k_y^3\}$	$\tilde{\Gamma}_3^-$	−
$\{B_x,B_y\}$	$\tilde{\Gamma}_3^+$	−
B_z	$\tilde{\Gamma}_2^+$	−

(1) *Time reversal:* $T=\Theta K$, where $\Theta=i\sigma_2\otimes 1$ and K is the complex conjugate operator.

(2) *Threefold rotation:* $R_3=e^{i(\Pi/2)\theta}$ with $\Pi=\sigma_3\otimes 1$ and $\theta=2\pi/3$.

(3) *Twofold rotation:* $R_2=i\sigma_1\otimes\tau_3$.

(4) *Inversion:* $P=1\otimes\tau_3$.

In the above, σ acts in the spin basis and τ acts in the basis of $P1^+$ and $P2^-$ subbands. According to the above transformation matrices, we can obtain the irreducible representation of each Γ matrix, details of which are derived in Appendix B. The invariance of the Hamiltonian requires that the function $d_i(\mathbf{k})$ $[d_{ij}(\mathbf{k})]$ should have the same behavior to the corresponding $\Gamma_i(\Gamma_{ij})$ matrix under the symmetry operation, which means that they should belong to the same representation of the crystal point group. In Table III Appendix B, we list the representation for both the Γ matrices and the polynomials of **k**, and also their transformation properties under time reversal. Since we hope to preserve both time reversal symmetry and crystal symmetry, we must choose the Γ matrices and polynomials of **k** with the same representation. For example, Γ_1 and Γ_2 carry the representation $\tilde{\Gamma}_3^-$ and are odd under time reversal and so are k_x and k_y. Therefore they can together form an invariant term for the Hamiltonian. Finally, up to $O(k^3)$, our model Hamiltonian yields

$$H'_{eff}=H'_0+H'_3,$$

$$H'_0=\epsilon_{\mathbf{k}}+\mathcal{M}(\mathbf{k})\Gamma_5+\mathcal{B}(k_z)\Gamma_4k_z+\mathcal{A}(k_\parallel)(\Gamma_1k_y-\Gamma_2k_x), \tag{16}$$

$$H'_3=R_1\Gamma_3(k_x^3-3k_xk_y^2)+R_2\Gamma_4(3k_x^2k_y-k_y^3), \tag{17}$$

where $\epsilon_{\mathbf{k}}=C_0+C_1k_z^2+C_2k_\parallel^2$, and $\mathcal{M}(\mathbf{k})=M_0+M_1k_z^2+M_2k_\parallel^2$ and $\mathcal{A}(k_\parallel)=A_0+A_2k_\parallel^2$ and $\mathcal{B}(k_z)=B_0+B_2k_z^2$ and $k_\parallel^2=k_x^2+k_y^2$. H'_0 preserves the in-plane rotation symmetry along the z direction, while H'_3 breaks the in-plane rotation symmetry down to threefold rotation symmetry. In the following, we find that the bulk H'_3 term will also lead to a correction to the effective surface Hamiltonian, which has been studied in Ref. 35.

The above Hamiltonian is the same as that presented by Zhang *et al.*,[1] which can be shown by performing the transformation

$$U_1=\begin{pmatrix}1&0&0&0\\0&-i&0&0\\0&0&1&0\\0&0&0&i\end{pmatrix}, \tag{18}$$

and the Hamiltonian is transformed into

$$H_{eff}=H_0+H_3,$$

$$H_0=U_1H'_0U_1^\dagger=\epsilon_{\mathbf{k}}+\begin{pmatrix}\mathcal{M}(\mathbf{k})&\mathcal{B}(k_z)k_z&0&\mathcal{A}(k_\parallel)k_-\\\mathcal{B}(k_z)k_z&-\mathcal{M}(\mathbf{k})&\mathcal{A}(k_\parallel)k_-&0\\0&\mathcal{A}(k_\parallel)k_+&\mathcal{M}(\mathbf{k})&-\mathcal{B}(k_z)k_z\\\mathcal{A}(k_\parallel)k_+&0&-\mathcal{B}(k_z)k_z&-\mathcal{M}(\mathbf{k})\end{pmatrix} \tag{19}$$

$$H_3=U_1H'_3U_1^\dagger=\frac{R_1(k_+^3+k_-^3)}{2}\begin{pmatrix}0&i&0&0\\-i&0&0&0\\0&0&0&i\\0&0&-i&0\end{pmatrix}+\frac{R_2(k_+^3-k_-^3)}{2}\begin{pmatrix}0&-i&0&0\\-i&0&0&0\\0&0&0&i\\0&0&i&0\end{pmatrix}. \tag{20}$$

Now we can see that H_0 is nearly the same as the model Hamiltonian (1) in Ref. 1, except the A_2 term and B_2 term, which represent the high order correction to the Fermi velocity A_0 and B_0. Since this correction is not important near the Γ point, we will neglect these two terms in the following. Derivation of our model Hamiltonians (19) and (20) from the symmetry principles is the central result of this section.

IV. MODEL HAMILTONIAN DERIVED FROM THE k·p PERTURBATION THEORY

Up to now we have obtained our model Hamiltonian from symmetry principles, or the theory of invariants.[10] In this

section, we will derive the model Hamiltonian in another way, $\mathbf{k}\cdot\mathbf{p}$ theory, and connect the parameters of the model Hamiltonian to the more fundamental matrix elements of momentum in $\mathbf{k}\cdot\mathbf{p}$ theory.

The basic idea of $\mathbf{k}\cdot\mathbf{p}$ theory is to use the wave function at the Γ point in the BZ as the zeroth-order wave function and treat

$$\hat{H}' = \frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{p} \tag{21}$$

as a perturbation, where $\mathbf{p}=-i\hbar\partial_{\mathbf{r}}$ is the momentum operator acting on the zeroth-order wave function and the crystal momentum $\mathbf{k}$ is regarded as a small parameter for the perturbation procedure. The model Hamiltonian is expanded in the powers of k. With the perturbation formalism (C18)–(C21), we can project the full Hamiltonian into the subspace spanned by the four states $|P1^+_-,1/2\rangle\equiv|1\rangle$, $|P2^-_+,1/2\rangle\equiv|2\rangle$, $|P1^+_-,-1/2\rangle\equiv|3\rangle$, and $|P2^-_+,-1/2\rangle\equiv|4\rangle$, which are used as the basis of our model Hamiltonian. All the other states are treated in the perturbation procedure and the details are given in Appendix C. The obtained model Hamiltonian will depend on a series of matrix elements of momentum $\langle\Lambda_1,\alpha|\mathbf{p}|\Lambda_2,\beta\rangle$, which can be simplified due to the symmetry of the crystal. For example, due to inversion symmetry, all the states at the Γ point have definite parity eigenvalues. Since the momentum $\mathbf{p}$ has odd parity, the matrix elements of momentum between two states with the same parity always vanish. The wave function at the Γ point can be obtained through *ab initio* calculations; consequently all these matrix elements can be calculated. With these matrix elements, we apply the perturbation formalism (C18)–(C21) to the system and recover our model Hamiltonians (19) and (20). The parameters of our model Hamiltonian C_0, C_1, C_2, M_0, M_1, M_2, A_0, B_0, R_1, and R_2 can be expressed as the function of the parameters P_{Λ_1,Λ_2}, Q_{Λ_1,Λ_2}, M_{Λ_1,Λ_2}, N_{Λ_1,Λ_2}, R_{Λ_1,Λ_2}, and S_{Λ_1,Λ_2} through Eqs. (C22)–(C31) in Appendix C. With these expressions, we can numerically calculate the values of the parameters of our model Hamiltonian, which is listed in Table IV. We note that the parameters given here are different from those in Refs. 1 and 41 where the parameters are determined by fitting to the energy dispersion, which has some ambiguity. In the present method, since we directly calculate the matrix elements of momentum from microscopic wave functions, there is no ambiguity.

The key result of this section are the parameters of our model Hamiltonian, given in Table IV. The fitted energy dispersions for Bi_2Se_3, Bi_2Te_3, and Sb_2Te_3 are plotted in Fig 4. It is shown that our model Hamiltonian is valid in the regime $k_{x,z}<0.04$ Å^{-1}. However the maximum of the valence band for Bi_2Se_3 and Bi_2Te_3 stays away from the Γ point, at about $k_x\approx 0.07$ Å^{-1}; therefore we need to keep in mind that there may be some discrepancies when we try to use our model Hamiltonian to describe Bi_2Se_3 quantitatively. Since our model Hamiltonian with four bands already captures the salient features of the band dispersion, especially the inverted band structure, in the following two sections, we still stay in the framework of our model Hamiltonian to discuss the topological surface states and the Landau levels in the magnetic field. Then in the last section, we will extend our model Hamiltonian to include eight bands, in order to describe these materials more quantitatively.

TABLE IV. The summary of the parameters in our model Hamiltonian with four bands.

	Bi_2Se_3	Bi_2Te_3	Sb_2Te_3
A_0 (eV Å)	3.33	2.87	3.40
B_0 (eV Å)	2.26	0.30	0.84
C_0 (eV)	−0.0083	−0.18	0.001
C_1 (eV Å^2)	5.74	6.55	−12.39
C_2 (eV Å^2)	30.4	49.68	−10.78
M_0 (eV)	−0.28	−0.30	−0.22
M_1 (eV Å^2)	6.86	2.79	19.64
M_2 (eV Å^2)	44.5	57.38	48.51
R_1 (eV Å^3)	50.6	45.02	103.20
R_2 (eV Å^3)	−113.3	−89.37	−244.67
g_{1z}	−25.4	−50.34	−14.45
g_{1p}	−4.12	−2.67	−2.43
g_{2z}	4.10	6.88	14.32
g_{2p}	4.80	3.43	16.55

V. SURFACE STATES

An important physical consequence of the nontrivial topology is the existence of topological surface states. In this section, we would like to study the surface state and its effective Hamiltonian based on our model Hamiltonian derived above with open boundary conditions.

Consider our model Hamiltonian (17) defined on the half space given by $z>0$. We can divide our model Hamiltonian into two parts

$$\hat{H} = \tilde{H}_0 + \tilde{H}_1, \tag{22}$$

$$\tilde{H}_0 = \tilde{\epsilon}(k_z) + \tilde{M}(k_z)\Gamma_5 + B_0\Gamma_4 k_z, \tag{23}$$

$$\tilde{H}_1 = C_2 k_\parallel^2 + M_2 k_\parallel^2\Gamma_5 + A_0(\Gamma_1 k_y - \Gamma_2 k_x) + H_3. \tag{24}$$

where $\tilde{\epsilon}(k_z)=C_0+C_1k_z^2$ and $\tilde{M}(k_z)=M_0+M_1k_z^2$. All k_z dependent terms are included in $\tilde{H}_0$. We replace k_z by $-i\partial_z$ and obtain the eigenvalue equation

$$\tilde{H}_0(k_z\to -i\partial_z)\Psi(z) = E\Psi(z). \tag{25}$$

Since $\Gamma_4=1\otimes\tau_2$ and $\Gamma_5=1\otimes\tau_3$ are both block diagonal, the Hamiltonian $\tilde{H}_0$ is also block diagonal and the eigenstates have the form

$$\Psi_\uparrow(z) = \begin{pmatrix}\psi_0\\ \mathbf{0}\end{pmatrix}, \quad \Psi_\downarrow(z) = \begin{pmatrix}\mathbf{0}\\ \psi_0\end{pmatrix}, \tag{26}$$

where $\mathbf{0}$ is a two-component zero vector. $\Psi_\uparrow(z)$ is related to $\Psi_\downarrow(z)$ by time-reversal operation. To obtain the surface states, the wave function $\psi_0(z)$ should be localized at the surface and satisfies the eigenequation

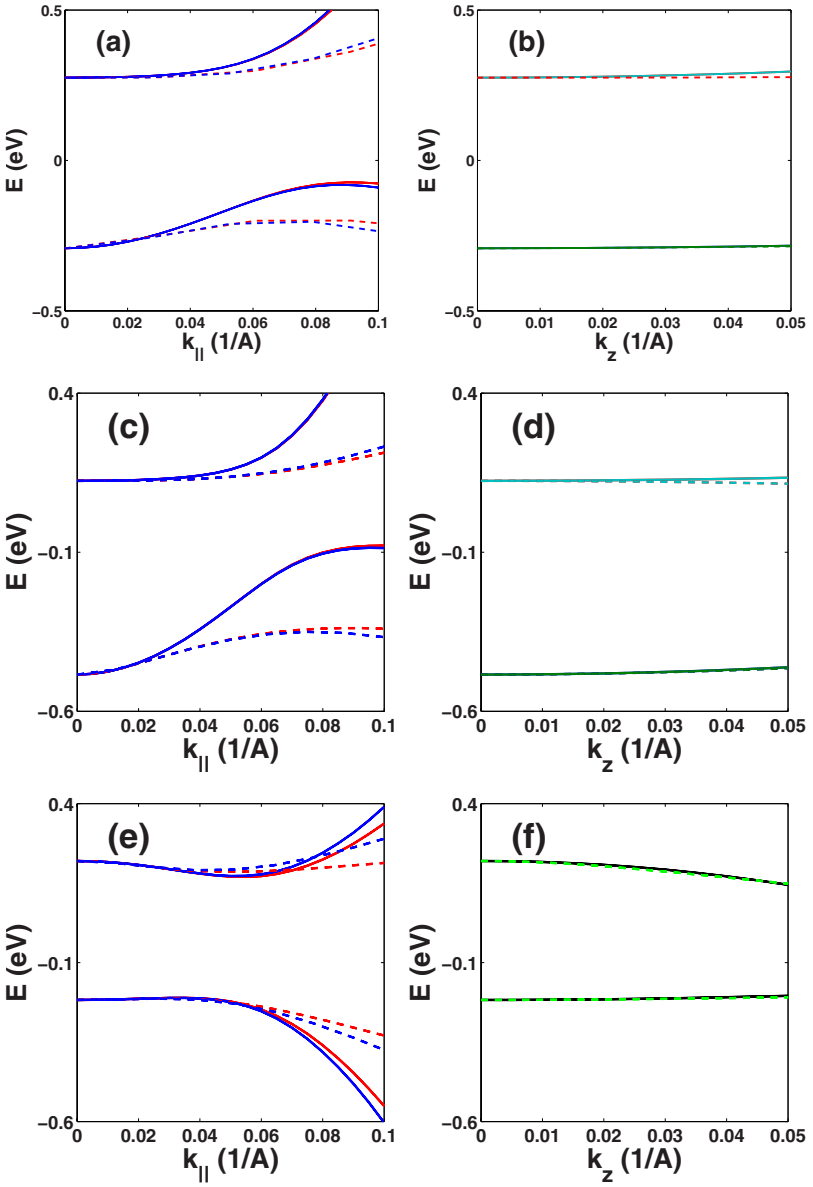

FIG. 4. (Color online) The energy dispersion obtained from our model Hamiltonian with four bands (solid line) is compared with that from *ab initio* calculation (dashed line). Here (a) and (b) is for Bi_2Se_3, (c) and (d) is for Bi_2Te_3, while (e) and (f) is for Sb_2Te_3. In (a), (c), and (e), the red line represents the dispersion along the k_x direction while the blue line is for the k_y direction.

$$(\tilde{\epsilon}(-i\partial_z) + \tilde{M}(-i\partial_z)\tau_3 - iB_0\tau_2\partial_z)\psi_0(z) = E\psi_0(z), \quad (27)$$

which has been solved analytically for an open boundary condition with different methods.[31,32,34,41,42] Here in order to show the existence of the surface states and to find the region where the surface states exist, we would like to briefly review the derivation for the explicit form of the surface states by neglecting $\tilde{\epsilon}$ for simplicity.[31]

After neglecting the $\tilde{\epsilon}$ term, the eigenequation (27) exhibits the particle hole symmetry; therefore we expect that special surface states with $E=0$ can exist. With the wave function ansatz $\psi_0=\phi e^{\lambda z}$, the above equation can be simplified as

$$(M_0 - M_1\lambda^2)\tau_1\phi = B_0\lambda\phi. \quad (28)$$

It is obvious that the two-component wave function ϕ should be the eigenstate of the Pauli matrix τ_1. Let us define $\tau_1\phi_\pm = \pm\phi_\pm$, then Eq. (28) is simplified to a quadratic equation for λ. Another important observation is that if λ is a solution for ϕ_+, then $-\lambda$ is a solution for ϕ_-. Consequently, the generic wave function is given by

$$\psi_0(z) = (ae^{\lambda_1 z} + be^{\lambda_2 z})\phi_+ + (ce^{-\lambda_1 z} + de^{-\lambda_2 z})\phi_-, \quad (29)$$

where $\lambda_{1,2}$ satisfy

$$\lambda_{1,2} = \frac{1}{2M_1}(-B_0 \pm \sqrt{4M_0M_1 + B_0^2}). \quad (30)$$

Similar to Ref. 31, the open boundary condition $\psi(0)=0$, together with the normalizability of the wave function in the region $z>0$, leads to the existence condition of the surface states, $\Re\lambda_{1,2}<0$ ($c=d=0$) or $\Re\lambda_{1,2}>0$ ($a=b=0$), which can only be satisfied with the band inversion condition $M_0M_1<0$. Furthermore, it is easy to show that when $B_0/M_1>0$, $\Re\lambda_{1,2}<0$, while $B_0/M_1<0$, $\Re\lambda_{1,2}>0$, thus the wave function for the surface states at Γ point is given by

$$\psi_0(m) = \begin{cases} a(e^{\lambda_1 z} - e^{\lambda_2 z})\phi_+, & B_0/M_1 > 0 \\ c(e^{-\lambda_1 z} - e^{-\lambda_2 z})\phi_-, & B_0/M_1 < 0. \end{cases} \quad (31)$$

We emphasize here that the sign change of B_0/M_1 will change the spin basis of the surface states, which will be the key point to determine the helicity of the Dirac Hamiltonian for the topological surface states. Another important quantity of the surface states is the decaying length, which can be defined as $l_c = \max\{\frac{1}{|\Re(\lambda_{1,2})|}\}$, given by

$$l_c = \begin{cases} \Re\left(\dfrac{B_0 + \sqrt{4M_0M_1 + B_0^2}}{2M_0}\right), & B_0 > 0,\ M_1 < 0 \\ \Re\left(\dfrac{B_0 - \sqrt{4M_0M_1 + B_0^2}}{2M_0}\right), & B_0 < 0,\ M_1 > 0 \\ \Re\left(-\dfrac{B_0 + \sqrt{4M_0M_1 + B_0^2}}{2M_0}\right), & B_0 > 0,\ M_1 > 0 \\ \Re\left(-\dfrac{B_0 - \sqrt{4M_0M_1 + B_0^2}}{2M_0}\right), & B_0 < 0,\ M_1 < 0, \end{cases} \quad (32)$$

where $\Re$ takes the real part.

In the above, we take a simple tight-binding model to show the existence condition and the form of the wave function for the surface states, which can help us to understand the underlying physics qualitatively, but not quantitatively. In the realistic materials, the detailed form of ψ_0 will depend on material details, such as the boundary condition or the detailed parameters; however the form of the wave function (26) remains valid. Therefore in the following, we just simply treat ψ_0 with some parameters. In the subspace $\Psi = [\Psi_\uparrow, \Psi_\downarrow]$, we find that

$$\langle\Psi|\Gamma_1|\Psi\rangle = \alpha_1\sigma_x, \quad \langle\Psi|\Gamma_2|\Psi\rangle = \alpha_1\sigma_y,$$

$$\langle\Psi|\Gamma_3|\Psi\rangle = \alpha_1\sigma_z, \quad \langle\Psi|\Gamma_4|\Psi\rangle = 0,$$

$$\langle\Psi|\Gamma_5|\Psi\rangle = \alpha_3, \quad (33)$$

with $\alpha_1 \equiv \langle\psi_0|\tau_1|\psi_0\rangle$ and $\alpha_3 \equiv \langle\psi_0|\tau_3|\psi_0\rangle$. With these expressions, the effective Hamiltonian for the surface states Ψ is given by[1]

$$\begin{aligned} H_{sur} &= \tilde{C}_0 + \tilde{C}_2 k_\parallel^2 + \tilde{A}(\sigma_x k_y - \sigma_y k_x) + \tilde{R}(k_+^3 + k_-^3)\sigma_z \\ &= (C_0 + \alpha_3 M_0) + (C_2 + \alpha_3 M_2)k_\parallel^2 + A_0\alpha_1(\sigma_x k_y - \sigma_y k_x) \\ &\quad + \frac{R_1\alpha_1}{2}(k_+^3 + k_-^3)\sigma_z, \end{aligned} \quad (34)$$

with $k_\pm = k_x \pm ik_y = k_\parallel e^{\pm i\theta}$. The k^3 terms have also been found in Ref. 35. In the following numerical calculation, the coefficient α_1 and α_3 are treated as two fitting parameters to the experiment, given by $\alpha_1 = \frac{\tilde{A}_{exp}}{A_0} = 0.99$ and $\alpha_3 = \frac{\tilde{C}_{exp} - C_0}{M_0} = -0.15$, where $\tilde{A}_{exp} = 3.29$ eV Å comes from the Fermi velocity of the surfaces states and $\tilde{C}_{exp} = 0.035$ eV comes from the position of the surface Dirac points.[10] Moreover we need to check the spin operators in this system. Again we use the wave function from *ab initio* calculations and project the spin operator into the subspace spanned by the four basis states. After we obtain the spin operators for our model Hamiltonian, we can use the eigenwave function (26) to project the spin operator onto the surface states subspace. Finally we find that $\langle\Psi|S_x|\Psi\rangle = S_{x0}\sigma_x$, $\langle\Psi|S_y|\Psi\rangle = S_{y0}\sigma_y$, and $\langle\Psi|S_z|\Psi\rangle = S_{z0}\sigma_z$ with $S_{x(y,z)0}$ to be some constants. This indicates that σ matrix in our model Hamiltonian (34) is proportional to the real spin.

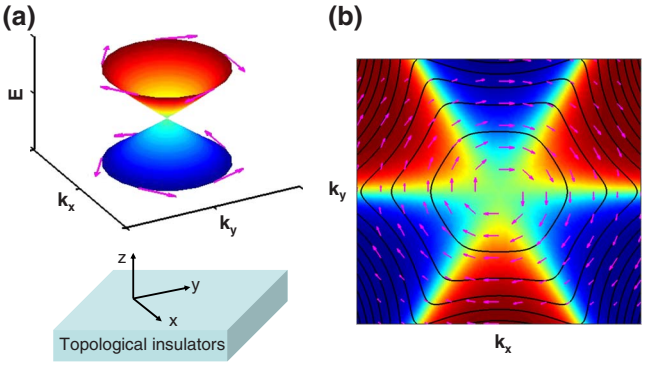

FIG. 5. (Color online) (a) Spin texture of the surface states near the Γ point. For the conduction band, the helicity is left handed while for the valence band, it is right handed. (b) Spin texture of the conduction band of the surface states in momentum space. The arrow represents the x-y planar spin polarization while the color indicates the z component of the spin polarization. Here red is for spin up while blue is for spin down. The black line gives the constant energy contours.

The derivation of the surface Hamiltonian (34) is the central result of this section. In the limit $k \to 0$, the linear term in Hamiltonian (34) will be dominant, then the surface states show the linear dispersion with helical spin texture, which has the opposite direction for the conduction and valence band, as shown in Fig. 5(a). Such type of spin texture is similar to one of the Fermi surfaces in the usual 2D electron gas with Rashba SOC,[40,43] which can be simply understood from the fact that the inversion symmetry is broken near the surface. The helical spin texture has also been calculated by *ab initio* methods[44] and already observed in the pioneering spin-resolved ARPES measurement.[18] From Eq. (34), the helicity of the spin texture is determined by the sign of the coefficient $\tilde{A} = A_0\alpha_1$, where $\alpha_1 \equiv \langle\psi_0|\tau_1|\psi_0\rangle$ is related to the spin basis of the surface states (31). Therefore the helicity is determined by the relative sign of A_0 and B_0/M_1. Furthermore, due to the inversion condition $M_1M_0 < 0$, the sign of M_1 is already determined by the gap M_0. Consequently, within our model Hamiltonian the helicity of the spin texture is given by the relative sign of the coefficients of two types of linear terms, A_0 and B_0.

To further explore the origin of the helicity of the spin texture in the atomic levels, we relate the coefficients A_0 and B_0 to the atomic SOC by using expressions (7)–(10) as

$$\begin{aligned} A_0 &= \frac{\hbar}{2m_0}\left\langle P1_-^+, \frac{1}{2}\middle| p_+ \middle| P2_+^-, -\frac{1}{2}\right\rangle \\ &= \frac{\hbar}{2m_0}[(u_-^{P1^+}v_+^{P2^-})^*\langle P1^+, p_z|p_+|P2^-, p_-\rangle \\ &\quad + (v_-^{P1^+}u_+^{P2^-})^*\langle P1^+, p_+|p_+|P2^-, p_z\rangle], \end{aligned} \quad (35)$$

$$B_0 = \frac{\hbar}{m_0}\langle P1_z^+, p_z|p_z|P2^-, p_z\rangle$$
$$= \frac{\hbar}{m_0}[(u_-^{P1^+})^* u_+^{P2^-}\langle P1^+, p_z|p_z|p2^-, p_z\rangle$$
$$+ (v_-^{P1^+})^* v_+^{P2^-}\langle P1^+, p_+|p_z|P2^-, p_+\rangle]. \quad (36)$$

Here $|\Lambda, \alpha\rangle$ ($\Lambda = P1^+$, $P2^-$ and $\alpha = p_x, p_y, p_z$) are the atomic orbitals without any SOC and all the dependence on SOC is included in the coefficients u^Λ and v^Λ. From Eq. (12), we find that for $u^{\Lambda_1} v^{\Lambda_2}$ it is proportional to λ_{Λ_1} (or λ_{Λ_2}), which indicates that A_0 depends on the sign of the atomic SOC, while it is only possible for $(u^{\Lambda_1})^* u^{\Lambda_2}$ and $(v^{\Lambda_1})^* v^{\Lambda_2}$ to be independent of $\lambda_{\Lambda_{1,2}}$ or be proportional to $\lambda_{\Lambda_1}\lambda_{\Lambda_2}$, thus the sign of B_0 will not depend on the atomic SOC. Finally, we conclude that the helicity of the spin texture is originally related to the atomic SOC.

In the above, we have shown how the linear term determines the spin texture of the surface states, which can also be affected by the quadratic and cubic terms in the effective Hamiltonian (34). We solve the eigenvalue problem of the whole effective Hamiltonian and obtain the eigenenergy and eigenstates as

$$E_\pm = \tilde{C}_0 + \tilde{C}_2 k_\parallel^2 \pm \sqrt{\tilde{A}^2 k_\parallel^2 + 4\tilde{R}^2 k^6 \cos^2 3\theta}, \quad (37)$$

$$\psi_\pm = \frac{1}{\sqrt{N}}\begin{pmatrix} \tilde{A}(k_y + ik_x) \\ d_\pm - \tilde{R}(k_+^3 + k_-^3) \end{pmatrix}, \quad (38)$$

with $d_\pm = \pm\sqrt{\tilde{A}^2 k_\parallel^2 + 4\tilde{R}^2 k^6 \cos^2 3\theta}$ and $N = \tilde{A}^2 k^2 + (\sqrt{\tilde{A}^2 k^2 + 4\tilde{R}^2 k^6 \cos^2 3\theta} - 2\tilde{R}k^3 \cos 3\theta)^2$. Consequently the spin polarization in k space is given by

$$\langle\psi_+|\sigma_x|\psi_+\rangle = \frac{2\tilde{A}k_y}{N}(d_+ - 2\tilde{R}k^3 \cos 3\theta), \quad (39)$$

$$\langle\psi_+|\sigma_y|\psi_+\rangle = -\frac{2\tilde{A}k_x}{N}(d_+ - 2\tilde{R}k^3 \cos 3\theta), \quad (40)$$

$$\langle\psi_+|\sigma_z|\psi_+\rangle = \frac{4\tilde{R}k^3 \cos 3\theta}{N}(d_+ - 2\tilde{R}k^3 \cos 3\theta), \quad (41)$$

which is plotted in Fig. 5(b). In the limit $k \to 0$, the spin polarization almost lies in the xy plane, which is due to the linear term and has been discussed in the above. When k is increased, the k-cubic term comes into play, which will not only induce the hexagonal warping of the constant energy contours[35] but also yield z direction spin polarization, similar to the situation in Bi_xSb_{1-x} studied by *ab initio* calculations.[45]

VI. MAGNETIC FIELD AND LANDAU LEVEL

In this section, we study the Landau level problem for both the bulk states and the surface states, which is important for predicting or understanding many properties of the system in a magnetic field, such as the Shubnikov–de Haas oscillation, surface quantum Hall effect and magneto-optics. In this regards, our model Hamiltonian has a unique advantage because the magnetic field effect cannot be incorporated in *ab initio* calculations. For a realistic finite sample, the bulk Landau levels will always coexist with the surface Landau levels; thus both the two types of Landau levels need to be taken into account. For simplicity, we solve the bulk Landau levels for an infinite sample and the surface Landau level for the semi-infinite sample. The mixing between bulk and surface Landau levels is neglected here.

For bulk states, there are two types of contribution from the magnetic field, the orbital effect and the Zeeman effect. The orbital effect can be included by Peierls substitution[46] $\mathbf{k} \to \pi = \mathbf{k} + \frac{e}{\hbar}\mathbf{A}$ with $\mathbf{A} = (0, B_z x, 0)$ for magnetic field along the z direction. We introduce the annihilation and creation operators $a = \frac{l_c}{\sqrt{2}}\pi_-$ and $a^\dagger = \frac{l_c}{\sqrt{2}}\pi_+$ with $l_c = \sqrt{\frac{\hbar}{eB_z}}$ for the harmonic oscillator function φ_n. a and $a^\dagger$ satisfy $a\varphi_N = \sqrt{N}\varphi_{N-1}$, $a^\dagger\varphi_N = \sqrt{N+1}\varphi_{N+1}$, and $[a, a^\dagger] = 1$. With the operators a and $a^\dagger$, the Hamiltonian (19) is written as

$$\hat{H}_{0B} = \tilde{\epsilon} + \begin{pmatrix} \tilde{\mathcal{M}} & B_0 k_z & 0 & A_0\frac{\sqrt{2}}{l_c}a \\ B_0 k_z & -\tilde{\mathcal{M}} & A_0\frac{\sqrt{2}}{l_c}a & 0 \\ 0 & A_0\frac{\sqrt{2}}{l_c}a^\dagger & \tilde{\mathcal{M}} & -B_0 k_z \\ A_0\frac{\sqrt{2}}{l_c}a^\dagger & 0 & -B_0 k_z & -\tilde{\mathcal{M}} \end{pmatrix}, \quad (42)$$

where $\tilde{\epsilon}(k_z, a^\dagger a) = C_0 + C_1 k_z^2 + \frac{2C_2}{l_c^2}(a^\dagger a + \frac{1}{2})$ and $\tilde{M}(k_z, a^\dagger a) = M_0 + M_1 k_z^2 + \frac{2M_2}{l_c^2}(a^\dagger a + \frac{1}{2})$. The k^3 term, which breaks the in-plane rotation symmetry, is neglected here; therefore the wave function should have the form of $\Psi_N = [f_1^N\varphi_{N-1}, f_2^N\varphi_{N-1}, f_3^N\varphi_N, f_4^N\varphi_N]^T$. With this wave function ansatz, the Hamiltonian is transformed to

$$\hat{H}_{0B}(k_z, N) = \begin{pmatrix} \tilde{\mathcal{M}}_{N-1}^+ & B_0 k_z & 0 & A_0\frac{\sqrt{2N}}{l_c} \\ B_0 k_z & \tilde{\mathcal{M}}_{N-1}^- & A_0\frac{\sqrt{2N}}{l_c} & 0 \\ 0 & A_0\frac{\sqrt{2N}}{l_c} & \tilde{\mathcal{M}}_N^+ & -B_0 k_z \\ A_0\frac{\sqrt{2N}}{l_c} & 0 & -B_0 k_z & -\tilde{\mathcal{M}}_N^- \end{pmatrix}, \quad (43)$$

with $\tilde{\mathcal{M}}_N^+ = \tilde{\epsilon}(k_z, N) + \tilde{\mathcal{M}}(k_z, N)$ and $\tilde{\mathcal{M}}_N^- = \tilde{\epsilon}(k_z, N) - \tilde{\mathcal{M}}(k_z, N)$.

To consider the Zeeman splitting, we need to further calculate the effective g factor.[40,47] The effective Zeeman type coupling can also be written down for our model Hamiltonian by symmetry principles. By a quick inspection of Table III, we find that the following terms:

$$\hat{H}_Z = \tilde{g}_{z1}\Gamma_{12}B_z + \tilde{g}_{z2}\Gamma_{34}B_z + \tilde{g}_{xy1}(\Gamma_{23} \quad \Gamma_{31}) \times \begin{pmatrix} \cos\phi_1 & \sin\phi_1 \\ -\sin\phi_1 & \cos\phi_1 \end{pmatrix} \begin{pmatrix} B_x \\ B_y \end{pmatrix} + \tilde{g}_{xy2}(\Gamma_{14} \quad \Gamma_{24}) \times \begin{pmatrix} \cos\phi_2 & \sin\phi_2 \\ -\sin\phi_2 & \cos\phi_2 \end{pmatrix} \begin{pmatrix} B_x \\ B_y \end{pmatrix} \tag{44}$$

are possible couplings to the magnetic field. Again ϕ_1 and ϕ_2 are phase factors which need to be determined from other methods, and here we take $\phi_1=\phi_2=0$ to coincide with the results from the $\mathbf{k}\cdot\mathbf{p}$ method and explicitly Eq. (44) can be written as

$$\hat{H}_Z = \frac{\mu_B}{2}\begin{pmatrix} g_{1z}B_z & 0 & g_{1p}B_- & 0 \\ 0 & g_{2z}B_z & 0 & g_{2p}B_- \\ g_{1p}B_+ & 0 & -g_{1z}B_z & 0 \\ 0 & g_{2p}B_+ & 0 & -g_{2z}B_z \end{pmatrix}, \tag{45}$$

with $\mu_B=\frac{e\hbar}{2m_0}$, and $\tilde{g}_{xy1}+\tilde{g}_{xy2}=\frac{\mu_B}{2}g_{1p}$, $\tilde{g}_{xy1}-\tilde{g}_{xy2}=\frac{\mu_B}{2}g_{2p}$, $\tilde{g}_{z1}+\tilde{g}_{z2}=\frac{\mu_B}{2}g_{1z}$, and $\tilde{g}_{z1}-\tilde{g}_{z2}=\frac{\mu_B}{2}g_{2z}$. This model Hamiltonian can also be derived from $\mathbf{k}\cdot\mathbf{p}$ theory and the parameters g_{1z}, g_{2z}, g_{1p}, and g_{2p} can be related to the matrix elements of the momentum operator $\mathbf{p}$ in the $\mathbf{k}\cdot\mathbf{p}$ theory, with full details given in Appendix C. Now our total Hamiltonian for the bulk states under the z-direction magnetic field is given by $\hat{H}_B=\hat{H}_{0B}+\hat{H}_Z$, which can be solved numerically to obtain the Landau level $E^{bulk}_{N,\eta}(B,k_z)$ with Landau level index N and band index η under the z-direction magnetic field.

A similar procedure can be applied to the surface effective Hamiltonian (34). With the wave function ansatz $\Psi_{sur,N}=[g_1^N\varphi_{N-1}, g_2^N\varphi_N]^T$, the surface Hamiltonian is changed to

$$\hat{H}_{sur,B0}(N) = \tilde{C}_0 + \frac{2NeB_z}{\hbar}\tilde{C}_2 - \frac{\tilde{C}_2 eB_z}{\hbar}\sigma_z - \sqrt{\frac{2NeB_z}{\hbar}}\tilde{A}\sigma_y, \tag{46}$$

and the Zeeman type term is given by

$$\hat{H}_{sur,Z} = \frac{\mu_B}{2}g_{sz}\sigma_z B_z + \frac{\mu_B}{2}g_{sp}(\sigma_x B_x + \sigma_y B_y), \tag{47}$$

with $\frac{\mu_B}{2}g_{sz}=\tilde{g}_{z1}+\tilde{g}_{z2}\alpha_3$ and $\frac{\mu_B}{2}g_{sp}=\tilde{g}_{xy1}+\tilde{g}_{xy2}\alpha_3$. The total Hamiltonian for the surface states yields $\hat{H}_{sur,B}=\hat{H}_{sur,B0}+\hat{H}_{sur,Z}$ and correspondingly the Landau level in z-direction magnetic field B_z is solved as

$$E_s^{sur}(N) = \tilde{C}_0 + \frac{2NeB_z}{\hbar}\tilde{C}_2 + s\sqrt{\left(-\frac{\tilde{C}_2 eB_z}{\hbar} + \frac{\mu_B}{2}g_{sz}B_z\right)^2 + \frac{2NeB_z}{\hbar}\tilde{A}^2}, \tag{48}$$

with $s=\pm$ for $N=1,2,\ldots$ and

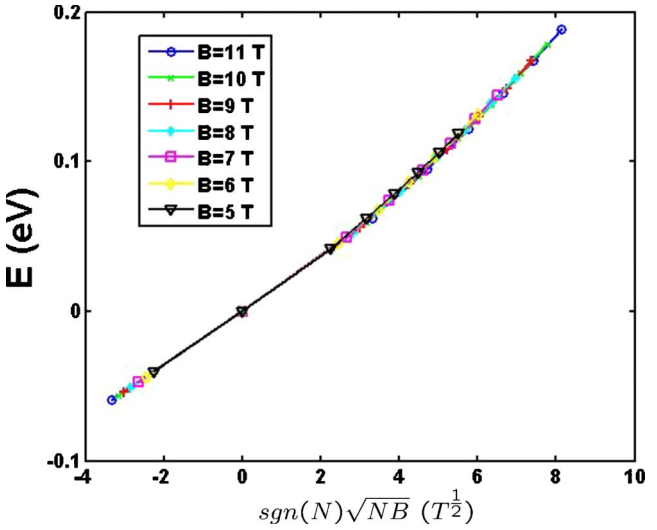

FIG. 6. (Color online) The energies of the Landau levels versus $sgn(N)\sqrt{NB}$ where N is the Landau level index and B is the magnetic field.

$$E^{sur}(0) = \tilde{C}_0 + \frac{eB_z}{\hbar}\tilde{C}_2 - \frac{\mu_B}{2}g_{sz}B_z \tag{49}$$

for the zero mode $N=0$. Here we note that due to the existence of the quadratic term $\tilde{C}_2k_\parallel^2$, the square-root dependence of the energy level versus magnetic field is only an approximation applicable for low Landau levels and low magnetic field. For high magnetic field, it will be a combination of the linear contribution and square root contribution. As shown in Fig. 6, the energy of the Landau levels is plotted as a function of $\mathrm{sgn}(N)\sqrt{NB_z}$ and the nonlinear behavior will appear for large $\sqrt{NB_z}$.

In Fig. 7(b), the Landau levels for both bulk and surface states are plotted as a function of magnetic field. Here we emphasize that for the bulk states, the Landau levels are plotted for $k_z=0$ [red line in Fig. 7(b)]. The dispersion along the z direction is also shown in Fig. 7(a) for $B=0$ T and (c) for $B=20$ T. From Fig. 7(a), we find that the maximum of

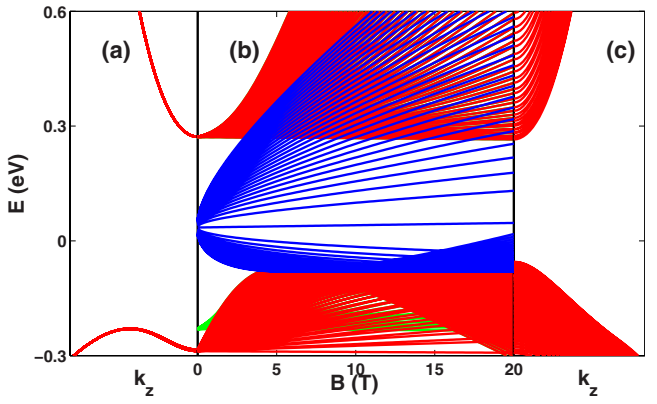

FIG. 7. (Color online) The Landau levels in the magnetic field for both the bulk states (red lines) and the surface states (blue lines) are shown in (b). Also the dispersion along the z direction at (a) $B=0$ T and (c) $B=20$ T is plotted.

the valence band is not located at $k_z=0$ for small magnetic fields, hence we plot the maximum of valence band as the green lines in Fig. 7(b). When the magnetic field increases, the bulk gap also decreases by a significant amount (about 160 meV for 5 T magnetic field), which is due to the double hump structure for the valence-band dispersion of Bi_2Se_3. Such decrease may be observed in a magneto-optical measurement.

In order to compare with the STM experiment,[22,23] it is helpful to investigate the local density of states (LDOS) at the surface. The LDOS for the surface states and the bulk states can be obtained by.[40]

$$D_{sur}(E,B)=\sum_{N,s}\frac{G}{\sqrt{2\pi}\Gamma}e^{-[E-E_{N,s}^{sur}(B)]^2/2\Gamma^2} \tag{50}$$

and

$$D_{bulk}(E,B)=L_0\int\frac{dk_z}{2\pi}\sum_{N,s}\frac{G}{\sqrt{2\pi}\Gamma}e^{-[E-E_{N,\eta}^{bulk}(B,k_z)]^2/2\Gamma^2}, \tag{51}$$

respectively, where $G=\frac{eB_z}{2\pi\hbar}=\frac{1}{2\pi l_c^2}$ is the degeneracy of each Landau level and Γ is the broadening. In order to compare the bulk LDOS with the surface LDOS, we require to introduce a length scale L_0 which represents the detection depth of STM. Here we simply take L_0 to be the thickness of one quintuple layer. Furthermore the surface states only exist near $\mathbf{k}=0$; thus we need to take a cutoff for the Landau level index N. With formulas (50) and (51), the LDOS for both bulk and surface Landau levels is shown in Fig. 8. The bulk LDOS shows a gap of about 0.3 eV and within the bulk gap, only surface LDOS remains and shows clearly the Landau levels as discrete peaks. The largest Landau gap for surface states is between the zeroth and first Landau levels, about 50 meV, which is large enough for the observation of the topological magnetoelectric effect.[36,37]

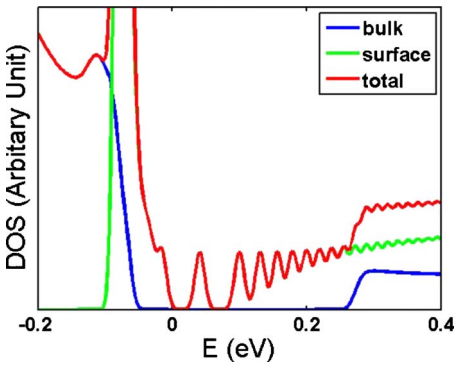

FIG. 8. (Color online) The density of states as a function of energy is plotted with $B=10$ T for the bulk (blue line), surface (green line), and total (red line) Landau levels.

VII. MODEL HAMILTONIAN WITH EIGHT BANDS

As we have described above, our model Hamiltonian can capture the salient topological features of the Bi_2Se_3 family of materials. However, for a full quantitative fitting to first principle calculations, we need to expand the basis set. By inspecting carefully the $\mathbf{k}\cdot\mathbf{p}$ matrix elements, we find that there are strong couplings between the state $|P1_-^+,\pm\frac{1}{2}\rangle$ and the state $|P2^-,\tilde{\Gamma}_{4,5}\rangle$ or $|P2_-^-,\pm\frac{1}{2}\rangle$. For example, at the valence-band maximum $k_x\approx 0.07$ Å^{-1}, we find that these couplings can be as large as the energy gap between these states. Therefore it is not surprising that our model Hamiltonian with four bands is not suitable in this regime. The strong couplings between these states indicate that if we want to describe this material more accurately, we need to further include the states $|P2^-,\tilde{\Gamma}_{4,5}\rangle$ and $|P2_-^-,\pm\frac{1}{2}\rangle$ into our model Hamiltonian. In the basis sequence $|P1_-^+,\frac{1}{2}\rangle$, $|P1_-^+,-\frac{1}{2}\rangle$, $|P2_+^-,\frac{1}{2}\rangle$, $|P1_+^-,-\frac{1}{2}\rangle$, $|P2^-,\tilde{\Gamma}_4\rangle$, $|P2^-,\tilde{\Gamma}_5\rangle$, $|P2_-^-,\frac{1}{2}\rangle$ and $|P2_-^-,-\frac{1}{2}\rangle$, following the similar perturbation procedure, we find that our model Hamiltonian is written as

$$\hat{H}=\frac{\hbar^2}{2m_0}\begin{pmatrix}
f_1(\mathbf{k}) & 0 & \frac{2}{\hbar}k_zQ_1 & \frac{2}{\hbar}P_1k_- & \frac{2}{\hbar}Q_2k_+ & \frac{2}{\hbar}k_+P_2 & \frac{2}{\hbar}k_zQ_3 & \frac{2}{\hbar}k_-P_3\\
 & f_1(\mathbf{k}) & \frac{2}{\hbar}k_+P_1^* & -\frac{2}{\hbar}k_zQ_1^* & -\frac{2}{\hbar}P_2^*k_- & \frac{2}{\hbar}Q_2^*k_- & \frac{2}{\hbar}P_3^*k_+ & -\frac{2}{\hbar}Q_3^*k_z\\
 & & f_3(\mathbf{k}) & 0 & g_{35}(\mathbf{k}) & g_{36}(\mathbf{k}) & f_{37}(\mathbf{k}) & -g_{47}^*(-\mathbf{k})\\
 & & & f_3(\mathbf{k}) & g_{36}^*(-\mathbf{k}) & -g_{35}^*(-\mathbf{k}) & g_{47}(\mathbf{k}) & f_{37}^*(-\mathbf{k})\\
 & & & & f_5(\mathbf{k}) & 0 & -g_{68}^*(-\mathbf{k}) & g_{58}(\mathbf{k})\\
 & & \text{H.c.} & & & f_5(\mathbf{k}) & g_{58}^*(-\mathbf{k}) & g_{68}(\mathbf{k})\\
 & & & & & & f_7(\mathbf{k}) & 0\\
 & & & & & & & f_7(\mathbf{k})
\end{pmatrix}, \tag{52}$$

TABLE V. The summary of the parameters in the eight band effective model.

	Bi_2Se_3	Bi_2Te_3	Sb_2Te_3
P_1 (eV Å)	3.33	2.87	3.40
Q_1 (eV Å)	2.26	0.30	0.84
P_2 (eV Å)	2.84	2.68	3.19
Q_2 (eV Å)	2.84	2.68	3.19
P_3 (eV Å)	−2.62	−1.94	−2.46
Q_3 (eV Å)	2.48	1.23	2.11
F_1 (eV Å^2)	3.73	7.16	3.82
K_1 (eV Å^2)	6.52	3.72	2.49
F_3 (eV Å^2)	−1.12	3.76	−32.03
K_3 (eV Å^2)	−14.0	−7.70	−59.28
F_5 (eV Å^2)	1.50	−0.62	−2.26
K_5 (eV Å^2)	−3.11	−7.17	−13.00
F_7 (eV Å^2)	2.71	3.77	5.04
K_7 (eV Å^2)	−5.08	22.27	2.40
$U_{35}=U_{36}^*$ (eV Å^2)	$-2.31-7.45i$	$-2.21-9.85i$	$-11.31-46.00i$
$V_{35}=-V_{36}^*$ (eV Å^2)	$-1.05-5.98i$	$-2.43-3.53i$	$-4.50-22.80i$
F_{37} (eV Å^2)	2.47	4.39	16.96
K_{37} (eV Å^2)	−8.52	−6.50	−24.17
U_{47} (eV Å^2)	−7.86	−4.29	−45.46
V_{47} (eV Å^2)	$-8.95i$	$-0.83i$	$-17.64i$
$U_{58}=-U_{68}^*$ (eV Å^2)	$-2.31-2.57i$	$-0.24-3.69i$	$-2.01-3.98i$
$V_{58}=V_{68}^*$ (eV Å^2)	$-0.64-4.29i$	$-0.85-6.64i$	$1.28-9.02i$
E_1 (eV)	−0.29	−0.48	−0.22
E_3 (eV)	0.28	0.12	0.22
E_5 (eV)	−0.57	−0.63	−0.88
E_7 (eV)	−0.98	−1.18	−1.51

with

$$f_{i(ij)}(\mathbf{k}) = F_{i(ij)}k_z^2 + K_{i(ij)}k_\parallel^2, \tag{53}$$

$$g_{ij}(\mathbf{k}) = U_{ij}k_z k_+ + V_{ij}k_-^2. \tag{54}$$

The parameters F_{ij}, K_{ij}, U_{ij}, and V_{ij} can now also be determined by perturbation theory. In this Hamiltonian, time-reversal symmetry is already satisfied. Furthermore, R_2 rotation symmetry requires that $U_{35}=U_{36}^*$, $V_{35}=-V_{36}^*$, $U_{58}=-U_{68}^*$, and $V_{58}=V_{68}^*$. The obtained parameters are listed in Table V and the band dispersion is found to agree well with that of *ab initio* calculation, as shown in Fig 9. This demonstrates that the eight band model is suitable to serve as a basis for the quantitative study of the Bi_2Se_3 family of materials.

We would like to make some more remarks about the eight band model. First, in our model Hamiltonian with four bands, the leading term that breaks the in-plane full rotation symmetry down to R_3 symmetry is third order in the perturbation, while in the eight band model, it is second order with the coupling coefficients g_{ij}. These type of terms exist because the states $|P2^-,\tilde{\Gamma}_{4,5}\rangle$ themselves break the rotation symmetry according to expressions (13) and (14). Second, it is interesting to compare the present eight band model with the well-known Kane model for usual III-V or II-VI group semiconductors with zinc-blend structure. In fact there is a one-to-one correspondence between the basis of these two models, in which $|P1_-^+,\pm\frac{1}{2}\rangle$ corresponds to the electron band (Γ_6), $|P2_+^-,\pm\frac{1}{2}\rangle$ and $|P2^-,\Gamma_{4,5}\rangle$ correspond to the light hole and heavy hole bands (Γ_8), respectively, and $|P2_-^-,\pm\frac{1}{2}\rangle$ corresponds to the spin-orbit split-off band (Γ_7). Therefore, from the symmetry point of view, our model here is nothing but an extension of the Kane model to a crystal structure with lower symmetry.

VIII. CONCLUSIONS

To summarize, based on the symmetry properties and the $\mathbf{k}\cdot\mathbf{p}$ perturbation theory, we systematically derived a model Hamiltonian for the 3D TI in the Bi_2Se_3 class of materials. Our model Hamiltonian captures the main low energy physics, such as the inverted band structure and topologically protected surface states. The topological surface states have well-defined spin texture, which can be traced back to the sign of the atomic SOC in these materials. Furthermore, the Landau levels of a z-direction magnetic field for both bulk states and surface states are calculated. The gap of bulk Landau levels is shown to decrease when the magnetic field increases, which may be observed with a magneto-optical spectroscopy. Within the bulk gap, the surface Landau levels appear as discrete peaks for the LDOS, which can be detected by STM. We also analyze the quantitative limitation of our model Hamiltonian with four bands and show that a model Hamiltonian with eight energy bands can describe Bi_2Se_3 family of materials quantitatively, which will be useful in the future comparisons with experiments.

ACKNOWLEDGMENTS

We would like to thank Yulin Chen, Aahron Kapitulnik, Zhixun Shen, and Qikun Xue for helpful discussions. This work is supported by the Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Contract No. DE-AC02-76SF00515 and by the Keck Foundation. C.X.L. acknowledges financial support by the Alexander von Humboldt Foundation of Germany. This work is also supported by the NSF of China, the National Basic Research Program of China (No. 2007CB925000), and the International Science and Technology Cooperation Program of China (No. 2008DFB00170).

APPENDIX A: SYMMETRY PROPERTY OF GROUP D_{3d}^5

As described in the text, the group D_{3d}^5 is generated by a threefold rotation operator R_3, a twofold rotation operator R_2, and an inversion operator P. It has six classes and correspondingly six irreducible representations, $\tilde{\Gamma}_1^\pm$, $\tilde{\Gamma}_2^\pm$ and $\tilde{\Gamma}_3^\pm$ with the upper index $\pm$ denoting the parity of the representation. Here we use $\tilde{\Gamma}$ to denote the representation at Γ point

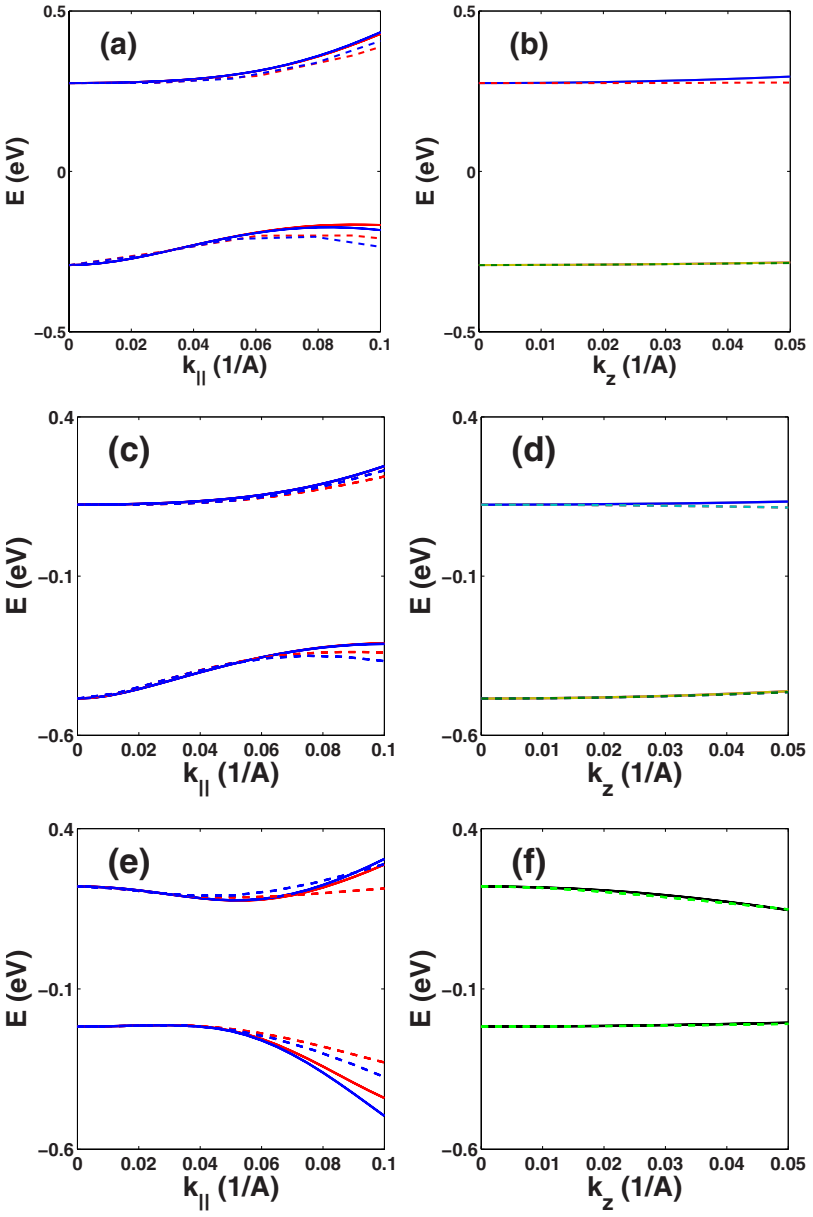

FIG. 9. (Color online) The energy dispersion obtained from the model Hamiltonian with eight bands (solid line) is compared with that from *ab initio* calculation (dashed line) for (a), (c), and (e) k_x and k_y directions and (b), (d), and (f) k_z direction. Here (a) and (b) is for Bi_2Se_3, (c) and (d) is for Bi_2Te_3, while (e) and (f) is for Sb_2Te_3. In (a), (c), and (e), the red line represents the dispersion along the k_x direction while the blue line is for the k_y direction.

in the BZ to avoid confusion with the Dirac Γ matrices. The character table of D_{3d}^5 is given in Table I.[39]

After taking into account the spin, the $\mathcal{C}=2\pi$ rotation induces a minus sign for the spin part so that the number of elements in the group is doubled, which is the so-called double group. For D_{3d}^5, the classes and irreducible representations of the double group are also doubled. The character table for the double group of D_{3d}^5 is given in Table II.[39]

When constructing the double group, it is useful to consider the decomposition of the direct product of $\tilde{\Gamma}_{1,2,3}^{\pm}$ and spinor representation $\tilde{\Gamma}_6$, which is given by

$$\tilde{\Gamma}_3^{\pm} \otimes \tilde{\Gamma}_6^{+} = \tilde{\Gamma}_4^{\pm} + \tilde{\Gamma}_5^{\pm} + \tilde{\Gamma}_6^{\pm}, \tag{A1}$$

$$\tilde{\Gamma}_1^{\pm} \otimes \tilde{\Gamma}_6^{+} = \tilde{\Gamma}_6^{\pm}, \tag{A2}$$

$$\tilde{\Gamma}_2^{\pm} \otimes \tilde{\Gamma}_6^{+} = \tilde{\Gamma}_6^{\pm}. \tag{A3}$$

Furthermore when considering the matrix elements of $\mathbf{k}\cdot\mathbf{p}$ theory, the following direct products will be helpful:

$$(\tilde{\Gamma}_6^{\pm})^* \otimes \tilde{\Gamma}_6^{\pm} = \tilde{\Gamma}_1^{+} + \tilde{\Gamma}_2^{+} + \tilde{\Gamma}_3^{+}, \tag{A4}$$

$$(\tilde{\Gamma}_6^{+})^* \otimes \tilde{\Gamma}_6^{-} = \tilde{\Gamma}_1^{-} + \tilde{\Gamma}_2^{-} + \tilde{\Gamma}_3^{-}, \tag{A5}$$

$$(\tilde{\Gamma}_6^{+})^* \otimes \tilde{\Gamma}_4^{\pm} = \tilde{\Gamma}_3^{\pm}, \tag{A6}$$

$$(\tilde{\Gamma}_6^{-})^* \otimes \tilde{\Gamma}_4^{\pm} = \tilde{\Gamma}_3^{\mp}, \tag{A7}$$

$$(\tilde{\Gamma}_6^{+})^* \otimes \tilde{\Gamma}_5^{\pm} = \tilde{\Gamma}_3^{\pm}, \tag{A8}$$

$$(\tilde{\Gamma}_6^{-})^* \otimes \tilde{\Gamma}_5^{\pm} = \tilde{\Gamma}_3^{\mp}, \tag{A9}$$

$$(\tilde{\Gamma}_{4(5)}^{+})^* \otimes \tilde{\Gamma}_{4(5)}^{-} = \tilde{\Gamma}_1^{-}, \tag{A10}$$

$$(\tilde{\Gamma}^{+}_{4(5)})^{*} \otimes \tilde{\Gamma}^{-}_{5(4)} = \tilde{\Gamma}^{-}_{2}. \tag{A11}$$

APPENDIX B: Γ MATRIX

The five Dirac Γ matrices can be defined as

$$\Gamma_1 = \sigma_1 \otimes \tau_1, \quad \Gamma_2 = \sigma_2 \otimes \tau_1, \quad \Gamma_3 = \sigma_3 \otimes \tau_1,$$

$$\Gamma_4 = 1 \otimes \tau_2, \quad \Gamma_5 = 1 \otimes \tau_3, \tag{B1}$$

which satisfies Clifford algebra $\{\Gamma_a, \Gamma_b\} = 2\delta_{ab}$. The other ten Γ matrices are given by $\Gamma_{ab} = [\Gamma_a, \Gamma_b]/2i$. Explicitly, Γ_{ab} is given by

$$\Gamma_{ij} = [\sigma_i \otimes \tau_1, \sigma_j \otimes \tau_1]/2i = \epsilon_{ijk}\sigma_k \otimes 1, \tag{B2}$$

$$\Gamma_{i4} = [\sigma_i \otimes \tau_1, 1 \otimes \tau_2]/2i = \sigma_i \otimes \tau_3, \tag{B3}$$

$$\Gamma_{i5} = [\sigma_i \otimes \tau_1, 1 \otimes \tau_3]/2i = -\sigma_i \otimes \tau_2, \tag{B4}$$

$$\Gamma_{45} = [1 \otimes \tau_2, 1 \otimes \tau_3]/2i = 1 \otimes \tau_1, \tag{B5}$$

where $i,j = 1,2,3$. Now let us check the properties of the fifteen Γ matrices under the time reversal T and inversion P. We assume the Γ matrices are written in the basis $|P1^{+}_{-}, \frac{1}{2}\rangle$, $|P2^{-}_{+}, \frac{1}{2}\rangle$, $|P1^{+}_{-}, -\frac{1}{2}\rangle$, and $|P2^{-}_{+}, -\frac{1}{2}\rangle$, then the transformation matrix of the symmetry operation has been obtained in Sec. III. With these transformation matrices, we have

$$T\Gamma_i T^{-1} = P\Gamma_i P^{-1} = -\Gamma_i, \quad i = 1,2,3,4, \tag{B6}$$

$$T\Gamma_5 T^{-1} = P\Gamma_5 P^{-1} = \Gamma_5. \tag{B7}$$

In fact P operator is exactly Γ_5 here,

$$T\Gamma_{ij}T^{-1} = -P\Gamma_{ij}P^{-1} = -\Gamma_{ij}, \tag{B8}$$

$$T\Gamma_{i4}T^{-1} = -P\Gamma_{i4}P^{-1} = -\Gamma_{i4}, \tag{B9}$$

$$T\Gamma_{i5}T^{-1} = -P\Gamma_{i5}P^{-1} = \Gamma_{i5}, \tag{B10}$$

$$T\Gamma_{45}T^{-1} = -P\Gamma_{45}P^{-1} = \Gamma_{45}, \tag{B11}$$

where $i,j = 1,2,3$.

Next let us consider R_2,

$$R_2\Gamma_{1,4}R_2^{-1} = -\Gamma_{1,4}, \tag{B12}$$

$$R_2\Gamma_{2,3,5}R_2^{-1} = \Gamma_{2,3,5} \tag{B13}$$

$$R_2\Gamma_{12,31,24,34,15,45}R_2^{-1} = -\Gamma_{12,31,24,34,15,45}, \tag{B14}$$

$$R_2\Gamma_{23,14,25,35}R_2^{-1} = \Gamma_{23,14,25,35}. \tag{B15}$$

Finally let us talk about the three fold rotation symmetry. Under the rotation operation $R_z(\theta)$, the Γ matrices are transformed as $\Gamma'(\theta) = e^{i\Sigma/2\theta}\Gamma e^{-i\Sigma/2\theta}$, then

$$\frac{d\Gamma'(\theta)}{d\theta} = \frac{i}{2}[\Sigma, \Gamma'(\theta)]. \tag{B16}$$

Therefore, the transformation properties of Γ matrices under the rotation operation are determined by the commutation relation $[\Sigma, \Gamma]$. The commutation relations for Γ matrices are listed as follows:

$$[\Sigma,\Gamma_1] = 2i\Gamma_2, \quad [\Sigma,\Gamma_2] = -2i\Gamma_1,$$

$$[\Sigma,\Gamma_3] = [\Sigma,\Gamma_4] = [\Sigma,\Gamma_5] = 0, \tag{B17}$$

$$[\Sigma,\Gamma_{12}] = 0, \quad [\Sigma,\Gamma_{34}] = 0, \tag{B18}$$

$$[\Sigma,\Gamma_{31}] = -2i\Gamma_{23}, \quad [\Sigma,\Gamma_{23}] = 2i\Gamma_{31}, \tag{B19}$$

$$[\Sigma,\Gamma_{14}] = 2i\Gamma_{24}, \quad [\Sigma,\Gamma_{24}] = -2i\Gamma_{14}, \tag{B20}$$

$$[\Sigma,\Gamma_{15}] = 2i\Gamma_{25}, \quad [\Sigma,\Gamma_{25}] = -2i\Gamma_{15}, \tag{B21}$$

$$[\Sigma,\Gamma_{35}] = 0, \quad [\Sigma,\Gamma_{45}] = 0. \tag{B22}$$

With the above commutation relations, we can easily solve Eq. (B16) and find that

$$\Gamma'_1(\theta) = \Gamma_1 \cos\theta - \Gamma_2 \sin\theta,$$

$$\Gamma'_2(\theta) = \Gamma_1 \sin\theta + \Gamma_2 \cos\theta, \tag{B23}$$

$$\Gamma'_3(\theta) = \Gamma_3, \quad \Gamma'_4(\theta) = \Gamma_4, \tag{B24}$$

$$\Gamma'_{23}(\theta) = \Gamma_{23} \cos\theta - \Gamma_{31} \sin\theta,$$

$$\Gamma'_{31}(\theta) = \Gamma_{31} \cos\theta + \Gamma_{23} \sin\theta$$

$$\Gamma'_{14}(\theta) = \Gamma_{14} \cos\theta - \Gamma_{24} \sin\theta, \tag{B25}$$

$$\Gamma'_{24}(\theta) = \Gamma_{14} \sin\theta + \Gamma_{24} \cos\theta, \tag{B26}$$

$$\Gamma'_{15}(\theta) = \Gamma_{15} \cos\theta - \Gamma_{25} \sin\theta,$$

$$\Gamma'_{25}(\theta) = \Gamma_{15} \sin\theta + \Gamma_{25} \cos\theta,$$

$$\Gamma'_5(\theta) = \Gamma_5, \quad \Gamma'_{34} = \Gamma_{34}, \quad \Gamma'_{12} = \Gamma_{12}, \tag{B27}$$

$$\Gamma'_{35} = \Gamma_{35}, \quad \Gamma'_{45} = \Gamma_{45}. \tag{B28}$$

The above results indicate that under the rotation $R_3\Gamma_{3,4,5}$ and $\Gamma_{12,34,35,45}$ behave as scalars (or pseudoscalars), while the three pairs of operators $\{\Gamma_{23},\Gamma_{31}\}$, $\{\Gamma_{14},\Gamma_{24}\}$, and $\{\Gamma_{15},\Gamma_{25}\}$ behave as vectors. The corresponding representation for each Γ matrix is given in table III.

APPENDIX C: PARAMETERS IN k·p THEORY

In this appendix, we show the detailed results from **k·p** theory. First let us consider the constraint for the matrix elements of momentum from the D^5_{3d} symmetry. As described above, the eigenstates can be denoted by $|\Lambda^{\pm}, \alpha\rangle$ with $\Lambda = P1_{\pm}, P2_{\pm}$ and $\alpha = \pm\frac{1}{2}, \pm\frac{3}{2}$. The states $|\Lambda, \pm 1/2\rangle$ belong to $\tilde{\Gamma}^{\pm}_6$ representation. For $|\Lambda, \pm 3/2\rangle$, as described above, we need to recombine these two states as

$$|\Lambda^{\pm},\tilde{\Gamma}_4\rangle = \frac{1}{\sqrt{2}}(|\Lambda^{\pm},3/2\rangle + |\Lambda^{\pm},-3/2\rangle), \tag{C1}$$

$$|\Lambda^{\pm},\tilde{\Gamma}_5\rangle = \frac{1}{\sqrt{2}}(|\Lambda^{\pm},3/2\rangle - |\Lambda^{\pm},-3/2\rangle), \tag{C2}$$

which belong to $\tilde{\Gamma}_4$ and $\tilde{\Gamma}_5$ representations, respectively. Expressions (A4)–(A11) give the decomposition of the direct product of these states. The momentum p_x, p_y belongs to the $\tilde{\Gamma}_3^-$ representation, while p_z belongs to the $\tilde{\Gamma}_2^-$ representation; therefore we require that the decomposition of the direct product of the eigenstates also include $\tilde{\Gamma}_3^-$ and $\tilde{\Gamma}_2^-$ to obtain nonzero matrix elements. For example, the direct product of $\tilde{\Gamma}_6^{\pm}$ and $\tilde{\Gamma}_{4,5}^{\pm}$ does not contain $\tilde{\Gamma}_2$, which indicates that the matrix element $\langle \Lambda_1, \pm 1/2|p_z|\Lambda_2,\tilde{\Gamma}_{4,5}\rangle$ is always zero.

The symmetry operation can further help us to obtain the relation between different matrix elements of the momentum. For example, due to the R_3 rotation symmetry, we have

$$\begin{aligned}\left\langle \Lambda_1^+,\frac{1}{2}\middle|p_x\middle|\Lambda_2^-,-\frac{1}{2}\right\rangle &= \left\langle \Lambda_1^+,\frac{1}{2}\middle|R_3^{\dagger}R_3p_xR_3^{\dagger}R_3\middle|\Lambda_2^-,-\frac{1}{2}\right\rangle \\ &= e^{-i(2\pi/3)}\left\langle \Lambda_1^+,\frac{1}{2}\middle|\left(p_x\cos\frac{2\pi}{3}\right.\right. \\ &\quad \left.\left. - p_y\sin\frac{2\pi}{3}\right)\middle|\Lambda_2^-,-\frac{1}{2}\right\rangle \\ &\rightarrow \left\langle \Lambda_1^+,\frac{1}{2}\middle|p_x\middle|\Lambda_2^-,-\frac{1}{2}\right\rangle \\ &= i\left\langle \Lambda_1^+,\frac{1}{2}\middle|p_y\middle|\Lambda_2^-,-\frac{1}{2}\right\rangle. \end{aligned} \tag{C3}$$

Finally we can define the independent components of the matrix elements as follows:

$$\begin{aligned}\left\langle \Lambda_1^+,\frac{1}{2}\middle|p_x\middle|\Lambda_2^-,-\frac{1}{2}\right\rangle &= \left\langle \Lambda_1^+,-\frac{1}{2}\middle|p_x\middle|\Lambda_2^-,\frac{1}{2}\right\rangle \\ &= i\left\langle \Lambda_1^+,\frac{1}{2}\middle|p_y\middle|\Lambda_2^-,-\frac{1}{2}\right\rangle \\ &= -i\left\langle \Lambda_1^+,-\frac{1}{2}\middle|p_y\middle|\Lambda_2^-,\frac{1}{2}\right\rangle = P_{\Lambda_1^+,\Lambda_2^-}, \end{aligned} \tag{C4}$$

$$\left\langle \Lambda_1^+,\frac{1}{2}\middle|p_z\middle|\Lambda^-,\frac{1}{2}\right\rangle = -\left\langle \Lambda_1^+,-\frac{1}{2}\middle|p_z\middle|\Lambda^-,-\frac{1}{2}\right\rangle = Q_{\Lambda_1^+,\Lambda_2^-}, \tag{C5}$$

$$\begin{aligned}\left\langle \Lambda_1^{\pm},\frac{1}{2}\middle|p_x|\Lambda_2^{\mp},\tilde{\Gamma}_4\right\rangle &= -i\left\langle \Lambda_1^{\pm},-\frac{1}{2}\middle|p_x|\Lambda_2^{\mp},\tilde{\Gamma}_4\right\rangle \\ &= -i\left\langle \Lambda_1^{\pm},\frac{1}{2}\middle|p_y|\Lambda_2^{\mp},\tilde{\Gamma}_4\right\rangle \\ &= \left\langle \Lambda_1^{\pm},-\frac{1}{2}\middle|p_y|\Lambda_2^{\mp},\tilde{\Gamma}_4\right\rangle = M_{\Lambda_1^{\pm},\Lambda_2^{\mp}}, \end{aligned} \tag{C6}$$

$$\begin{aligned}\left\langle \Lambda_1^{\pm},\frac{1}{2}\middle|p_x|\Lambda_2^{\mp},\tilde{\Gamma}_5\right\rangle &= i\left\langle \Lambda_1^{\pm},-\frac{1}{2}\middle|p_x|\Lambda_2^{\mp},\tilde{\Gamma}_5\right\rangle \\ &= -i\left\langle \Lambda_1^{\pm},\frac{1}{2}\middle|p_y|\Lambda_2^{\mp},\tilde{\Gamma}_5\right\rangle \\ &= -\left\langle \Lambda_1^{\pm},-\frac{1}{2}\middle|p_y|\Lambda_2^{\mp},\tilde{\Gamma}_5\right\rangle = N_{\Lambda_1^{\pm},\Lambda_2^{\mp}}, \end{aligned} \tag{C7}$$

$$\langle \Lambda_1^+,\tilde{\Gamma}_4|p_z|\Lambda^-,\tilde{\Gamma}_4\rangle = R_{\Lambda_1^+,\Lambda_2^-}, \tag{C8}$$

$$\langle \Lambda_1^+,\tilde{\Gamma}_5|p_z|\Lambda^-,\tilde{\Gamma}_5\rangle = S_{\Lambda_1^+,\Lambda_2^-}. \tag{C9}$$

Here it is more convenient to use $p_{\pm}=p_x \pm ip_y$, which leads to

$$\left\langle \Lambda_1^+,\frac{1}{2}\middle|p_+\middle|\Lambda_2^-,-\frac{1}{2}\right\rangle = \left\langle \Lambda_1^+,-\frac{1}{2}\middle|p_-\middle|\Lambda_2^-,\frac{1}{2}\right\rangle = 2P_{\Lambda_1^+,\Lambda_2^-}, \tag{C10}$$

$$\left\langle \Lambda_1^+,\frac{1}{2}\middle|p_-\middle|\Lambda_2^-,-\frac{1}{2}\right\rangle = \left\langle \Lambda_1^+,-\frac{1}{2}\middle|p_+\middle|\Lambda_2^-,\frac{1}{2}\right\rangle = 0, \tag{C11}$$

$$\left\langle \Lambda_1^{\pm},\frac{1}{2}\middle|p_-|\Lambda_2^{\mp},\Gamma_4\right\rangle = -i\left\langle \Lambda_1^{\pm},-\frac{1}{2}\middle|p_+|\Lambda_2^{\mp},\Gamma_4\right\rangle = 2M_{\Lambda_1^{\pm},\Lambda_2^{\mp}}, \tag{C12}$$

$$\left\langle \Lambda_1^{\pm},\frac{1}{2}\middle|p_+|\Lambda_2^{\mp},\Gamma_4\right\rangle = -i\left\langle \Lambda_1^{\pm},-\frac{1}{2}\middle|p_-|\Lambda_2^{\mp},\Gamma_4\right\rangle = 0, \tag{C13}$$

$$\left\langle \Lambda_1^{\pm},\frac{1}{2}\middle|p_-|\Lambda_2^{\mp},\Gamma_5\right\rangle = i\left\langle \Lambda_1^{\pm},-\frac{1}{2}\middle|p_+|\Lambda_2^{\mp},\Gamma_5\right\rangle = 2N_{\Lambda_1^{\pm},\Lambda_2^{\mp}}, \tag{C14}$$

$$\left\langle \Lambda_1^{\pm},\frac{1}{2}\middle|p_+|\Lambda_2^{\mp},\Gamma_5\right\rangle = i\left\langle \Lambda_1^{\pm},-\frac{1}{2}\middle|p_-|\Lambda_2^{\mp},\Gamma_5\right\rangle = 0. \tag{C15}$$

Time-reversal symmetry indicates that $P_{\Lambda_1^+,\Lambda_2^-}$, $Q_{\Lambda_1^+,\Lambda_2^-}$ can be chosen to be real ($P_{\Lambda_1^+,\Lambda_2^-}=P^*_{\Lambda_1^+,\Lambda_2^-}$, $Q_{\Lambda_1^+,\Lambda_2^-}=Q^*_{\Lambda_1^+,\Lambda_2^-}$) while $M_{\Lambda_1^{\pm},\Lambda_2^{\mp}}=iN^*_{\Lambda_1^{\pm},\Lambda_2^{\mp}}$ and $R_{\Lambda_1^+,\Lambda_2^-}=-S^*_{\Lambda_1^+,\Lambda_2^-}$. Since the matrix ele-

ment between $|P1_-^+, \pm\frac{1}{2}\rangle$ and $|P2_+^-, \pm\frac{1}{2}\rangle$ is quite important, we denote

$$\left\langle P1_-^+, \frac{1}{2}\middle| p_x \middle| P2_+^-, -\frac{1}{2}\right\rangle = \left\langle P1_-^+, -\frac{1}{2}\middle| p_x \middle| P2_+^-, \frac{1}{2}\right\rangle = i\left\langle P1_-^+, \frac{1}{2}\middle| p_y \middle| P2_+^-, -\frac{1}{2}\right\rangle = -i\left\langle P1_-^+, -\frac{1}{2}\middle| p_x \middle| P2_+^-, \frac{1}{2}\right\rangle = P_0, \tag{C16}$$

$$\left\langle P1_-^+, \frac{1}{2}\middle| p_z \middle| P2_+^-, \frac{1}{2}\right\rangle = -\left\langle P1_-^+, -\frac{1}{2}\middle| p_z \middle| P2_+^-, -\frac{1}{2}\right\rangle = Q_0. \tag{C17}$$

Now we consider perturbation theory. The degenerate perturbation formalism is given by

$$H_{mm'}^{(0)} = E_m \delta_{mm'}, \tag{C18}$$

$$H_{mm'}^{(1)} = H'_{mm'}, \tag{C19}$$

$$H_{mm'}^{(2)} = \frac{1}{2}\sum_l H'_{ml}H'_{lm'}\left(\frac{1}{E_m - E_l} + \frac{1}{E_{m'} - E_l}\right), \tag{C20}$$

$$H_{mm'}^{(3)} = -\frac{1}{2}\sum_{l,m''}\left[\frac{H'_{ml}H'_{lm''}H'_{m''m'}}{(E_{m'} - E_l)(E_{m''} - E_l)} + \frac{H'_{mm''}H'_{m''l}H'_{lm'}}{(E_m - E_l)(E_{m''} - E_l)}\right] + \frac{1}{2}\sum_{l,l'} H'_{ml}H'_{ll'}H'_{l'm'}\left[\frac{1}{(E_m - E_l)(E_m - E_{l'})} + \frac{1}{(E_{m'} - E_l)(E_{m'} - E_{l'})}\right]. \tag{C21}$$

Here m and m' are taken from $|P1_-^+, 1/2\rangle = |1\rangle$, $|P2_+^-, 1/2\rangle = |2\rangle$, $|P1_-^+, -1/2\rangle = |3\rangle$ and $|P2_+^-, -1/2\rangle = |4\rangle$ with the energy $E_1 = E_3$ and $E_2 = E_4$ and $E_1 < E_2$. l is taken from the other bands except for these four bands. The expression from the perturbative calculation of our model Hamiltonian with four bands is given as follows and the values of the parameters are listed in Table IV. For our model Hamiltonian with eight bands, the perturbation procedure is the same as our model Hamiltonian with four bands and here we only list the values of the parameters in Table V.

$$C_0 + M_0 = E_1, \tag{C22}$$

$$C_0 - M_0 = E_2, \tag{C23}$$

$$C_1 + M_1 = \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2}\sum_{\Lambda^-}\frac{|Q_{P1^+,\Lambda^-}|^2}{E_1 - E_{\Lambda^-,1/2}}, \tag{C24}$$

$$C_2 + M_2 = \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2}\sum_{\Lambda^-}\left(\frac{|P_{P1^+,\Lambda^-}|^2}{E_1 - E_{\Lambda^-,-1/2}} + \frac{|M_{P1^+,\Lambda^-}|^2}{E_1 - E_{\Lambda^-,\Gamma_4}} + \frac{|N_{P1^+,\Lambda^-}|^2}{E_1 - E_{\Lambda^-,\Gamma_5}}\right), \tag{C25}$$

$$C_1 - M_1 = \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2}\sum_{\Lambda^-}\frac{|Q_{\Lambda^+,P2^-}|^2}{E_1 - E_{\Lambda^+,1/2}}, \tag{C26}$$

$$C_2 - M_2 = \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2}\sum_{\Lambda^-}\left(\frac{|P_{\Lambda^+,P2^-}|^2}{E_1 - E_{\Lambda^+,-1/2}} + \frac{|M_{\Lambda^+,P2^-}|^2}{E_1 - E_{\Lambda^+,\Gamma_4}} + \frac{|N_{\Lambda^+,P2^-}|^2}{E_1 - E_{\Lambda^+,\Gamma_5}}\right), \tag{C27}$$

$$A_0 = \frac{\hbar}{m_0}P_0, \tag{C28}$$

$$B_0 = \frac{\hbar}{m_0}Q_0. \tag{C29}$$

For R_1 and R_2 term we have

$$R_1 - R_2 = \frac{\hbar^3}{m_0^3}\left[\sum_{\Lambda}\frac{|M_{P1^+,\Lambda^-}|^2 P_{P1^+,P2^-}}{(E_{P2^-} - E_{\Lambda^-})(E_{P1^+} - E_{\Lambda^-})} - \sum_{\Lambda_1^-,\Lambda_2^+} M_{P1^+\Lambda_1^-}M^*_{\Lambda_2^+\Lambda_1^-}P_{\Lambda_2^+P2^-} \times\left(\frac{1}{(E_{P1^+} - E_{\Lambda_1^-})(E_{P1^+} - E_{\Lambda_2^+})} + \frac{1}{(E_{P2^-} - E_{\Lambda_1^-})(E_{P2^-} - E_{\Lambda_2^+})}\right)\right], \tag{C30}$$

$$R_1 + R_2 = \frac{\hbar^3}{m_0^3}\left[-\sum_{\Lambda}\frac{P_{P1^+,P2^-}|M_{P2^-,\Lambda^-}|^2}{(E_{P2^-} - E_{\Lambda^+})(E_{P1^+} - E_{\Lambda^+})} - \sum_{\Lambda_1^-,\Lambda_2^+} P_{P1^+\Lambda_1^-}M_{\Lambda_1^-\Lambda_2^+}M^*_{P2^-\Lambda_2^+} \times\left(\frac{1}{(E_{P1^+} - E_{\Lambda_1^-})(E_{P1^+} - E_{\Lambda_2^+})} + \frac{1}{(E_{P2^-} - E_{\Lambda_1^-})(E_{P2^-} - E_{\Lambda_2^+})}\right)\right]. \tag{C31}$$

Now we study the effect of a magnetic field. Under a magnetic field, there are two different kinds of contribution. One is the orbital term, which induces the Landau levels and has been considered in Sec. VI. The other one is the Zeeman type term, which is described by an effective g factor. In the following we will discuss the effective g factor in detail. There are two kinds of contributions to the effective g factor. One comes from the atomic g factor, which can be estimated

from the *ab initio* calculation. In an atom, the electron spin and orbital angular momentum couples to the magnetic field as $\hat{H}_{Zee}=\frac{\mu_B}{\hbar}(g_l\mathbf{L}+g_s\mathbf{S})\cdot\mathbf{B}=\frac{\mu_B}{\hbar}g_0\mathbf{J}\cdot\mathbf{B}$, where $\mathbf{J}=\mathbf{S}+\mathbf{L}$ is the total angular momentum and g_0 is so called Landé g-factor. The wave functions for the basis of our model Hamiltonian have been calculated from *ab initio* calculation, which can be projected onto the atomic orbitals. Since for each atomic orbitals, the g factor is simply given by $g_0=1+\frac{J(J+1)-L(L+1)+S(S+1)}{2J(J+1)}$, the effective g_0 can be easily calculated, and is found to be $g_0\approx 1.2$. Another contribution to the effective g factor originates from the second order perturbation, and is related to the correction to the effective mass term. The relation between the effective mass and effective g factor in the ordinary semiconductors is known as the Roth's formula.[47] Here the second-order correction to the g factor is given by

$$g_{1z}^{(2)}=\frac{4}{m_0}\sum_{\Lambda^-,\alpha}\left(\frac{|P_{P1^+,\Lambda^-}|^2}{E_1-E_{\Lambda^-,-1/2}}-\frac{|M_{P1^+,\Lambda^-}|^2}{E_1-E_{\Lambda^-,\Gamma_4}}-\frac{|N_{P1^+,\Lambda^-}|^2}{E_1-E_{\Lambda^-,\Gamma_5}}\right), \tag{C32}$$

$$g_{1p}^{(2)}=\frac{4}{m_0}\sum_{\Lambda^-}\frac{Q_{P1^+,\Lambda^-}P^*_{P1^+,\Lambda^-}}{E_{P1^+}-E_{\Lambda^-,1/2}}, \tag{C33}$$

$$g_{2z}^{(2)}=\frac{4}{m_0}\sum_{\Lambda^+,\alpha}\left(\frac{|P_{\Lambda^+,P2^-}|^2}{E_2-E_{\Lambda^+,-1/2}}-\frac{|M_{\Lambda^+,P2^-}|^2}{E_2-E_{\Lambda^+,\Gamma_4}}-\frac{|N_{\Lambda^+,P2^-}|^2}{E_2-E_{\Lambda^+,\Gamma_5}}\right), \tag{C34}$$

$$g_{2p}^{(2)}=\frac{4}{m_0}\sum_{\Lambda^+}\frac{Q^*_{\Lambda^+,P2^-}P_{\Lambda^+,P2^-}}{E_{P2^-}-E_{\Lambda^+}}. \tag{C35}$$

where $g_{1z(p)}$ and $g_{2z(p)}$ are defined in Eq. (45). Therefore our effective g factor is the summation of the above two different contributions,

$$g_\alpha=g_0+g_\alpha^{(2)},\quad \alpha=1z,2z,1p,2p, \tag{C36}$$

and the values of the effective g factor are given in Table IV. From Table IV, we find that for the $|P1^+_-,\pm\frac{1}{2}\rangle$ band, there is a strong anisotropy which comes from the large contribution of the second order perturbative correction of the states $|P2^-,\pm\frac{3}{2}\rangle$ and $|P2^-_-,\pm\frac{1}{2}\rangle$.

[1] H. Zhang, C. Liu, X. Qi, X. Dai, Z. Fang, and S. Zhang, Nat. Phys. **5**, 438 (2009).
[2] X. Qi and S. Zhang, Phys. Today **63** (1), 33 (2010).
[3] J. E. Moore, Nature (London) **464**, 194 (2010).
[4] M. Hasan and C. Kane, arXiv:1002.3895 (unpublished).
[5] B. A. Bernevig, T. L. Hughes, and S. C. Zhang, Science **314**, 1757 (2006).
[6] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).
[7] A. Roth, C. Brune, H. Buhmann, L. W. Molenkamp, J. Maciejko, X. Qi, and S. Zhang, Science **325**, 294 (2009).
[8] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).
[9] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **452**, 970 (2008).
[10] Y. Xia *et al.*, Nat. Phys. **5**, 398 (2009).
[11] Y. L. Chen *et al.*, Science **325**, 178 (2009).
[12] D. Hsieh *et al.*, Phys. Rev. Lett. **103**, 146401 (2009).
[13] Y. S. Hor, A. J. Williams, J. G. Checkelsky, P. Roushan, J. Seo, Q. Xu, H. W. Zandbergen, A. Yazdani, N. P. Ong, and R. J. Cava, Phys. Rev. Lett. **104**, 057001 (2010).
[14] Z. Alpichshev, J. G. Analytis, J. Chu, I. R. Fisher, Y. L. Chen, Z. X. Shen, A. Fang, and A. Kapitulnik, Phys. Rev. Lett. **104**, 016401 (2010).
[15] D. Kong, J. C. Randel, H. Peng, J. J. Cha, S. Meister, K. Lai, Y. Chen, Z. Shen, H. C. Manoharan, and Y. Cui, Nano Lett. **10**, 329 (2010).
[16] J. J. Cha, J. R. Williams, D. Kong, S. Meister, H. Peng, A. J. Bestwick, P. Gallagher, D. Goldhaber-Gordon, and Y. Cui, Nano Lett. **10**, 1076 (2010).
[17] H. Peng, K. Lai, D. Kong, S. Meister, Y. Chen, X. Qi, S. Zhang, Z. Shen, and Y. Cui, Nature Mater. **9**, 225 (2010).
[18] D. Hsieh *et al.*, Nature (London) **460**, 1101 (2009).
[19] 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).
[20] T. Zhang *et al.*, Phys. Rev. Lett. **103**, 266803 (2009).
[21] Y. Zhang, K. He, C. Chang, C. Song, L. Wang, X. Chen, J. Jia, Z. Fang, X. Dai, W. Shan, S. Shen, Q. Niu, X. Qi, S. Zhang, X. Ma, and Q. Xue, arXiv:0911.3706, Nat. Phys. (to be published).
[22] P. Cheng, C. Song, T. Zhang, Y. Zhang, Y. Wang, J. Jia, J. Wang, Y. Wang, B. Zhu, X. Chen, X. Ma, K. He, L. Wang, X. Dai, Z. Fang, X. Xie, X. Qi, C. Liu, S. Zhang, and Q. Xue, arXiv:1001.3220 (unpublished).
[23] T. Hanaguri, K. Igarashi, M. Kawamura, H. Takagi, and T. Sasagawa, arXiv:1003.0100 (unpublished).
[24] J. Chen, H. Qin, F. Yang, J. Liu, T. Guan, F. Qu, G. Zhang, J. Shi, X. Xie, C. Yang, K. Wu, Y. Li, and L. Lu, arXiv:1003.1534 (unpublished).
[25] N. Butch, K. Kirshenbaum, P. Syers, A. Sushkov, G. Jenkins, H. Drew, and J. Paglione, Phys. Rev. B **81**, 241301(R) (2010).
[26] D. Teweldebrhan, V. Goyal, and A. A. Balandin, Nano Lett. **10**, 1209 (2010).
[27] H. Steinberg, D. Gardner, Y. Lee, and P. Jarillo-Herrero, arXiv:1003.3137 (unpublished).
[28] H. Tang, D. Liang, R. Qiu, and X. Gao, arXiv:1003.6099 (unpublished).
[29] O. E. Ayala-Valenzuela, J. G. Analytis, J. Chu, M. M. Altarawneh, I. R. Fisher, and R. D. McDonald, arXiv:1004.2311 (unpublished).
[30] Z. Alpichshev, J. Analytis, J. Chu, I. Fisher, and A. Kapitulnik, arXiv:1003.2233 (unpublished).
[31] M. König, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, J. Phys. Soc. Jpn. **77**, 031007

(2008).
[32] J. Linder, T. Yokoyama, and A. Sudbo, Phys. Rev. B **80**, 205401 (2009).
[33] C. X. Liu, H. J. Zhang, B. Yan, X. L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S. C.Zhang, Phys. Rev. B **81**, 041307 (2010).
[34] H. Z. Lu, W. Y. Shan, W. Yao, Q. Niu, and S. Q. Shen, Phys. Rev. B **81**, 115407 (2010).
[35] L. Fu, Phys. Rev. Lett. **103**, 266801 (2009).
[36] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Phys. Rev. B **78**, 195424 (2008).
[37] X. Qi, R. Li, J. Zang, and S. Zhang, Science **323**, 1184 (2009).
[38] W. C. Lee, C. Wu, D. P. Arovas, and S. C. Zhang, Phys. Rev. B **80**, 245439 (2009).
[39] A. J. Mildred, S. Dresselhaus, and Gene Dresselhaus, *Group Theory: Application to the Physics of Condensed Matter* (Springer, Berlin, 2008).
[40] R. Winkler, *Spin-orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*, Springer Tracts in Modern Physics (Springer, Berlin, 2003).
[41] W. Shan, H. Lu, and S. Shen, New J. Phys. **12**, 043048 (2010).
[42] B. Zhou, H. Z. Lu, R. L. Chu, S. Q. Shen, and Q. Niu, Phys. Rev. Lett. **101**, 246807 (2008).
[43] Y. A. Bychkov and E. I. Rashba, J. Phys. C **17**, 6039 (1984).
[44] W. Zhang, R. Yu, H. Zhang, X. Dai, and Z. Fang, New J. Phys. **12**, 065013 (2010).
[45] H. J. Zhang, C. X. Liu, X. L. Qi, X. Y. Deng, X. Dai, S. C. Zhang, and Z. Fang, Phys. Rev. B **80**, 085307 (2009).
[46] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).
[47] L. M. Roth, B. Lax, and S. Zwerdling, Phys. Rev. **114**, 90 (1959).