OPEN ACCESS

First-principles studies of the three-dimensional strong topological insulators Bi_2Te_3, Bi_2Se_3 and Sb_2Te_3

New Journal of Physics
The open-access journal for physics

First-principles studies of the three-dimensional strong topological insulators Bi_2Te_3, Bi_2Se_3 and Sb_2Te_3

Wei Zhang, Rui Yu, Hai-Jun Zhang, Xi Dai and Zhong Fang
Beijing National Laboratory for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, People's Republic of China
E-mail: zfang@iphy.ac.cn

New Journal of Physics **12** (2010) 065013 (14pp)
Received 23 March 2010
Published 17 June 2010
Online at http://www.njp.org/
doi:10.1088/1367-2630/12/6/065013

Abstract. Bi_2Se_3, Bi_2Te_3 and Sb_2Te_3 compounds have recently been predicted to be three-dimensional (3D) strong topological insulators. In this paper, based on *ab initio* calculations, we study in detail the topological nature and the surface states of this family of compounds. The penetration depth and the spin-resolved Fermi surfaces of the surface states are analyzed. We also present a procedure from which a highly accurate effective Hamiltonian can be constructed based on projected atomic Wannier functions (which keep the symmetries of the systems). Such a Hamiltonian can be used to study the semi-infinite systems or slab-type supercells efficiently. Finally, we discuss the 3D topological phase transition in the $Sb_2(Te_{1-x}Se_x)_3$ alloy system.

Contents

1. Introduction

Topological insulators (TIs) [1]–[5] are interesting not only because of their fundamental importance but also because of their great potential for future applications [6]–[8]. The TI is a new state of quantum matter and is distinct from a simple metal or insulator in the sense that its bulk is insulating (with a bulk band gap), while its surface (or edge) is metallic due to the presence of gapless surface (edge) states. Those surface states are spin split but with double degeneracy at the Dirac point, which is protected by the time-reversal symmetry. The surface states, consisting of an odd number of massless Dirac cones, are robust against time-reversal-invariant perturbations, and also are very different from graphene where spin degeneracy is reserved (the concept of pseudo-spin describing the sublattice degeneracy is used for graphene). In early studies of TIs, due to the absence of realistic materials, most discussions were based on the model Hamiltonians [1]–[5]. However, within the last couple of years this field has been strongly motivated by the discovery of several real compounds.

The two-dimensional (2D) TI (also called the quantum spin Hall insulator) has been proposed and realized in the HgTe/CdTe quantum well structure by varying the well thickness [9]–[11]. The three-dimensional (3D) topological nature of the $Bi_{1-x}Sb_x$ alloy has been demonstrated by fine-tuning the alloy concentration [12]–[14]. In both cases, the bulk band gap is of the order of 10 meV, too small for the potential applications. In our recent studies, however, we predicted that the Bi_2Se_3 family of compounds [15] (i.e. Bi_2Te_3, Bi_2Se_3 and Sb_2Te_3), which are well-known thermoelectric materials [16]–[18], are strong TIs, with surface states consisting of a single Dirac cone at the Γ point. In particular, for Bi_2Se_3 a bulk band gap of 0.3 eV is predicted, much larger than the energy scale of room temperature. The compounds are stoichiometric, chemically very stable, and easy to synthesize, and yet their surface states are very simple. The existence of such surface states in this family of compounds has been experimentally confirmed on samples ranging from bulk to thin film [19]–[24].

More and more studies are now in progress on the Bi_2Se_3 family of compounds, and quantitative studies and comparison with experiments are of increasing importance. However, fully self-consistent *ab initio* calculations are quite limited due to the large system size. In order to study both the bulk and the surface behavior simultaneously, which is necessary to identify its topological nature, a large slab-type supercell is needed, and demanding calculations are involved. To simplify future studies, it is therefore desirable to have a highly accurate effective Hamiltonian, which is one of the main purposes of this paper. It is well known that the maximally localized Wannier functions (MLWF) method introduced by Marzari and co-workers [25, 26] has played an important role in constructing the effective Hamiltonian. However, the algorithm involved does not hold symmetry when the MLWF are constructed from Bloch functions. To solve this problem we introduce a set of projected atomic Wannier functions (PAWF), whose symmetry can be easily controlled while maintaining high accuracy. With this set of PAWF bases, a highly accurate effective Hamiltonian can be constructed. Using this Hamiltonian, we demonstrate that the topological nature as well as the details of surface states can all be well reproduced.

The organization of the present paper is as follows. In section 2, we describe the crystal structure and symmetries of the Bi_2Se_3 family of compounds. Section 3 is devoted to the construction of the effective Hamiltonian from PAWF. In section 4, this Hamiltonian is applied to the Bi_2Se_3 system, and the topological nature and surface properties are analyzed. In section 5, we discuss the 3D topological phase transition by doping Se into Sb_2Te_3. Final remarks are summarized in section 6.

2. Crystal structure and symmetries

The Bi_2Se_3 family of compounds has a rhombohedral crystal structure with space group $\mathbf{D_{3d}^5}$ ($\mathbf{R\bar{3}m}$); we take Bi_2Se_3 as an example in the following. As shown in figure 1(a), the system has a layered structure with five atomic layers as a basic unit (cell), named a quintuple layer (QL). The inter-layer bonding within the QLs is strong because of the dominant covalent character, but the bonding between the QLs is much weaker due to the van der Walls-type interaction. The binary (with twofold rotation symmetry), bisectrix (appearing in the reflection plane) and trigonal (with threefold rotation symmetry) axes are taken as the x, y and z axes, respectively, and the primitive translation vectors $\mathbf{t_{1,2,3}}$ shown in figure 1 are

$$\begin{aligned} \mathbf{t_1} &= (-a/2, -\sqrt{3}a/6, c/3), \\ \mathbf{t_2} &= (a/2, -\sqrt{3}a/6, c/3), \\ \mathbf{t_3} &= (0, \sqrt{3}a/3, c/3). \end{aligned} \tag{1}$$

Here, a and c are lattice constants in the hexagonal cell. The corresponding reciprocal vectors $\mathbf{s_{1,2,3}}$, defined by $\mathbf{s_i} \cdot \mathbf{t_j} = \mathbf{2}\pi\,\delta_{\mathbf{ij}}$, are given as

$$\begin{aligned} \mathbf{s_1} &= (-1, -\sqrt{3}/3, b)h, \\ \mathbf{s_2} &= (1, -\sqrt{3}/3, b)h, \\ \mathbf{s_3} &= (0, 2\sqrt{3}/3, b)h, \end{aligned} \tag{2}$$

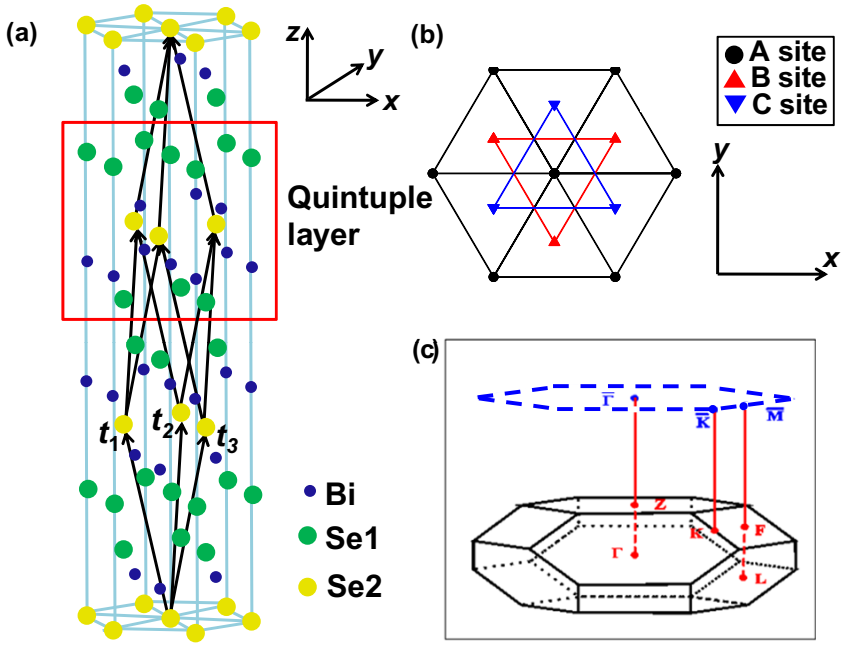

Figure 1. Crystal structure of the Bi_2Se_3 family of compounds. (a) The hexagonal supercell containing 15 atomic layers and primitive translation vectors $\mathbf{t}_{1,2,3}$. (b) The top view of a QL in the triangle lattice. Three sets of different sites, labeled as A, B and C sublattices, respectively, are presented. Owing to the $\mathbf{D}^5_{3d}$ symmetry, the stacking of atomic layers along the z-direction is in the order of . . . –C(Se1)–A(Se1)–B(Bi)–C(Se2)–A(Bi)–B(Se1)–C(Se1) . . . . (c) The first BZ. Four nonequivalent TRIM points Γ (0, 0, 0), $L(\pi, 0, 0)$, $F(\pi, \pi, 0)$ and $Z(\pi, \pi, \pi)$ are denoted in the 3D BZ. The corresponding surface 2D BZ is represented by the dashed blue hexagon, and $\overline{\Gamma}$, $\overline{M}$ and $\overline{K}$ are the corresponding TRIM special k points in the surface BZ.

with

$$b = a/c,$$
$$h = 2\pi/a. \tag{3}$$

As shown in figure 1(a), we take Se2 to be at the origin (0, 0, 0); then two Bi sites are at $(\pm\mu, \pm\mu, \pm\mu)$, and two Se1 are at $(\pm\nu, \pm\nu, \pm\nu)$, defined in the unit of primitive translation vectors. All the experimental lattice parameters and internal parameters μ and ν are listed in table 1. Figure 1(c) shows the 3D first Brillouin zone (BZ) and the 2D surface BZ of Bi_2Se_3. $\Gamma(0, 0, 0)$, $L(\pi, 0, 0)$, $F(\pi, \pi, 0)$ and $Z(\pi, \pi, \pi)$ are four time-reversal invariant momentum (TRIM) points in the 3D BZ. $\Gamma(0, 0, 0)$ and $Z(\pi, \pi, \pi)$ are projected as $\overline{\Gamma}$, and $L(\pi, 0, 0)$ and $F(\pi, \pi, 0)$ are projected as $\overline{M}$ in the surface BZ. For the choice of our cell, Bi_2Se_3 has the inversion symmetry with inversion center at Se2. The space group $\mathbf{R\bar{3}m}$ can be constructed from three symmetry generators: I (inversion), c_{3z} (threefold rotation around z) and σ_x (mirror plane with its normal along x).

Table 1. Experimental lattice parameters and the internal parameters of the Bi_2Se_3 family of compounds [27].

		Sb_2Te_3	Bi_2Te_3	Bi_2Se_3
Lattice constant	a (Å)	4.250	4.383	4.138
	c (Å)	30.35	30.487	28.64
Inner coordinates	μ	0.400	0.400	0.399
	ν	0.211	0.212	0.206

3. The effective Hamiltonian from projected atomic Wannier functions (PAWF)

The construction of the PAWF basis and the effective Hamiltonian is a post-process of *ab initio* calculations. The fully self-consistent *ab initio* calculations are performed for the bulk compounds in the framework of density functional theory (DFT) [28] using the BSTATE (Beijing Simulation Tool for Atom Technology) package [29] with the plane-wave ultra-soft pseudo-potential method [30]. Generalized gradient approximation (GGA) of the Perdew–Burke–Ernzerhof (PBE) type is used for the exchange-correlation potential [31]. To guarantee the convergence, we use 340 eV for the cut-off energy of wave-function expansion and a $10 \times 10 \times 10$ K-points mesh in the BZ. All structure parameters of Bi_2Se_3, Sb_2Te_3 and Bi_2Te_3 are chosen from the experimental data listed in table 1.

After completing the self-consistent *ab initio* calculations two steps are followed to construct the PAWF basis and the effective Hamiltonian. First, we disentangle an isolated group of bands through minimizing the invariant part of the spread functional of wave functions (Ω_{I}) [26]. In the second step, the Hilbert space of this isolated group of bands is directly projected on to atomic p orbitals (keeping the angular symmetry). Finally, a set of unitary rotation matrices $\mathcal{U}(\mathbf{k})$ can be obtained, which can be used straightly to construct the PAWF and the effective Hamiltonian. The details are elucidated in the following.

3.1. Extracting the isolated group of bands

For a simple band insulator, the disentangling procedure is, typically, unnecessary, and it is straightforward to choose a set of Wannier bases expanding the identical space of occupied Bloch states. However, in our case, aiming to describe the electronic characters near the Fermi level, both occupied and unoccupied bands are necessary, which are not detached from higher bands. By partial density of states (PDOS) analysis, we know that the states near the Fermi level mainly come from the contribution of p orbitals of both Bi and Se atoms. Since each atom has three p orbitals, the total number of orbitals of interest is $N = 15$, consisting of nine valence bands and six unoccupied bands. Following the procedure described in MLWF [26], we introduce a big enough outer energy window (−20 eV, 20 eV), and a small inner window (−6 eV, 2 eV). A smooth subspace (of N bands) can then be constructed by optimizing the spread functional, namely Ω_{I}, as suggested in [26].

3.2. The PAWF basis and the effective Hamiltonian

The Bloch functions $\psi_{m\mathbf{k}}$ of the isolated group of N bands are directly projected on to the atomic p orbitals, and the unitary rotation matrices $\mathcal{U}(\mathbf{k})$ can be obtained by proper

orthonormalization [25, 32]. With these unitary rotation matrices $\mathcal{U}(\mathbf{k})$, we can obtain well-localized PAWF by rotating the original Bloch functions in the following way:

$$|\tilde{\psi}_{n\mathbf{k}}\rangle = \sum_{m=1}^{N} \mathcal{U}_{mn}(\mathbf{k})|\psi_{m\mathbf{k}}\rangle, \tag{4}$$

$$|W_n^R\rangle = \frac{1}{N_k}\sum_k \mathrm{e}^{-\mathrm{i}kR}|\tilde{\psi}_{n\mathbf{k}}\rangle, \tag{5}$$

where n and $m = 1, 2, \ldots, N$ are band indices, N_k is the total number of k-points and W_n^R denotes the nth PAWF orbital, which is centered at the lattice vector R.

Using the PAWF as the basis, the effective Hamiltonian $H^W(k)$ can be obtained correspondingly by rotating the original Hamiltonian,

$$H^W(\mathbf{k}) = \mathcal{U}^\dagger H(\mathbf{k})\mathcal{U}. \tag{6}$$

Because our PAWF are constructed from atomic orbitals, which have atomic symmetry, we can further symmetrize the Hamiltonian using their atomic characters. Using Fourier transformation, the PAWF effective Hamiltonian in real space can be given by

$$H^W(\mathbf{R}) = \frac{1}{N_k}\sum_{\mathbf{k}} \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{R}} H^W(\mathbf{k}). \tag{7}$$

The Hamiltonians at any other k-point can be now obtained by transforming $H^W(\mathbf{R})$ back into k space,

$$H^W(\mathbf{k}) = \sum_{\mathbf{R}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{R}} H^W(\mathbf{R}). \tag{8}$$

All the hopping matrix elements in the effective Hamiltonian are calculated directly by constructing the PAWF. This approach differs considerably from the conventional Slater–Koster TB method [33], where hopping integrals up to certain neighbors are obtained by a fitting method. The present method has several special advantages. Firstly, it is not necessary to maximize the localization of Wannier functions (WF). Such a step will usually break the local symmetry of WF. In our case, the WF can be constrained to satisfy local crystal symmetry, and the PAWF effective Hamiltonian can be symmetrized while keeping the necessary accuracy. Secondly, the atomic spin–orbit coupling (SOC) can be easily implemented. Since the WF are constructed from atomic orbitals, they are already quite localized although not maximally localized by definition. The PAWF effective Hamiltonian is expected to be useful for large supercell calculations without extra errors from the symmetry problem.

3.3. The spin–orbit coupling

The SOC is important for the topological nature of the Bi_2Se_3 family of compounds, and should be included in all analyses. Since SOC is mostly local atomic physics and has little $\mathbf{k}$-dependence, the simplest way to supplement SOC on top of our PAWF effective Hamiltonian (denoted as H^0 from now on) is to include an additional local term H^{soc} in the total Hamiltonian [34],

$$H^{\mathrm{tot}} = H^0 + H^{\mathrm{soc}}, \tag{9}$$

where

$$H^{\mathrm{soc}} = (\hbar/4m^2c^2)[\nabla \mathbf{V} \times \mathbf{P}] \cdot \sigma. \tag{10}$$

H^{soc} comes from the SOC interaction, while $\mathbf{V}$ and σ represent total potential and Pauli spin matrices, respectively.

As discussed above, our PAWF are constructed from atomic p orbitals (having atomic angular symmetry); the SOC Hamiltonian can be straightforwardly written down, in terms of those PAWF bases,

$$|p_x, \uparrow\rangle, |p_y, \uparrow\rangle, |p_z, \uparrow\rangle, |p_x, \downarrow\rangle, |p_y, \downarrow\rangle, |p_z, \downarrow\rangle, \tag{11}$$

where $\uparrow$ ($\downarrow$) indicates the spin. With this basis set for each atom, the H^{soc} part can be expressed as

$$H^{\mathrm{soc}} = \frac{\lambda}{2} \begin{bmatrix} 0 & -\mathrm{i} & 0 & 0 & 0 & 1 \\ \mathrm{i} & 0 & 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 & -1 & \mathrm{i} & 0 \\ 0 & 0 & -1 & 0 & \mathrm{i} & 0 \\ 0 & 0 & -\mathrm{i} & -\mathrm{i} & 0 & 0 \\ 1 & \mathrm{i} & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{12}$$

where λ denotes the SOC parameter. We take the SOC parameters of Bi, Se, Sb and Te atoms from Wittel's spectral data ($\lambda_{\mathrm{Bi}} = 1.25\,\mathrm{eV}$, $\lambda_{\mathrm{Se}} = 0.22\,\mathrm{eV}$, $\lambda_{\mathrm{Sb}} = 0.4\,\mathrm{eV}$, $\lambda_{\mathrm{Te}} = 0.49\,\mathrm{eV}$) [35].

3.4. Accuracy of the effective Hamiltonian: results for bulk

To demonstrate the quality of the PAWF effective Hamiltonian, we compare here several important properties calculated from the PAWF Hamiltonian with those obtained from the *ab initio* calculations. The first property to be compared is the calculated bulk band structure. As shown in figure 2(a), for Bi_2Se_3 without SOC the band structures obtained from the PAWF effective Hamiltonian almost exactly reproduce the *ab initio* band structures. After taking into account the SOC interaction (as shown in figure 2(b)), the quality is slightly reduced, but it is still reasonably high since only the **k**-independent SOC interaction is implemented in the effective Hamiltonian. For Bi_2Te_3 and Sb_2Te_3, the PAWF effective Hamiltonian shows similar quality (shown in figures 3(a) and (b)).

Parity is another important quantity to distinguish the topological character of the Bi_2Se_3 family of compounds. Because the present systems possess inversion symmetry, the method proposed by Fu and Kane [5] is used here. For all three systems we calculate the parity eigenvalues of nine occupied bands and the first conduction band at the TRIM points Γ, L, F and Z. The results for the Γ point, obtained from the effective Hamiltonian, are listed in table 2. As can be clearly seen, the parity products of occupied bands are -1 for the Γ points. For the other TRIM points, L, F and Z, the detailed numbers are not listed here, but our results give +1 for three of the compounds. From the parity numbers, we can therefore identify the corresponding Z_2 invariants [1, 4, 5] as 1 for all the three compounds. This means that they have a nontrivial topological nature, in good agreement with our previous studies [15]. From the above comparisons we conclude that the quality of the PAWF effective Hamiltonian is as good as that obtained from the *ab initio* calculations.

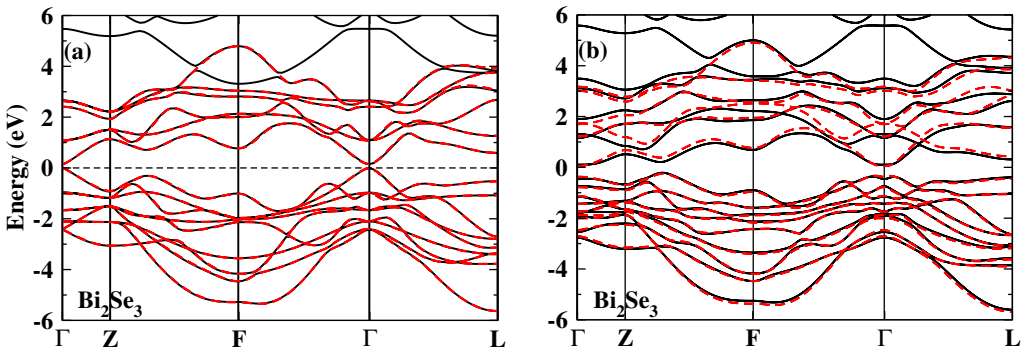

Figure 2. Comparison of PAWF bands (red dashed lines) with the *ab initio* band structures (black solid lines) of Bi_2Se_3. (a) The results without SOC. The PAWF effective Hamiltonian almost exactly reproduces the *ab initio* band structures. (b) The results with SOC for Bi_2Se_3.

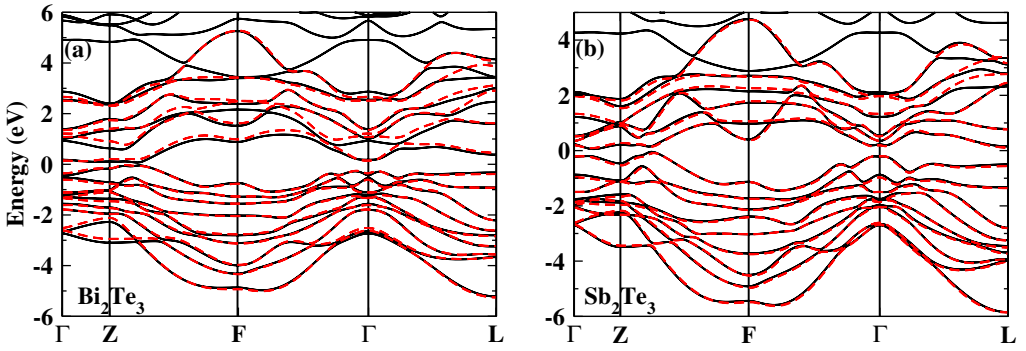

Figure 3. The comparison of PAWF bands (red dashed lines) with the *ab initio* band structures (black solid lines) of (a) Bi_2Te_3 and (b) Sb_2Te_3. Only the results obtained after including the SOC interaction are shown here.

4. The properties of surface states

Once the effective Hamiltonian has been constructed from the above procedures, we are able to calculate the surface states of semi-infinite systems from the iterative Green's function method [36], as described in our previous work [15]. The existence of gapless spin-filter surface states is the direct manifestation of the topological nature. The crystal structure of the Bi_2Se_3 family of compounds can be understood as the stacking of QLs along the z-direction. The inter-layer bonding between two QLs is much weaker than that inside the QL; it is natural to expect that the cleavage plane should be between two QLs. This fact has been well confirmed by recent experiments on the layer-by-layer MBE growth of ultra-thin film [21]–[24]. We therefore focus on this type of surface termination (with Se1 atomic layer as the top most layer). Our previous work has shown that gapless surface states with a single Dirac point (located at $\overline{\Gamma}$) exist for all three compounds, Bi_2Se_3, Bi_2Te_3 and Sb_2Te_3, but not for Sb_2Se_3. The dispersion of surface states around the Dirac point is highly linear, and these surface states can be described by a simple continuous model [15]. We will not repeat those results here; however, we focus on

Table 2. The parity eigenvalues of the energy bands of the Bi_2Se_3 family of compounds at Γ point, obtained from the effective Hamiltonian. The parity numbers of nine occupied bands and the first conduction band are shown (the corresponding band energy increases from left to right). The parity products of nine occupied bands are given in parentheses.

	Parity									
Bi_2Se_3 (−1)	−1	1	−1	−1	1	1	−1	−1	1;	−1
Bi_2Te_3 (−1)	−1	1	1	−1	−1	1	−1	−1	1;	−1
Sb_2Te_3 (−1)	1	−1	−1	1	−1	1	−1	−1	1;	−1

some other details of surface states, including the penetration depth, the spin-resolved Fermi surface and the chiral spin texture of the surface states.

4.1. The penetration depth of surface states

The spread of surface states, in other words the spatial penetration depth into the bulk, is an essential quantity for the potential applications of these surface states, if any. In order to clarify this point, a free standing slab model is constructed using our PAWF effective Hamiltonian. We use a slab consisting of 25 QLs, which is thick enough to avoid the direct coupling between the two surfaces, i.e. the top and the bottom surfaces of the slab. In order to make the calculations more realistic, we further take into account the surface correction. The bulk PAWF effective Hamiltonian is used to construct the Hamiltonian of the slab; however, for those layers close to the surface, the on-site energies of PAWF should be modified from the bulk counterparts due to the presence of vacuum (i.e the surface potential). By comparison with the fully self-consistent *ab initio* calculations of thin slab, all the correction of on-site energies for PAWF ($\Delta E_n = E_n^{\text{surface}} - E_n^{\text{bulk}}$) can be obtained.

The calculated band structures of our 25-QLs Bi_2Se_3 slab along $\overline{\Gamma} \to \overline{K}$ and $\overline{\Gamma} \to \overline{M}$ lines are shown in figures 4 and 5. Clearly, within the bulk energy gap, there exist two surface bands (marked by Σ_1 and Σ_2), which are degenerate at $\overline{\Gamma}$ point (where the so-called Dirac cone exists). The two surface bands (Σ_1 and Σ_2) are sampled by several k-points (with k_i, $i = 1, 2, 3$). The real space distribution of eigen wavefunctions for those sampling k-points are plotted in figures 4(b) and (c) and figures 5(b) and (c), respectively. For those k-points close to the Dirac point (such as k_1 and k_2), the distribution of wavefunctions are very localized to the surface region, with a typical spread of about 2 QLs (about 2 nm in thickness). By moving away from the Dirac point, the penetration depth increases, and finally for k_3 point, where the surface state almost merges with the bulk states, the eigen wavefunction becomes an extended state. For both $\overline{\Gamma}$–$\overline{M}$ and $\overline{\Gamma}$–$\overline{K}$ lines, the same behaviors are observed. To further observe the oscillating behavior of surface state distribution over atomic layers, we plot the distribution for k_1 in the inset. Finally, we can conclude that surface states are very localized to the surface region, and the penetration depth is about 2 or 3 QLs.

To make further comparison, we perform fully self-consistent *ab initio* calculations for a three-QL slab. The real space charge distribution of surface states can thus be obtained by integrating the states within an energy window of 10 meV around the Dirac point. As shown in figure 6, the charge distribution is localized at the surface region. Furthermore, the charges

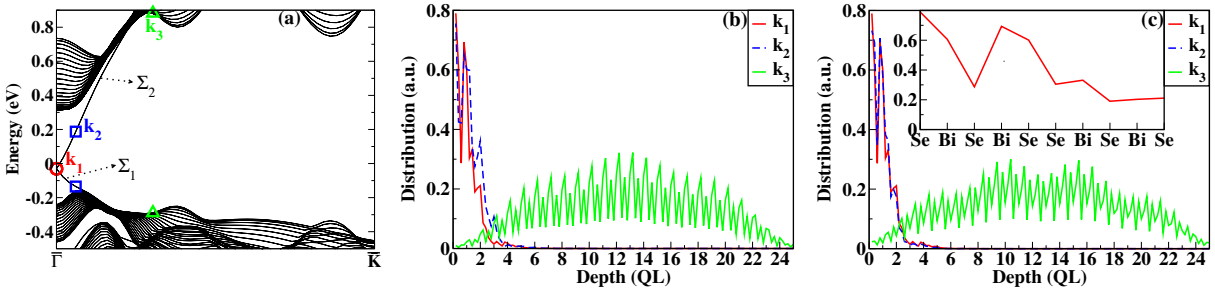

Figure 4. (a) The calculated band structures of the Bi_2Se_3 slab with a film thickness of 25 QLs. The bands along the $\overline{\Gamma} \rightarrow \overline{K}$ direction are shown. The Fermi level is located at energy zero. The two surface bands in the bulk gap (around the $\overline{\Gamma}$ point) are denoted as Σ_1 and Σ_2. (b, c) The real-space distribution of eigen wavefunctions ($|\psi_{nk}(\mathbf{r})|^2$) for different k-points (k_1, k_2 and k_3 as indicated in (a)). (b) and (c) are for k-points along the Σ_1 or Σ_2 lines, respectively. The inset of (c) is the zoomed-in picture for the k_1 point eigen wavefunction, where the atomic layer resolution is visible.

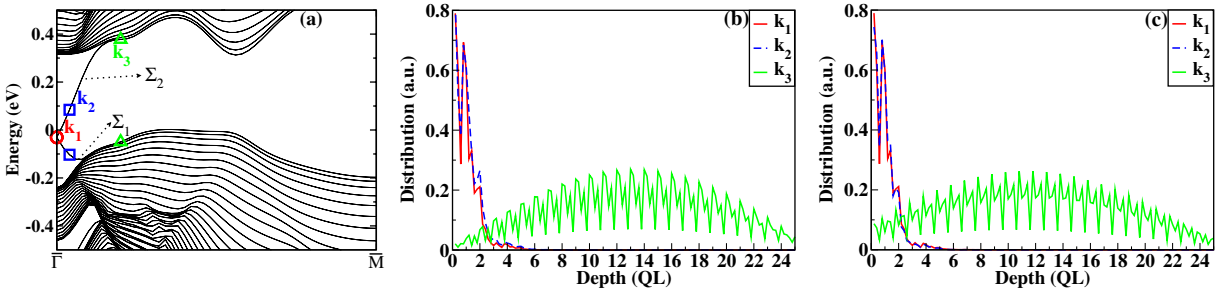

Figure 5. The same as figure 4, but for bands along the $\overline{\Gamma} \rightarrow \overline{M}$ direction.

are not centered at the atom, but rather distributed mostly between the Bi and Se atoms, where strong covalent bonding is expected.

4.2. The spin-resolved Fermi surface

Probing the π Berry phase enclosed by the Fermi surface of surface states is one of the most direct methods for distinguishing TIs. We now analyze the spin-resolved Fermi surfaces around the Dirac point of the semi-infinite Bi_2Se_3 system. With the Green's functions calculated from the effective Hamiltonian, the spin-filter surface states and the corresponding Fermi surfaces can be obtained directly.

As shown in figure 7, when the Fermi level is close to the Dirac point the corresponding Fermi surface is nearly a perfect circle, while if the Fermi level is away from the Dirac point the properties of the surface states are significantly affected by the bulk states and thus satisfy the crystal symmetry. For example, the Fermi surface for $E_f = 0.1$ eV (shown in figure 7(a)) looks like a small circle, while the Fermi surface for $E_f = 0.25$ eV (shown in figure 7(b)) looks like a hexagon satisfying c_{3z} symmetry. The spin orientation for states around the Fermi surface is marked as well by arrows. The magnitude of the spin z-component is very small, only about

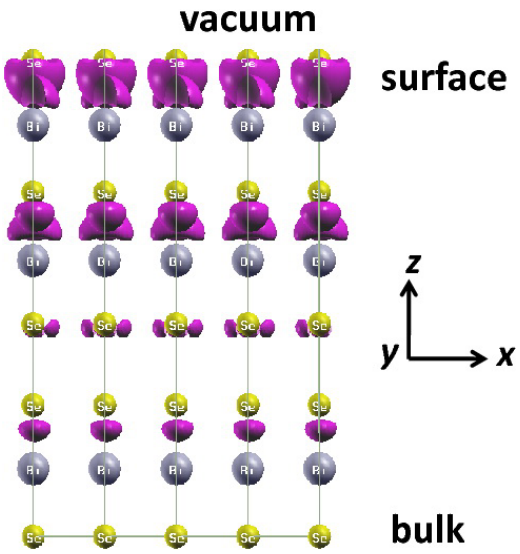

Figure 6. The spatial charge distribution within an energy window of 10 meV around the Fermi level. The top atomic layer means the surface of the slab. It is clearly seen that electrons are mainly localized at the surface of the slab.

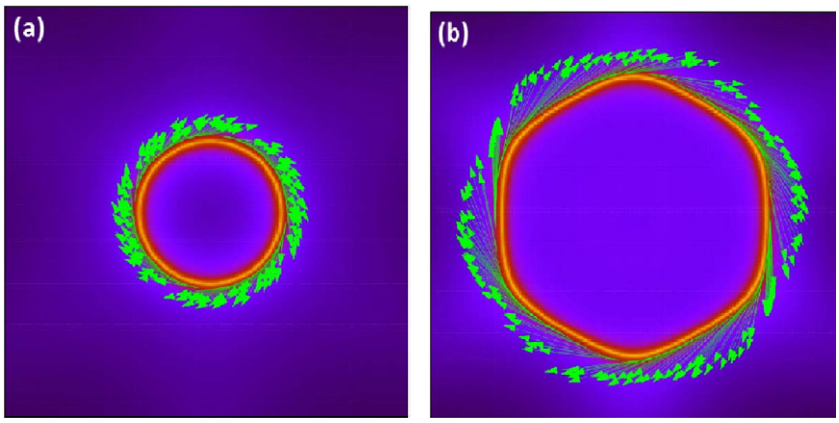

Figure 7. The spin-resolved Fermi surface of surface states. (a, b) The Fermi surfaces for Fermi energy at 0.1 and 0.25 eV, respectively. The Fermi surfaces are denoted by a red circle, and the in-plane spin orientation is indicated by green arrows.

1/40 of the in-plane component. Thus the spin lies almost completely in the plane. Moving around the Fermi surface, the spin orientation rotates simultaneously, forming a spin–orbit ring, which carries a π Berry phase. This signifies the topological nontrivial properties of Bi_2Se_3. In addition, as shown in figure 7, the spin orientation of the ring belongs to the left chirality (the normal direction of the semi-infinite system is defined as the z-direction), which is again one of the important manifestations of nontrivial topological characters.

4.3. The linear dispersion of surface states

The gapless surface states that connect the bulk valence and conduction bands have almost linear energy dependence near the Dirac point. A linear band structure in 2D should lead to linear

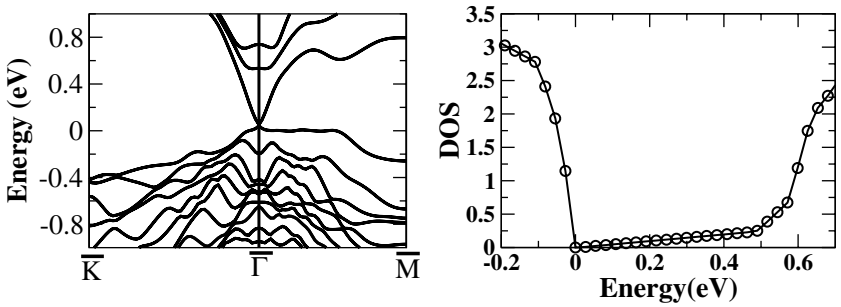

Figure 8. (a) Energy bands of a three-QL slab. (b) The corresponding DOS. The DOS near the Fermi level is highly linear to energy due to the presence of Dirac cone-type surface states.

density of states (DOS). To be specific, in the bulk energy gap the energy bands are mostly from the surface states, so the DOS in the bulk energy gap are expected to be highly linear. Figure 8(a) shows the band structures of a three-QL slab. Its corresponding DOS (presented in figure 8(b)) shows very nice linear energy dependence within the bulk energy gap, as expected. This type of DOS can be easily measured by low-temperature scanning tunneling spectroscopy (STS), which will provide an indirect method to probe the existence of linear surface states.

5. Three-dimensional topological phase transition

The above discussions presented in sections 2 and 3 concentrate on the effective Hamiltonian and detailed understanding of the properties of surface states. In such a way, we can analyze the topological nature of a specific compound. However, the topological nature can be also understood from simple bulk studies, where a gap close–reopening transition (a topological phase transition) can be obtained by tuning some parameters [37]. We show in this section that such a phase transition may be obtained in the $Sb_2(Te_{1-x}Se_x)_3$ alloy system.

The virtual crystal approximation (VCA) method proposed by Bellaiche [38] is used here to simulate the doping process. The solid solution elements are treated as different atomic species with their own weights, rather than being treated as a whole virtual atom. With this method we investigate the Se doping dependence of the band structure of $Sb_2(Te_{1-x}Se_x)_3$. We assume that crystal parameters and positions of inner atoms change linearly with doping, and they are obtained by linear interpolation in calculation. The evolution of the gap energy of $Sb_2(Te_{1-x}Se_x)_3$ as a function of Se concentration is calculated and illustrated in figure 9. With increasing Se doping, the valence band maximum and conduction band minimum gradually become closer ($0 < x \leqslant 0.94$), which is attributed to the gradually decreasing SOC strength. Consequently, the two bands cross at the critical point $x = 0.94$, resulting in a 3D topological quantum phase transition. A band inversion appears when the doping concentration further increases. Experimentally, the most stable crystal structure of Sb_2Se_3 is orthorhombic (with space group D_{2h}^{16}), which is a structure distorted from the rhombohedral structure of Sb_2Te_3; this makes the realization of the topological phase transition discussed above difficult. A wide range of solid solutions of $Sb_2(Te_{1-x}Se_x)_3$ having rhombohedral structure (identical to Sb_2Te_3) have been reported experimentally for $0 < x < 0.67$ in the literature [39]; however, to realize

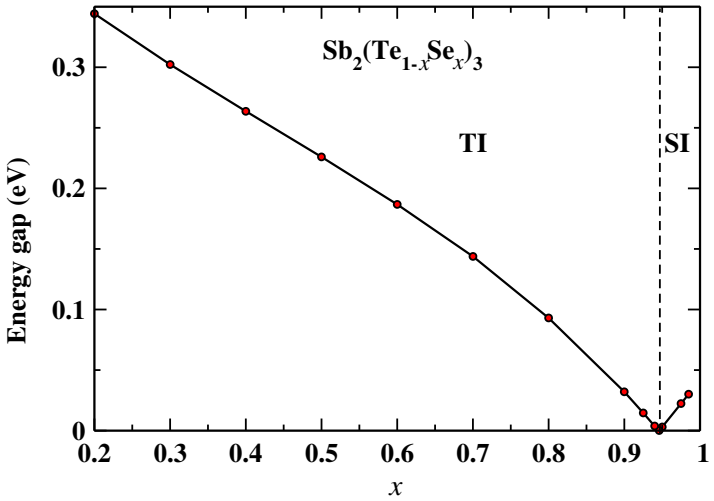

Figure 9. The gap energy of $Sb_2(Te_{1-x}Se_x)_3$ at the Γ point as a function of Se concentration x. TI denotes the TI, whereas SI means the simple insulator.

the topological phase transition, further solid solutions (with x around 0.94) are needed. This is of course a difficult condition; however, it is worthy of being tried, particularly for samples prepared in thin film form or under pressure. At the topological phase transition point, the 3D Dirac cone can be expected in the bulk band structure [37].

6. Conclusions

Bi_2Se_3, Bi_2Te_3 and Sb_2Te_3 systems are a new class of TI. In the present paper, we construct the effective Hamiltonian for this family of compounds based on the PAWF method. The effective Hamiltonian can well reproduce the *ab initio* band structures and their topological nature. The penetration depth and the spin-resolved Fermi surfaces of surface states are calculated and analyzed in detail. At the end of the paper, we discuss the 3D topological phase transition in $Sb_2(Te_{1-x}Se_x)_3$ and suggest that the 3D Dirac cone can be obtained by doping Se in Sb_2Te_3. Finally, we hope that our effective Hamiltonian can lead to further understanding of these materials in future.

Acknowledgments

We acknowledge valuable discussions with C X Liu, X L Qi, S C Zhang, S Q Shen, Q Niu and support from the NSF of China, the 973 Program of China and the International Science and Technology Cooperation Program of China.

References

[1] Kane C L and Mele E J 2005 *Phys. Rev. Lett.* **95** 146802
[2] Bernevig B A and Zhang S C 2006 *Phys. Rev. Lett.* **96** 106802
[3] Moore J E and Balents L 2007 *Phys. Rev.* B **75** 121306
[4] Liang Fu, Kane C L and Mele E J 2007 *Phys. Rev. Lett.* **98** 106803

[5] Liang Fu and Kane C L 2007 *Phys. Rev.* B **76** 045302
[6] Liang Fu and Kane C L 2008 *Phys. Rev. Lett.* **100** 096407
[7] Qi X L, Hughes T L and Zhang S C 2008 *Phys. Rev.* B **78** 195424
Qi X-L, Li R, Zang J and Zhang S-C 2009 *Science* **323** 1184
[8] Essin A M, Moore J E and Vanderbilt D 2009 *Phys. Rev. Lett.* **102** 146805
[9] Bernevig B A, Hughes T L and Zhang S C 2006 *Science* **314** 1757
[10] König M, Wiedmann S, Brüne C, Roth A, Buhmann H, Molenkamp L W, Qi X L and Zhang S C 2007 *Science* **318** 766
[11] Dai X, Hughes T L, Qi X-L, Fang Z and Zhang S-C 2008 *Phys. Rev.* B **77** 125319
[12] Hsieh D, Qian D, Wray L, Xia Y, Hor Y S, Cava R J and Hasan M Z 2008 *Nature* **452** 970
[13] Teo J C Y, Fu L and Kane C L 2008 *Phys. Rev.* B **78** 045426
[14] Zhang H-J, Liu C-X, Qi X-L, Deng X Y, Dai X, Zhang S-C and Fang Z 2009 *Phys. Rev.* B **80** 085307
[15] Zhang H J, Liu C X, Qi X L, Dai X, Fang Z and Zhang S-C 2009 *Nat. Phys.* **5** 438
[16] Snyder G J and Toberer E S 2008 *Nat. Mater.* **7** 105
[17] Mishra S K, Satpathy S and Jepsen O 1997 *J. Phys.: Condens. Matter* **9** 461
[18] Larson P 2006 *Phys. Rev.* B **74** 205113
[19] Xia Y *et al* 2009 *Nat. Phys.* **5** 398
[20] Chen Y L *et al* 2009 *Science* **325** 178
[21] Zhang G, Qin H, Teng J, Guo J, Guo Q, Dai X, Fang Z and Wu K 2009 *Appl. Phys. Lett.* **95** 053114
[22] Li Y-Y *et al* 2009 arXiv.org:0912.5054
[23] Zhang Y *et al* 2009 arXiv.org:0911.3706
[24] Zhang T *et al* 2009 *Phys. Rev. Lett.* **103** 266803
[25] Marzari N and Vanderbilt D 1997 *Phys. Rev.* B **56** 12847
[26] Souza I, Marzari N and Vanderbilt D 2001 *Phys. Rev.* B **65** 035109
[27] Wyckoff R W G 1964 *Crystal Structures* vol 2 (New York: Wiley)
[28] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[29] Fang Z and Terakura K 2002 *J. Phys.: Condens. Matter* **14** 3001
[30] Vanderbilt D 1990 *Phys. Rev.* B **41** 7892
[31] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[32] Anisimov V I *et al* 2005 *Phys. Rev.* B **71** 125119
[33] Slater J C and Koster G F 1954 *Phys. Rev.* **94** 1498
[34] Chadi D J 1977 *Phys. Rev.* B **16** 790
[35] Wittel K and Manne R 1974 *Theor. Chim. Acta* **33** 347
[36] Sancho M P L, Sancho J M L and Rubio J 1984 *J. Phys. F: Met. Phys.* **14** 1205
Sancho M P L, Sancho J M L and Rubio J 1985 *J. Phys. F: Met. Phys.* **15** 851
[37] Murakami S 2008 *Prog. Theor. Phys. Suppl.* **176** 279
[38] Bellaiche L 2000 *Phys. Rev.* B **61** 7877
[39] Procarione W and Wood C 1970 *Phys. Status Solidi* **42** 871