Symmetry analysis of strain, electric and magnetic fields in the Bi2Se3-class of topological insulators

Paper

Brems, Mathias Rosdahl; Paaske, Jens; Lunde, Anders Mathias; Willatzen, Morten

Published in:
New Journal of Physics

Link to article, DOI:
10.1088/1367-2630/aabcfc

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

PAPER • OPEN ACCESS

Symmetry analysis of strain, electric and magnetic fields in the Bi_2Se_3-class of topological insulators

To cite this article: Mathias Rosdahl Brems *et al* 2018 *New J. Phys.* **20** 053041

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
29 December 2017

REVISED
15 March 2018

ACCEPTED FOR PUBLICATION
10 April 2018

PUBLISHED
16 May 2018

PAPER

Symmetry analysis of strain, electric and magnetic fields in the Bi_2Se_3-class of topological insulators

Mathias Rosdahl Brems[1], **Jens Paaske**[2], **Anders Mathias Lunde**[2] **and Morten Willatzen**[3,4]

[1] Technical University of Denmark, Department of Photonics Engineering, Kgs-Lyngby, DK-2800, Denmark
[2] Center for Quantum Devices, Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark
[3] Beijing Institute of Nanoenergy and Nanosystems, Chinese Academy of Sciences, Beijing, China and Technical University of Denmark, Department of Photonics Engineering, Kgs-Lyngby, DK-2800, Denmark
[4] Author to whom any correspondence should be addressed.

E-mail: mortenwillatzen@binn.cas.cn **and** morwi@fotonik.dtu.dk

Keywords: topological insulators, electronic band structure, group theory, strain, electric field

Abstract

Based on group theoretical arguments we derive the most general Hamiltonian for the Bi_2Se_3-class of materials including terms to third order in the wave vector, first order in electric and magnetic fields, first order in strain and first order in both strain and wave vector. We determine analytically the effects of strain on the electronic structure of Bi_2Se_3. For the most experimentally relevant surface termination we analytically derive the surface state (SS) spectrum, revealing an anisotropic Dirac cone with elliptical constant energy contours giving rise to a direction-dependent group velocity. The spin-momentum locking of strained Bi_2Se_3 is shown to be modified. Hence, strain control can be used to manipulate the spin degree of freedom via the spin–orbit coupling. We show that for a thin film of Bi_2Se_3 the SS band gap induced by coupling between the opposite surfaces changes opposite to the bulk band gap under strain. Tuning the SS band gap by strain, gives new possibilities for the experimental investigation of the thickness dependent gap and optimization of optical properties relevant for, e.g., photodetector and energy harvesting applications. We finally derive analytical expressions for the effective mass tensor of the Bi_2Se_3 class of materials as a function of strain and electric field.

1. Introduction

Topological insulators have an inverted band gap which engenders topologically protected surface states (SS). Exhibiting linear dispersion, the electrons at the surface resemble massless helical Dirac fermions, with spin locked to the momentum. The prime examples of three-dimensional topological insulators are among Bi_2Se_3-class of materials [1]. This class, also known as the tetradymite group, contains compounds M_2X_3 where M is either Bi or Sb and X is a combination of Se, S and Te. The crystal structure consists of unit layers of five atomic layers, so-called quintuple layers (QL). For the simplest case of a (111) surface termination, where the surface is parallel to the quintuple layer, the Dirac cone is fully isotropic for small in-plane momentum, perturbed only by a hexagonal warping to third order in momentum [2]. For other surface terminations the Dirac cone of the SS becomes anisotropic with elliptical curves of constant energy, as reported for the (221) surface in [3]. Anisotropic Dirac fermions have interesting transport properties due to the different group velocity in different directions [4], and have attracted attention in other Dirac materials [5, 6].

For both fundamental research and applications it is interesting to be able to tune the physical properties of these materials. One way to do this is by the application of strain. It was reported in [7–9] that strain strongly affects the band gap at the Γ point of bulk Bi_2Se_3, and could even close the gap at large strain values, i.e. induce a topological phase transition. Whether this is possible to achieve in real materials is not known, but a recent study points to the feasibility of strain induced effects in Bi_2Se_3 where strain up to 3% was induced by lattice mismatch [10].

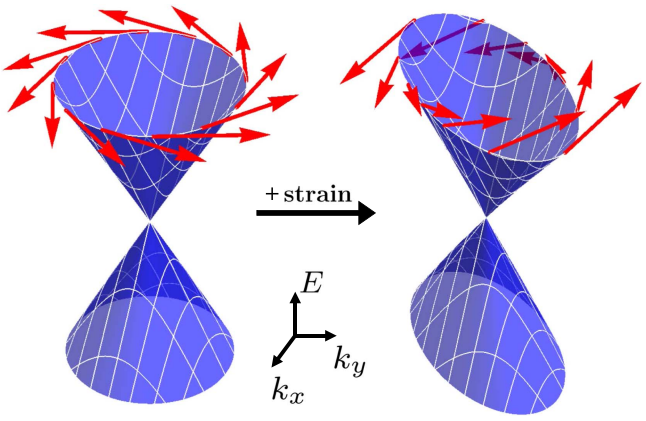

Figure 1. For unstrained Bi_2Se_3 the linear dispersion of the helical surface states is isotropic in the plane to linear order in the wave vector. In the case of strain, the in-plane rotational symmetry is broken by terms to first order in both strain and wave vector leading to a Dirac cone with elliptical contours of constant energy, here shown for $\epsilon_{xy} = 5\%$. Hence, the group velocity of the surface electrons becomes dependent on the direction. The broken rotational symmetry also allows for a spin component perpendicular to the surface, as shown by the red arrows.

In this paper, we use the method of invariants [11] to derive the most general Hamiltonian at the Γ point to third order in wave vector, first order in strain, first order in electric and magnetic fields as well as terms to first order in both wave vector and strain. The allowed third order terms in the wave vector include three terms that were neglected in a previous analysis [12]. Since this model is based solely on the symmetry of the crystal, it is valid for all materials in the Bi_2Se_3 class.

Using this model, we determine analytical expressions for the modified bulk band structure and the effective mass tensor near the Γ-point. Changes in the effective mass tensor with strain and electric field are important for transport and optical properties [13]. Comparing the strain dependence of the band gap to recent density functional theory (DFT) calculations, allows us to determine some of the strain related parameters in the band structure. From this bulk spectrum, we go on to investigate the effects of strain on the SS of a semi-infinite topological insulator. We analytically derive the spectrum of the SS, applying hard-wall boundary conditions at the surface for a semi-infinite topological insulator with a (111) surface termination. In contrast to the unstrained case, the full in-plane rotational symmetry to linear order in wave vector is broken by terms linear in both strain and wave vector, leading to an anisotropic Dirac spectrum. The spin expectation values of the SS are calculated, and it is shown that the spin-momentum locking is affected by strain, revealing a non-zero spin component perpendicular to the surface. The modified Dirac cone and spin structure of the SS is shown in figure 1. Finally, we show that strain affects the localization of the SS wave functions and thereby the band gap arising in thin films due to the coupling between SS on opposite surfaces. We show that the SS band gap changes oppositely to the bulk band gap under strain: increasing the bulk band gap, localizes the SS and decreases the SS band gap, and vice versa. This shows that not only the bulk, but also the SS band gap can be tuned via strain.

2. Model Hamiltonian

In line with [12], we wish to derive a 4-band model Hamiltonian describing the Bi_2Se_3 class, which now includes all allowed terms to orders cubic in the wave vector **k**, linear in the strain tensor ϵ_{ij}, linear in magnetic field **B** or electric field **E**, and linear in both **k** and ϵ_{ij}.

In general any 4×4 Hamiltonian can be written in terms of Dirac Γ matrices defined in terms of the Pauli matrices by:

$$\Gamma_1 = \sigma_1 \otimes \tau_1, \quad \Gamma_2 = \sigma_2 \otimes \tau_1, \quad \Gamma_3 = \sigma_3 \otimes \tau_1, \quad \Gamma_4 = \sigma_0 \otimes \tau_2, \quad \Gamma_5 = \sigma_0 \otimes \tau_3, \tag{1}$$

and their commutators $\Gamma_{ij} = \frac{1}{2i}[\Gamma_i, \Gamma_j]$. A general 4×4 matrix can then be written:

$$H_m = \varepsilon \mathbf{I} + \sum_{i=1}^{5} d_i \Gamma_i + \sum_{i,j} d_{ij} \Gamma_{ij}, \tag{2}$$

where the coefficients ε, d_i, d_{ij} must be real to ensure hermiticity. In the present case the coefficients are functions of **k**, ϵ_{ij}, **B** and **E**.

Here we use the basis $|P1^+_-, \tfrac{1}{2}\rangle, -\mathrm{i}|P2^-_+, \tfrac{1}{2}\rangle, |P1^+_-, -\tfrac{1}{2}\rangle, \mathrm{i}|P2^-_+, -\tfrac{1}{2}\rangle$, derived from the *p* orbitals of the Bi and Se atoms. The upper sign denotes the inversion eigenvalue and the $\pm\frac{1}{2}$ the total angular momentum. For a more

Table 1. Multiplication table for the irreducible representations of the group D_{3d}, (a), and the coupling constants for the basis functions of the relevant products $\Gamma^{(2)} \otimes \Gamma^{(3)}$, (b) and $\Gamma^{(3)} \otimes \Gamma^{(3)}$, (c). Note that the multiplication table is completely general since it is in principle concerning equivalence classes of representations, while the coupling constants are only valid for a specific representation, here the representation with basis functions $\{k_+, k_-\}$. Adopted from [16].

(a)

$\otimes$	$\Gamma^{(1)}$	$\Gamma^{(2)}$	$\Gamma^{(3)}$
$\Gamma^{(1)}$	$\Gamma^{(1)}$	$\Gamma^{(2)}$	$\Gamma^{(3)}$
$\Gamma^{(2)}$	$\Gamma^{(2)}$	$\Gamma^{(1)}$	$\Gamma^{(3)}$
$\Gamma^{(3)}$	$\Gamma^{(3)}$	$\Gamma^{(3)}$	$\Gamma^{(1)} \oplus \Gamma^{(2)} \oplus \Gamma^{(3)}$

(b)

	$u^{(2)}v_1^{(3)}$	$u^{(2)}v_2^{(3)}$
$\psi_1^{(3)}$	i	0
$\psi_2^{(3)}$	0	$-$i

(c)

	$u_1^{(3)}v_1^{(3)}$	$u_1^{(3)}v_2^{(3)}$	$u_2^{(3)}v_1^{(3)}$	$u_2^{(3)}v_2^{(3)}$
$\psi^{(1)}$	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	0
$\psi^{(2)}$	0	$\frac{\mathrm{i}}{\sqrt{2}}$	$\frac{-\mathrm{i}}{\sqrt{2}}$	0
$\psi_1^{(3)}$	0	0	0	1
$\psi_2^{(3)}$	1	0	0	0

complete discussion see [12]. In this basis the σ matrices do not represent spin, but are related to spin by:

$$s_x = \frac{\sigma_1 \otimes \tau_3}{2} \quad s_y = \frac{\sigma_2 \otimes \tau_3}{2} \quad s_z = \frac{\sigma_3 \otimes \tau_0}{2}, \tag{3}$$

as noted in [14].

In this basis, the transformation operators corresponding to the symmetries of the crystal are inversion $I = \sigma_0 \otimes \tau_3$, three-fold rotation around the z-axis $R_3 = \mathrm{e}^{\mathrm{i}\pi\sigma_3/3} \otimes \tau_0$, two fold rotation around the x-axis $R_2 = \mathrm{i}\,\sigma_1 \otimes \tau_3$ and time-reversal $T = \mathrm{i}(\sigma_2 \otimes \tau_0)K$, where K is the complex conjugation operator. Now the Γ-matrices can be characterized by their irreducible representation under the group D_3 as well as their eigenvalues under I and T.

2.1. Group theory

Next, we construct polynomials in wave vector, strain, electric and magnetic fields transforming according to the irreducible representations of the group D_{3d}. This group is a direct product group of D_3 with the group of the inversion operator. Here we will consider D_3 separately and simply add inversion eigenvalues as well as time-reversal eigenvalues. The group D_3 has three irreducible representations: $\Gamma^{(1)}$, which is the trivial representation, $\Gamma^{(2)}$, which is one-dimensional as well and the two-dimensional representation $\Gamma^{(3)}$.

The wave vector component k_z transforms according to the one-dimensional representation $\Gamma^{(2)}$, i.e. it only changes by multiplication of a constant under the transformations of D_3. The in-plane components, k_x and k_y, transform according to the two-dimensional irreducible representation $\Gamma^{(3)}$. Here we will use the basis $\{k_+, k_-\}$, where $k_\pm = k_x \pm \mathrm{i}k_y$. As an example we will show how to construct the second order terms transforming according to irreducible representations including only in-plane momentum. First we note that

$$\Gamma^{(3)} \otimes \Gamma^{(3)} = \Gamma^{(1)} \oplus \Gamma^{(2)} \oplus \Gamma^{(3)}, \tag{4}$$

i.e. the second order terms of in-plane momenta transform according to a representation equivalent to the sum of these three irreducible representations. The basis functions for the irreducible representations can be formed using table 1(c), with $\{u_1, u_2\} = \{v_1, v_2\} = \{k_+, k_-\}$, and we get:

$$\Gamma^{(1)}: \quad \psi^{(1)} = \frac{1}{\sqrt{2}}k_+k_- + \frac{1}{\sqrt{2}}k_-k_+ \propto k_{||}^2, \tag{5}$$

$$\Gamma^{(2)}: \quad \psi^{(2)} = \frac{\mathrm{i}}{\sqrt{2}}k_+k_- + \frac{-\mathrm{i}}{\sqrt{2}}k_-k_+ = 0, \tag{6}$$

$$\Gamma^{(3)}: \quad \{\psi_1^{(3)}, \psi_2^{(3)}\} = \{k_-^2, k_+^2\}, \tag{7}$$

Table 2. Polynomials of **k**, strain and magnetic fields and the Γ matrices under the transformations of the group D_3, inversion and time reversal. To simplify the notation we have introduced $\epsilon_{\|} = \epsilon_{xx} + \epsilon_{yy}$, $\epsilon_{\pm} = \epsilon_{xx} - \epsilon_{yy} \pm 2i\epsilon_{xy}$, $\epsilon_{z\pm} = \epsilon_{zx} \pm i\epsilon_{zy}$. In the 2-dimensional representations we have changed basis to $\Gamma^{i,j}_{\pm} = \Gamma_i \pm i\Gamma_j$. These have been constructed such that they are real, and each pair is hermitian conjugates of each other and the pairs transform like $\{k_+, k_-\}$ under the group D_3.

T	I	$\Gamma^{(1)}$	$\Gamma^{(2)}$	$\Gamma^{(3)}$
		dim = 1	dim = 1	dim = 2
−	−	$\mathrm{Re}(k_+^3)$ $\mathrm{Im}(k_+\epsilon_{z-})$ $\mathrm{Re}(k_+\epsilon_+)$ Γ_4	k_z k_z^3 $k_{\|}^2 k_z$ $\mathrm{Im}(k_+^3)$ $\epsilon_{zz}k_z$ $\epsilon_{\|}k_z$ $\mathrm{Re}(k_+\epsilon_{z-})$ $\mathrm{Im}(k_+\epsilon_+)$ Γ_3	$\{k_+, k_-\}$ $k_{\|}^2\{k_+, k_-\}$ $k_z^2\{k_+, k_-\}$ $\{ik_zk_-^2, -ik_zk_+^2\}$ $\epsilon_{zz}\{k_+, k_-\}$ $\epsilon_{\|}\{k_+, k_-\}$ $k_z\{\epsilon_{z+}, \epsilon_{z-}\}$ $k_z\{i\epsilon_-, -i\epsilon_+\}$ $\{ik_-\epsilon_{z-}, -ik_+\epsilon_{z+}\}$ $\{k_-\epsilon_+, k_+\epsilon_-\}$ $\{\Gamma_+^{1,2}, \Gamma_-^{1,2}\}$
+	+	k_z^2 $k_{\|}^2$ ϵ_{zz} $\epsilon_{\|}$ Γ_5		$\{ik_zk_+, -ik_zk_-\}$ $\{k_-^2, k_+^2\}$ $\{i\epsilon_{z+}, -i\epsilon_{z-}\}$ $\{\epsilon_-, \epsilon_+\}$
+	−	Γ_{45}	E_z Γ_{35}	$\{E_+, E_-\}$ $\{i\Gamma_+^{25,51}, -i\Gamma_-^{25,51}\}$
−	+		B_z Γ_{12} Γ_{34}	$\{B_+, B_-\}$ $\{\Gamma_+^{13,23}, \Gamma_-^{23,31}\}$ $\{\Gamma_+^{14,24}, \Gamma_-^{14,24}\}$

where we have defined $k_{\|}^2 = k_x^2 + k_y^2$. Since $\{u_1, u_2\} = \{v_1, v_2\}$ the $\Gamma^{(2)}$ term is simply 0, and we have 3 independent second order terms in the in-plane momenta: $k_{\|}^2$, which is invariant, and the pair $\{k_-^2, k_+^2\}$ transforming according to the two-dimensional irreducible representation $\Gamma^{(3)}$. The inversion and time-reversal eigenvalues follow the simple rule $\pm \otimes \pm = +$ and $\pm \otimes \mp = -$. Proceeding like this we can go to any desired order, and include external fields as well. We note that the strain tensor ϵ_{ij} transforms as k_ik_j [15]. This result follows from the definition of strain

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \tag{8}$$

where **u** is the displacement vector. The resulting irreducible basis functions are summarized in table 2.

2.2. Model Hamiltonian

For the Hamiltonian to be invariant under the group D_3 it can only contain terms from the invariant representation, $\Gamma^{(1)}$. As seen in the multiplication table 1(a), $\Gamma^{(1)}$ terms can only be formed by combining terms from the same irreducible representation. For invariance under inversion (I) and time-reversal (TR), we must combine terms and matrices with the same eigenvalues, i.e. from the same cell in table 2. For the $\Gamma^{(3)}$ representation we can use table 1(c) to construct an invariant term, $\Gamma^{(1)}$ and $\Gamma^{(2)}$ are 1 dimensional so we simply take the product.

For example, the terms $\{ik_zk_-^2, -ik_zk_+^2\}$ belong to the irreducible representation $\Gamma^{(3)}$ and are odd under both TR and I. The pair of matrices $\{\Gamma_+^{1,2}, \Gamma_-^{1,2}\}$, defined as $\Gamma^{i,j}_{\pm} = \Gamma_i \pm i\Gamma_j$, transform exactly the same way, and using table 1(c) we can therefore combine them as follows:

$$\begin{aligned}\Gamma^{(1)} : \psi^{(1)} &= \frac{1}{\sqrt{2}}ik_zk_-^2\Gamma_-^{1,2} - \frac{1}{\sqrt{2}}ik_zk_+^2\Gamma_+^{1,2} \\ &\propto k_z(ik_-^2\Gamma_-^{1,2} - ik_+^2\Gamma_+^{1,2}).\end{aligned} \tag{9}$$

This forms an invariant, and thus it is an allowed term in the Hamiltonian. Any multiple of this term is of course also an invariant, and we get a free parameter in front, which we denote Z_3. Proceeding this way we achieve the following Hamiltonian, for now excluding electric and magnetic fields:

$$H_m = \mathcal{E}_0 + \begin{pmatrix} \mathcal{M} & \beta_i^* k_i & 0 & \alpha_i^* k_i \\ \beta_i k_i & -\mathcal{M} & \alpha_i^* k_i & 0 \\ 0 & \alpha_i k_i & \mathcal{M} & -\beta_i k_i \\ \alpha_i k_i & 0 & -\beta_i^* k_i & -\mathcal{M} \end{pmatrix} - Z_1 \,\mathrm{Re}(k_+^3)\Gamma_4 + Z_2 \,\mathrm{Im}(k_+^3)\Gamma_3 + Z_3 k_z (\mathrm{i}k_-^2 \Gamma_-^{1,2} - \mathrm{i}k_+^2 \Gamma_+^{1,2}), \tag{10}$$

where repeated indices are summed over and where:

$$\mathcal{E}_0 = C + C_1 \epsilon_{zz} + C_2 \epsilon_{||} + D_1 k_z^2 + D_2 k_{||}^2, \tag{11a}$$

$$\mathcal{M} = M + M_1 \epsilon_{zz} + M_2 \epsilon_{||} - B_1 k_z^2 - B_2 k_{||}^2, \tag{11b}$$

$$\alpha_1 = A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} + A_{23}k_z^2 + A_{24}k_{||}^2 + \mathrm{i}Y_3\epsilon_{z-} + Y_4\epsilon_+, \tag{11c}$$

$$\alpha_2 = \mathrm{i}A_2 + \mathrm{i}A_{21}\epsilon_{zz} + \mathrm{i}A_{22}\epsilon_{||} + \mathrm{i}A_{23}k_z^2 + \mathrm{i}A_{24}k_{||}^2 + Y_3\epsilon_{z-} - \mathrm{i}Y_4\epsilon_+, \tag{11d}$$

$$\alpha_3 = Y_1\epsilon_{z+} + \mathrm{i}Y_2\epsilon_-, \tag{11e}$$

$$\beta_1 = X_1\epsilon_{zx} + 2X_2\epsilon_{xy} - \mathrm{i}X_3\epsilon_{zy} + \mathrm{i}X_4(\epsilon_{xx} - \epsilon_{yy}), \tag{11f}$$

$$\beta_2 = X_1\epsilon_{zy} + X_2(\epsilon_{xx} - \epsilon_{yy}) + \mathrm{i}X_3\epsilon_{zx} - 2\mathrm{i}X_4\epsilon_{xy}, \tag{11g}$$

$$\beta_3 = A_1 + A_{11}\epsilon_{zz} + A_{12}\epsilon_{||} + A_{13}k_z^2 + A_{14}k_{||}^2, \tag{11h}$$

with $\epsilon_{||} = \epsilon_{xx} + \epsilon_{yy}, \epsilon_\pm = \epsilon_{xx} - \epsilon_{yy} \pm 2\mathrm{i}\epsilon_{xy}, \epsilon_{z\pm} = \epsilon_{zx} \pm \mathrm{i}\epsilon_{zy}$. From the method of invariants it follows that all model parameters designated by capital roman letters must be real parameters, which are not determined by the symmetries of the crystal. The Hamiltonian above contains all allowed terms to third order in $\mathbf{k}$, first order in strain and terms first order in both. Neglecting strain, this Hamiltonian reduces to the one derived in [12], except for three terms combining the in-plane and out-of-plane components of the wave vector, $k_z^2(k_+\Gamma_-^{1,2} + k_-\Gamma_+^{1,2})$, $k_{||}^2 k_z \Gamma_3$ and $k_z(\mathrm{i}k_-^2\Gamma_-^{1,2} - \mathrm{i}k_+^2\Gamma_+^{1,2})$, which were not included in [12]. The $\mathbf{k}$ independent strain terms depend only on $\epsilon_{||}$ and ϵ_{zz}, and are either proportional to the identity matrix or Γ_5. Hence, the effects of these terms can be described within the unstrained model by making the parameters strain-dependent $M \to \tilde{M}(\boldsymbol{\epsilon})$ and $C \to \tilde{C}(\boldsymbol{\epsilon})$, where:

$$\tilde{M}(\boldsymbol{\epsilon}) = M + M_1\epsilon_{zz} + M_2\epsilon_{||}, \tag{12a}$$

$$\tilde{C}(\boldsymbol{\epsilon}) = C + C_1\epsilon_{zz} + C_2\epsilon_{||}. \tag{12b}$$

Similarly the terms linear in $\mathbf{k}$ and $\epsilon_{||}$ or ϵ_{zz}, can be described by strain-dependent model parameters in the unstrained model. The shear strain terms and the terms with $\epsilon_{xx} - \epsilon_{yy}$ give new terms linear in $\mathbf{k}$, which cannot be described by making the parameters of the unstrained model strain-dependent. The effects of ϵ_{zx} are similar to the effects of ϵ_{xy}, since they contribute to the real parts of β_1, α_2 and α_3 and the imaginary parts of α_1 and β_2. For the same reason $\epsilon_{xx} - \epsilon_{yy}$ and ϵ_{zy} have similar effects. As we demonstrate below, the new terms linear in the wave vector give rise to new physics at the (111) surface of a semi-infinite topological insulator.

Higher order terms in strain will always be both inversion and time-reversal symmetric, and in table 2 we see that the only matrices symmetric under both I and TR are the identity matrix and Γ_5, and k-independent strain terms may therefore be lumped into the strain-dependent parameters $\tilde{M}(\boldsymbol{\epsilon})$ and $\tilde{C}(\boldsymbol{\epsilon})$.

2.3. Magnetic field

The Hamiltonian in equation (10) is invariant under both TR and inversion, thus all bands are doubly degenerate. The TR symmetry can be broken by a magnetic field, which must be included in the following form:

$$H_B = \frac{\mu_B}{2}\begin{pmatrix} g_{1z}B_z & 0 & g_{1p}B_- & 0 \\ 0 & g_{2z}B_z & 0 & g_{2p}B_- \\ g_{1p}B_+ & 0 & -g_{1z}B_z & 0 \\ 0 & g_{2p}B_+ & 0 & -g_{2z}B_z \end{pmatrix}, \tag{13}$$

where $g_{1z,2z,1z,2z}$ are real parameters and $B_\pm = B_x \pm \mathrm{i}B_y$. This contribution was discussed in [12]. The Hamiltonian in equation (13) is responsible for the Zeeman effect. Orbital effects can be included by invoking minimal coupling, $\mathbf{k} \to \mathbf{k} + \frac{e}{\hbar}\mathbf{A}$, with vector potential $\mathbf{A}$, in the strained Hamiltonian equations (10)–(a), and the effects of strain on the Landau levels can be calculated along the lines of Liu *et al* [12]. We shall not pursue these effects further in this work.

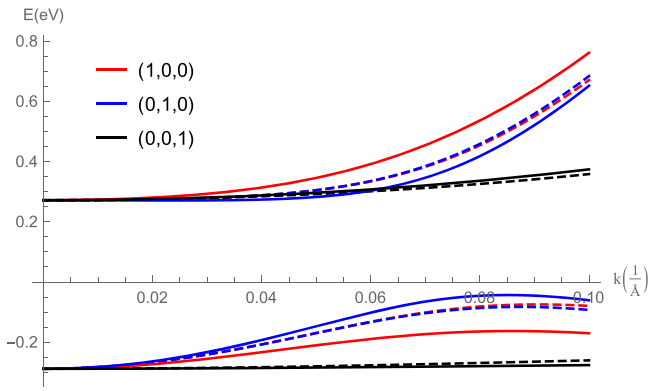

Figure 2. Band structure close to Γ, with $\epsilon_{zy} = 10\%$ (solid) and without strain (dashed) in the absence of electric and magnetic fields. Without strain the in-plane rotational symmetry is broken only by terms to third order in the wave vector, and only a small splitting between different in-plane direction is seen. Including strain allows terms to first order in the wave vector that breaks the in-plane rotational symmetry, and we see that a splitting of the in-plane directions occurs at lower k values. Here we have used the parameters of [12] for the unstrained model, and $X_i = Y_i = 10$ eV Å. We note that $\epsilon_{xx} - \epsilon_{yy} = 10\%$ gives the same dispersion for the parameters used here.

2.4. Electric field

Inversion symmetry can be broken by an electric field, giving rise to a Hamiltonian of the form:

$$\begin{aligned} H_E &= W_{||}(\mathrm{i}E_-\Gamma_+^{25,51} - \mathrm{i}E_+\Gamma_-^{25,51}) + W_z E_z \Gamma_{35} \\ &= \begin{pmatrix} 0 & \mathrm{i}W_z E_z & 0 & \mathrm{i}W_{||}E_- \\ -\mathrm{i}W_z E_z & 0 & -\mathrm{i}W_{||}E_- & 0 \\ 0 & \mathrm{i}W_{||}E_+ & 0 & -\mathrm{i}W_z E_z \\ -\mathrm{i}W_{||}E_+ & 0 & \mathrm{i}W_z E_z & 0 \end{pmatrix}, \end{aligned} \tag{14}$$

where $W_{||,z}$ are real parameters and $E_\pm = E_x \pm \mathrm{i}E_y$.

3. Bulk spectrum and band-gap

In this section we will analyze the effects of strain and electric field on the bulk band structure close to the Γ point. The Hamiltonian $H = H_m + H_E$ of equation (10) and (14) can be diagonalized analytically giving:

$$\begin{aligned} E = \mathcal{E}_0 \pm (&\mathcal{M}^2 + |\alpha_i k_i + \mathrm{i}Z_3 k_z k_-^2|^2 \\ &+ |\beta_i k_i - \mathrm{i}Z_1 \,\mathrm{Re}(k_+^3) + Z_2 \,\mathrm{Im}(k_+^3)|^2 + W_z^2 E_z^2 + W_{||}^2 E_{||}^2 \\ &\pm (2|(\alpha_i k_i + \mathrm{i}Z_3 k_z k_-^2) W_{||} E_+ \\ &+ (\beta_i k_i - \mathrm{i}Z_1 \,\mathrm{Re}(k_+^3) + Z_2 \,\mathrm{Im}(k_+^3)) W_z E_z|^2 \\ &+ 4W_{||}^2 E_{||}^2 \,|\beta_i k_i - \mathrm{i}Z_1 \,\mathrm{Re}(k_+^3) + Z_2 \,\mathrm{Im}(k_+^3)|^2 \\ &+ 4W_z^2 E_z^2 \,|\alpha_i k_i + \mathrm{i}Z_3 k_z k_-^2|^2 \\ &- 2\,\mathrm{Re}((\alpha_i k_i + \mathrm{i}Z_3 k_z k_-^2) W_{||} E_+ \\ &- (\beta_i k_i - \mathrm{i}Z_1 \,\mathrm{Re}(k_+^3) + Z_2 \,\mathrm{Im}(k_+^3)) W_z E_z))^{\frac{1}{2}})^{\frac{1}{2}}. \end{aligned} \tag{15}$$

In figure 2 an example of the band structure with and without strain is plotted. The terms in the Hamiltonian linear in both strain and wave vector split the bands in the x and y directions closer to the Γ point compared to the case without strain, where the bands in the x and y directions are only split by the $\mathbf{k}^3$ terms. The electric field breaks the inversion symmetry, and the degeneracies of the bulk conduction and valence bands are split. At the Γ point the degeneracies are protected by TR symmetry which is not broken. Another important quantity of the electronic structure is the gap between the valence and conduction band at the Γ point, given by:

$$\Delta_\Gamma = 2\sqrt{(M + M_1\epsilon_{zz} + M_2\epsilon_{||})^2 + W_z^2 E_z^2 + W_{||}^2 E_{||}^2}. \tag{16}$$

Note that the gap can only be increased by applying an electric field, whereas strain can increase or decrease the gap depending on the sign of the strain. The effects of strain have been investigated earlier by DFT calculations. In [7] both Bi_2Se_3 and Bi_2Te_3 were investigated. Here the lattice constant in the xy-plane or the z-direction was fixed relative to the known unstrained lattice constant, and the lattice constant in the other direction was relaxed before the calculations were performed. An approximate linear relationship between the in-plane, and out-of-plane

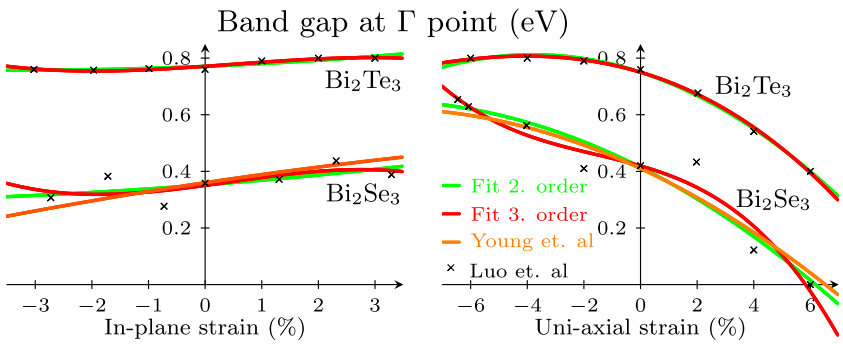

Figure 3. The band gap at Γ point as a function of strain according to the DFT calculations of [7, 8]. The curves from [8] are calculated using the approximate linear relationships between the in-plane and out-of-plane lattice constant reported in [7]. The fits to the data from [7] show good agreement, especially for Bi_2Te_3. Going to third order makes only a slight improvement compared to second order.

Table 3. Parameters used for the calculations in this work. The parameters X_i and Y_i for $i = 1, 2, 3, 4$ are not calculated, but have been chosen to qualitatively show the effect.

C(eV)	−0.0083 [12]
D_1(eVÅ^2)	5.74 [12]
D_2(eVÅ^2)	30.4 [12]
M(eV)	−0.28 [12]
B_1(eVÅ^2)	−6.86 [12]
B_2(eVÅ^2)	−44.5 [12]
A_1(eVÅ)	2.26 [12]
A_2(eVÅ)	3.33 [12]
Z_1(eVÅ^3)	50.6 [12]
Z_2(eVÅ^3)	−113.3 [12]
g_{1z}	−25.4 [12]
g_{1p}	−4.12 [12]
g_{2z}	4.10 [12]
g_{2p}	4.80 [12]
M_1(eV)	2.635 [8]
M_{12}(eV)	15.025 [8]
X_i, Y_i(eVÅ)	10

lattice constants was reported [7]. In [8] the band gap was calculated for many different strain configurations and then fitted to a second order polynomial of the strain components. In figure 3 the results of these calculations are shown together with second and third order fits to the results of [7]. The results show good agreement and both predict a topological phase transition at approximately 6% uniaxial strain. Similar results were reported in [9]. According to our model the topological phase transition will occur when:

$$\epsilon_{zz} = -\frac{M}{M_1} - \frac{M_2}{M_1}\epsilon_{||}. \tag{17}$$

In general, the parameters of this symmetry-adapted model can be determined by using DFT calculations. The values of the parameters are highly dependent on the computational details as can be seen by comparing [1, 12, 17]. In this paper, we use the values from [12] for the parameters not related to strain.

It should be emphasized that strain can only change the gap by a term proportional to the matrix Γ_5 in the Hamiltonian, which is equivalent to simply changing the band gap parameter M. In the following sections we focus on the on the response of the band gap under uniaxial strain, and include a second order contribution as well since this can easily be done using the result of [8]:

$$\tilde{M}(\epsilon_{zz}) = M + M_1\epsilon_{zz} + M_{12}\epsilon_{zz}^2, \tag{18}$$

with the parameters given in table 3.

3.1. Effective masses

An important quantity for transport and optical properties of a solid is the effective mass tensor. Here we calculate the diagonal components of the inverse effective mass tensor including both strain and electric field.

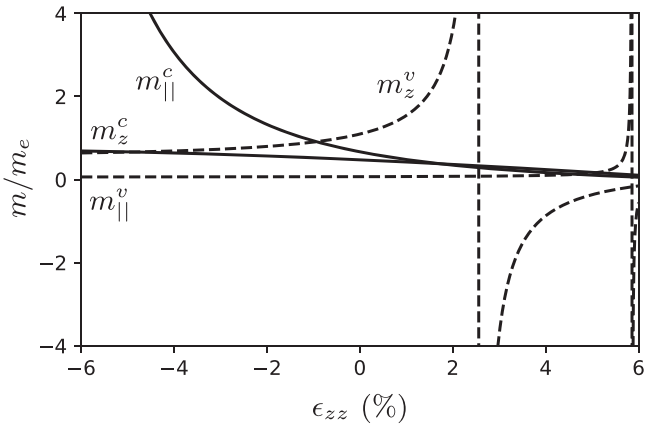

Figure 4. The bulk conduction and valence band effective masses of Bi_2Se_3 plotted as a function of the strain component ϵ_{zz} and in the absence of electric and magnetic fields. In the calculations, we have used the DFT computed strain parameters from [7] and [12] for the parameters not related to strain. The solid (dashed) lines refer to the conduction (valence) band effective masses.

The inverse effective mass tensor is given by:

$$\left[\frac{1}{m^*}\right]_{ij} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k_i \partial k_j}. \tag{19}$$

With $\mathbf{E} = 0$ we get the effective masses along the x and y axes:

$$\left[\frac{\hbar^2}{m}\right]_{ii} = 2D_2 \pm \left(\frac{|\alpha_i|^2 + |\beta_i|^2 - 2(M + M_1\epsilon_{zz} + M_2\epsilon_{||})B_2}{\frac{\Delta_\Gamma}{2}}\right), \tag{20}$$

for $i = x, y$ and along the z-axis

$$\left[\frac{\hbar^2}{m}\right]_{zz} = 2D_1 \pm \left(\frac{|\alpha_3|^2 + |\beta_3|^2 - 2(M + M_1\epsilon_{zz} + M_2\epsilon_{||})B_1}{\frac{\Delta_\Gamma}{2}}\right). \tag{21}$$

The plus (minus) sign corresponds to the conduction (valence) band. For a non-zero electric field each of the two bands is split into two subbands having the same effective mass, but differing by linear and cubic terms in the wave vector. The effective masses including both strain and electric field are given in the appendix.

In figure 4 the bulk effective masses of Bi_2Se_3 are plotted as a function of the strain component ϵ_{zz} and in the absence of electric and magnetic fields. Conduction (valence) band masses are shown as solid (dashed) lines and the effective mass superscript c (v) refers to the conduction (valence) band. In this case where only $\epsilon_{zz} \neq 0$ and the electric field is zero the effective masses along the x and y directions are the same. Note that since the maximum of the valence band shifts away from the Γ point at a sufficiently large strain value, the valence band effective masses change as a function of strain from positive to negative and reach infinite values in-between.

4. SS spectrum and spin structure

We shall now proceed to calculate the SS spectrum and spin structure for the model Hamiltonian (10). In this and the following section we exclude $\mathbf{k}^3$ terms, i.e. the last three terms in equation (10) and the $\mathbf{k}$ dependent terms in α_i and β_i. First, we consider a semi-infinite topological insulator filling the $z < 0$ half space, implemented via the boundary conditions $\Psi(0) = 0$ and $\Psi(z) \to 0$ for $z \to -\infty$. The translational invariance is broken in the z-direction and we make the substitution $k_z \to -i\partial_z$. Using the ansatz $\psi_\lambda e^{z\lambda}$ the stationary Schrödinger equation implies the following secular equation:

$$\begin{aligned} 0 = {} & D_+D_-\lambda^4 + (2(E - \tilde{C}(\boldsymbol{\epsilon}) - D_2k_{||}^2)D_1 \\ & - 2(\tilde{M}(\boldsymbol{\epsilon}) - B_2k_{||}^2)B_1 + |\alpha_3|^2 + |\beta_3|^2)\lambda^2 \\ & + i(2\mathrm{Re}(\alpha_1\alpha_3^* + \beta_1\beta_3^*)k_x + 2\mathrm{Re}(\alpha_2\alpha_3^* + \beta_2\beta_3^*)k_y)\lambda \\ & + (E - \tilde{C}(\boldsymbol{\epsilon}) - D_2k_{||}^2)^2 - (\tilde{M}(\boldsymbol{\epsilon}) - B_2k_{||}^2)^2 \\ & - |\alpha_1k_x + \alpha_2k_y|^2 - |\beta_1k_x + \beta_2k_y|^2, \end{aligned} \tag{22}$$

where $D_\pm = D_1 \pm B_1$. This equation has 4 complex solutions λ_i, each with two corresponding eigenspinors $\psi_j(\lambda_i)$ given by:

$$\psi_1(\lambda_i) = \begin{pmatrix} E - \mathcal{E}_0 + \mathcal{M} \\ \beta_1 k_x + \beta_2 k_y - \mathrm{i}\beta_3 \lambda_i \\ 0 \\ \alpha_1 k_x + \alpha_2 k_y - \mathrm{i}\alpha_3 \lambda_i \end{pmatrix}, \tag{23}$$

$$\psi_2(\lambda_i) = \begin{pmatrix} \beta_1^* k_x + \beta_2^* k_y - \mathrm{i}\beta_3^* \lambda_i \\ E - \mathcal{E}_0 - \mathcal{M} \\ \alpha_1 k_x + \alpha_2 k_y - \mathrm{i}\alpha_3 \lambda_i \\ 0 \end{pmatrix}. \tag{24}$$

The general solution to the Schrödinger equation can then be written on the form:

$$\Psi(z) = \sum_{j=1}^{2}\sum_{i=1}^{4} C_{ij}\psi_j(\lambda_i)\mathrm{e}^{\lambda_i z}, \tag{25}$$

where the coefficients C_{ij} are to be determined from the boundary conditions. For SS we require that $\Psi(z) \to 0$ for $z \to -\infty$, implying that $\mathrm{Re}(\lambda_i) > 0$. To satisfy the boundary condition at the surface we need two different solutions with positive real part. By complex conjugation of equation (22) we see that for a solution λ_i, $-\lambda_i^*$ will also be a solution. Hence, the solutions are either imaginary or occur in pairs related by $\lambda_i = -\lambda_j^*$. This allows for three distinct cases: (i) All solutions are imaginary. (ii) Two complex solutions related by $\lambda_1 = -\lambda_2^*$ and two imaginary solutions. (iii) Two pairs of complex solutions $\lambda_1 = -\lambda_3^*$ and $\lambda_2 = -\lambda_4^*$. Existence of SSs is only possible in case (iii) and we therefore assume that λ_1 and λ_2 have positive real parts. Hence, we limit the sum in equation (25) to $i \in \{1, 2\}$. Applying the boundary condition at $z = 0$ we obtain the following secular equation for nontrivial solution to the coefficients C_{ij}:

$$(\lambda_1 + \lambda_2)^2 = \frac{|\alpha_3|^2 + |\beta_3|^2}{-D_+D_-}, \tag{26}$$

and combining this with equation (22) we find the SS spectrum:

$$\begin{aligned} E = {} & \tilde{C}(\boldsymbol{\epsilon}) + \frac{D_1}{B_1}\tilde{M}(\boldsymbol{\epsilon}) \\ & \pm \sqrt{1 - \frac{D_1^2}{B_1^2}\left(|\alpha_1 k_x + \alpha_2 k_y|^2 + |\beta_1 k_x + \beta_2 k_y|^2 \right.} \\ & \left. - \frac{\mathrm{Re}((\alpha_1 k_x + \alpha_2 k_y)\alpha_3^* + (\beta_1 k_x + \beta_2 k_y)\beta_3^*)^2}{|\alpha_3|^2 + |\beta_3|^2}\right)^{\frac{1}{2}} \\ & + \left(D_2 - \frac{D_1}{B_1}B_2\right)k_{||}^2. \end{aligned} \tag{27}$$

For small $k_{||}$ we see that the spectrum is linear but due to the strain, the group velocity has now become anisotropic and the contours of constant energy have become elliptical. Note that the position of the Dirac-point can be shifted by changing $\tilde{M}(\boldsymbol{\epsilon})$, however the relative position within the bulk band gap is not changed. Notice also that the second order term is not changed by strain, since we have systematically retained terms of order $\epsilon^1 k^1$ and discarded terms of order $\epsilon^1 k^2$. Both the directional dependence of the group velocity and the ellipticity of the contours of constant energy are shown in figure 5.

For $k_{||} = 0$, the linear term in equation (22) vanishes and λ_1 and λ_2 can be calculated analytically:

$$\lambda_\alpha = \sqrt{\frac{-F + (-1)^\alpha\sqrt{R}}{2D_+D_-}}, \tag{28}$$

where:

$$F = |\beta_3|^2 + |\alpha_3|^2 - 2\tilde{M}(\boldsymbol{\epsilon})B_1 + 2ED_1 - 2\tilde{C}(\boldsymbol{\epsilon})D_1, \tag{29a}$$

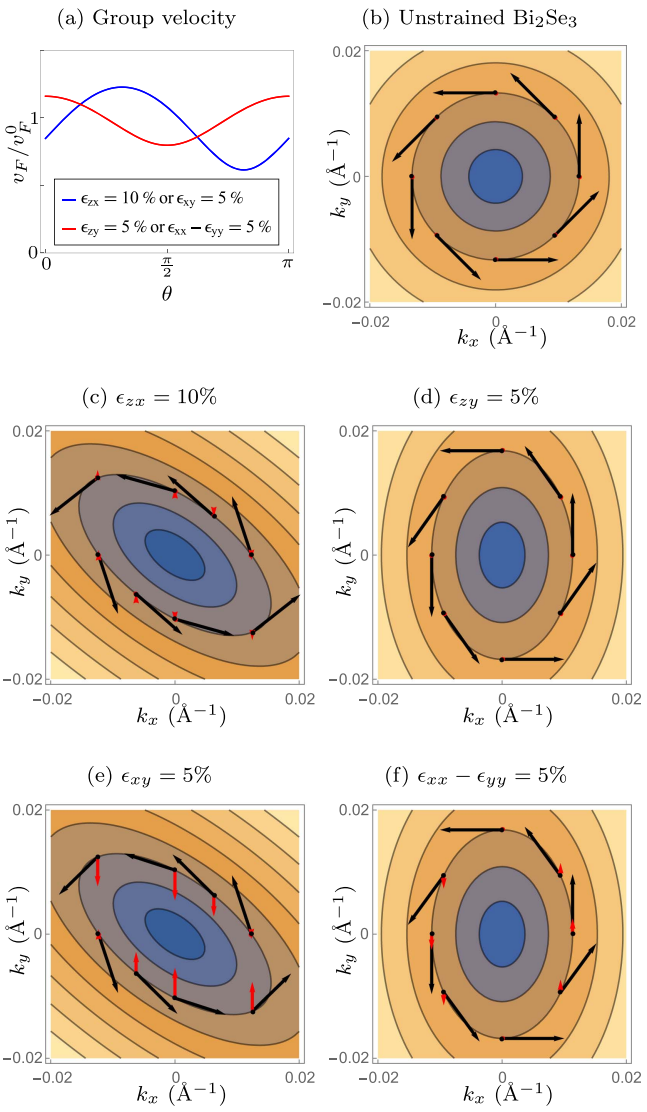

Figure 5. (a) Group velocity relative to the velocity in unstrained Bi_2Se_3 given by $v_F^0 = A_2\sqrt{1 - \frac{D_1^2}{B_1^2}}$ as a function of the angle $\theta = \arctan(\frac{k_y}{k_x})$ in the k_x, k_y plane. (b)–(f) Contour plots showing the elliptical curves of constant energy of the upper Dirac cone of the surface states in the (k_x, k_y) plane. The black arrows indicate the direction in-plane part of the expectation value of the spin, while the red arrows show the expectation value of the spin in the z-direction. Spin-momentum locking is still present, but spin and momentum are no longer perpendicular. Here we have used the parameters of [12] for the parameters of the unstrained model, and $X_i = Y_i = 10$ eV Å.

$$R = F^2 - 4D_+D_-(E - \tilde{C}(\epsilon) + \tilde{M}(\epsilon))(E - \tilde{C}(\epsilon) - \tilde{M}(\epsilon)). \tag{29b}$$

Imposing the boundary condition $\Psi(0) = 0$ with equation (25) we find the exact wave functions of the two degenerate eigenstates at $k_{||} = 0$, which are related by TR:

$$\Psi_1 = N\begin{pmatrix} i\,\mathrm{sgn}(D_+)\sqrt{\frac{|D_+|}{|2B_1|}} \\ \frac{\beta_3}{\sqrt{|\alpha_3|^2 + |\beta_3|^2}}\sqrt{\frac{|D_-|}{|2B_1|}} \\ 0 \\ \frac{\alpha_3}{\sqrt{|\alpha_3|^2 + |\beta_3|^2}}\sqrt{\frac{|D_-|}{|2B_1|}} \end{pmatrix}(e^{\lambda_1 z} - e^{\lambda_2 z}), \tag{30a}$$

$$\Psi_2 = N\begin{pmatrix} 0 \\ \frac{\alpha_3^*}{\sqrt{|\alpha_3|^2 + |\beta_3|^2}}\sqrt{\frac{|D_-|}{|2B_1|}} \\ \mathrm{i}\,\mathrm{sgn}(D_+)\sqrt{\frac{|D_+|}{|2B_1|}} \\ \frac{-\beta_3^*}{\sqrt{|\alpha_3|^2 + |\beta_3|^2}}\sqrt{\frac{|D_-|}{|2B_1|}} \end{pmatrix}(\mathrm{e}^{\lambda_1 z} - \mathrm{e}^{\lambda_2 z}). \tag{30b}$$

Note that the α_3 term couples the spin blocks and in contrast to the case without strain the eigenstates at the $k_{||} = 0$ point are mixed up/down spinstates. The conditions for existence of SS at $k_{||} = 0$ are:

$$D_+ D_- < 0 \quad \text{and} \quad \tilde{M}(\boldsymbol{\epsilon}) B_1 > 0. \tag{31}$$

As expected, the SS can only be destroyed by changing the sign of $\tilde{M}(\boldsymbol{\epsilon})$, i.e. closing and reopening the bulk band gap.

Due to TR-symmetry the spin of the SS at (k_x, k_y) must be opposite of the spin of the SS at $(-\mathrm{k}_x, -\mathrm{k}_y)$. For a (111) surface without strain, the model Hamiltonian is invariant under rotations in the xy-plane of any angle. Hence $\langle S_z \rangle = 0$ since (k_x, k_y) and $(-k_x, -k_y)$ are related by a rotation around the z-axis which does not change S_z. The terms linear in both wave vector and strain break the full rotation symmetry, and non-zero $\langle S_z \rangle$ values are allowed. At finite $k_{||}$, equation (22) was solved numerically, and using $\Psi(0) = 0$ a numerical expression for the SS spinor was determined. Finally, the expectation values of the spin operators were calculated and are shown in figure 5. As expected, non-zero values of $\langle S_z \rangle$ occur. Such out-of-plane spin components also show up in the absence of strain to third order in k in the case of hexagonal warping [2] and has been experimentally verified in [18]. The actual values of S_z shown in figure 5(a) are of course strongly dependent on the chosen parameters, but the non-zero S_z-component is generally present for any values of X_i and Y_i.

4.1. Effective 2D model

Using the SS at $k_{||} = 0$ as basis we can derive an effective 2D model for the SS. We use the wave functions in equation (30), but with $\epsilon_{ij} = 0$ to have a basis independent of strain. We note that in the absence of strain, ψ_1 and ψ_2 represent spin up/down states respectively. The matrix elements of the Γ-matrices in this basis read:

$$\langle \Psi_i | \Gamma_1 | \Psi_j \rangle = -\mathrm{sgn}(B_1 A_1)\sqrt{1 - \frac{D_1^2}{B_1^2}}\,[\sigma_y]_{ij}, \tag{32a}$$

$$\langle \Psi_i | \Gamma_2 | \Psi_j \rangle = \mathrm{sgn}(B_1 A_1)\sqrt{1 - \frac{D_1^2}{B_1^2}}\,[\sigma_x]_{ij}, \tag{32b}$$

$$\langle \Psi_i | \Gamma_3 | \Psi_j \rangle = 0, \tag{32c}$$

$$\langle \Psi_i | \Gamma_4 | \Psi_j \rangle = -\mathrm{sgn}(B_1 A_1)\sqrt{1 - \frac{D_1^2}{B_1^2}}\,[\sigma_z]_{ij}, \tag{32d}$$

$$\langle \Psi_i | \Gamma_5 | \Psi_j \rangle = \frac{D_1}{B_1}[\sigma_0]_{ij}. \tag{32e}$$

The effective 2D model including $\mathbf{k}$ up to second order becomes:

$$H^{2D} = \tilde{C}(\boldsymbol{\epsilon}) + \frac{D_1}{B_1}\tilde{M}(\boldsymbol{\epsilon}) + \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma} + \left(D_2 - \frac{D_1}{B_1}B_2\right)k_{||}^2, \tag{33}$$

where $\mathbf{d}(\mathbf{k}) = (d_1(\mathbf{k}), d_2(\mathbf{k}), d_3(\mathbf{k}))$ is given by:

$$\begin{aligned} d_1(\mathbf{k}) &= \sqrt{1 - \frac{D_1^2}{B_1^2}}\,\mathrm{Im}(\alpha_1 k_x + \alpha_2 k_y) \\ &= \sqrt{1 - \frac{D_1^2}{B_1^2}}\,((Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy})k_x \\ &\quad + (A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} - Y_3\epsilon_{zy} - Y_4(\epsilon_{xx} - \epsilon_{yy}))k_y), \end{aligned} \tag{34a}$$

$$d_2(\mathbf{k}) = -\sqrt{1 - \frac{D_1^2}{B_1^2}}\,\mathrm{Re}(\alpha_1 k_x + \alpha_2 k_y)$$
$$= -\sqrt{1 - \frac{D_1^2}{B_1^2}}\,((Y_3\epsilon_{zy} + 2Y_4\epsilon_{xy})k_y + (A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} + Y_3\epsilon_{zy} + Y_4(\epsilon_{xx} - \epsilon_{yy}))k_x), \quad (34b)$$

$$d_3(\mathbf{k}) = -\sqrt{1 - \frac{D_1^2}{B_1^2}}\,\mathrm{Im}(\beta_1 k_x + \beta_2 k_y)$$
$$= \sqrt{1 - \frac{D_1^2}{B_1^2}}\,((X_3\epsilon_{zy}(\epsilon_{xx} - \epsilon_{yy}))k_x + (-X_3\epsilon_{zx} + 2X_4\epsilon_{xy})k_y). \quad (34c)$$

Diagonalizing the effective Hamiltonian gives the spectrum:

$$E = \tilde{C}(\boldsymbol{\epsilon}) + \frac{D_1}{B_1}\tilde{M}(\boldsymbol{\epsilon}) + \left(D_2 - \frac{D_1}{B_1}B_2\right)k_{||}^2$$
$$\pm \sqrt{1 - \frac{D_1^2}{B_1^2}}\,\sqrt{|\alpha_1 k_x + \alpha_2 k_y|^2 + \mathrm{Im}(\beta_1 k_x + \beta_2 k_y)^2}. \quad (35)$$

This spectrum is different from equation (27), obtained from the 3D model. The two spectra become equal, however, if we take $\alpha_3 = 0$. We note that the Berry curvature of the SS band vanishes identically, i.e. $\Omega(\mathbf{k}) = \hat{d} \cdot [(\partial_{k_x}\hat{d}) \times (\partial_{k_y}\hat{d})] = 0$.

5. Localization of SS and finite size effect

In the case of a finite slab, a gap will open in the SS spectrum at $k_{||} = 0$. This is due to the overlap of SS on opposite surfaces. For a given thickness the overlap will be highly dependent on the penetration depth of the SS. From the wave functions the expectation value of the distance from the surface $d = -\langle z \rangle$ can be calculated to:

$$d = \frac{\frac{B_1}{2\tilde{M}(\boldsymbol{\epsilon})} + \frac{-D_+D_-}{|\alpha_3|^2 + |\beta_3|^2}}{\sqrt{\frac{-D_+D_-}{|\alpha_3|^2 + |\beta_3|^2}}}, \quad (36)$$

and we see that d depends on strain through the band gap parameter, and the quantity $\sqrt{|\alpha_3|^2 + |\beta_3|^2}$ involving the coefficients of the terms first order in k_z.

The eigenenergies in the finite system can be calculated by imposing the boundary condition $\Psi(\pm L/2) = 0$, where L is the slab thickness. Now the wave function can be written as equation (25), but using all 4 solutions to equation (22). At the Γ point, the solutions are $\pm\lambda_{1,2}$ with $\lambda_{1,2}$ given by equation (28). The secular equation of the nontrivial solution for the coefficients gives:

$$\frac{(E - \tilde{C}(\boldsymbol{\epsilon}) + \tilde{M}(\boldsymbol{\epsilon}) + D_+\lambda_1^2)\lambda_2}{(E - \tilde{C}(\boldsymbol{\epsilon}) + \tilde{M}(\boldsymbol{\epsilon}) + D_+\lambda_2^2)\lambda_1} = \left(\frac{\tanh(\lambda_1 L/2)}{\tanh(\lambda_2 L/2)}\right)^{\pm 1}, \quad (37)$$

for even/odd states respectively. This is the same form as found using the model without strain [19] , but here the parameters $\tilde{M}(\boldsymbol{\epsilon})$ and $\tilde{C}(\boldsymbol{\epsilon})$ as well as $\lambda_{1,2}$ from equation (28) depend on strain. The parameter $\tilde{C}(\boldsymbol{\epsilon})$ is not important, since it simply shifts all energies. Again the important quantities are $\tilde{M}(\boldsymbol{\epsilon})$ and $\sqrt{|\alpha_3|^2 + |\beta_3|^2}$.

To find the gap of the SSs, Δ, equation (37) has been solved numerically for $L = n \cdot 9.547$ Å for integers n between 2 and 6 where $L = 9.547$ Å is the thickness of 1 QL. First we have analyzed the effect of the bulk band gap changed by strain, and taking $\sqrt{|\alpha_3|^2 + |\beta_3|^2} = A_1$ the value without strain. We see that increasing the bulk band gap, decreases the SS band gap. Close to the phase transition this behavior is evident, since the SS must approach bulk states extending further into the material as the bulk gap approaches zero, giving a large coupling between opposite surfaces. As we saw in section 3, the bulk gap is most sensitive to ϵ_{zz}. Using equation (18) we have plotted the SS gap as function of ϵ_{zz} in figure 6(a). A similar result was reported in [9] for a slab of 8 QL using DFT, where a gap opens when tensile strain is applied. Our results points to the physical origin of this effect being the modification of the spatial distribution of the SS wave functions.

In figure 6(b), we have plotted the dependence of the SS gap on $\sqrt{|\alpha_3|^2 + |\beta_3|^2}$ assuming a constant bulk band gap parameter of $M = 0.28$ eV from [12] and setting $M_1 = M_2 = 0$. Without strain $\sqrt{|\alpha_3|^2 + |\beta_3|^2} = A_1 = 2.26$ eVÅ according to [12], but the quantitative dependence on strain has not been determined. We see that the SS gap closes at low values, which is related to the fact that for

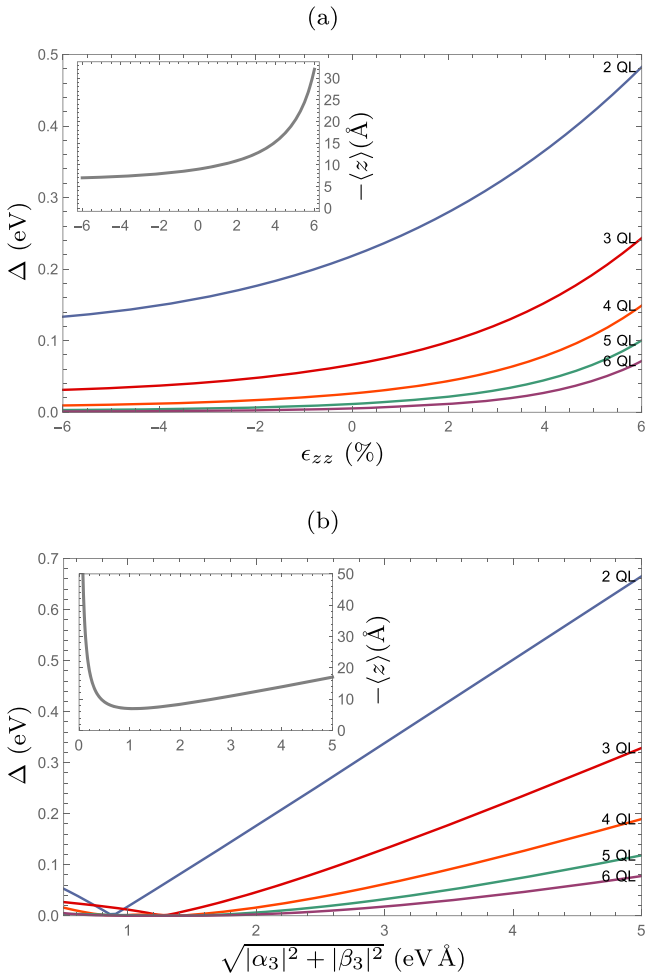

Figure 6. Band gap for thicknesses of 2 to 6 quintuple layers (QL). In (a) the effect of changing the bulk band gap by uniaxial strain on the SS gap Δ is shown. We have used a relation between the bulk band gap and ϵ_{zz} from [8]. In this plot we have neglected the strain dependence of $\sqrt{|\alpha_3|^2 + |\beta_3|^2}$ by setting $A_{11} = 0$. In (b) we show the SS gap as a function of $\sqrt{|\alpha_3|^2 + |\beta_3|^2}$, assuming a constant bulk band gap, i.e., setting $M_1 = M_2 = 0$. The insets shows the expectation value of the distance to the surface for the SS in the semi-infinite case, with the same horizontal axis. A qualitative agreement between the localization of the SS on single surface and the gap induced by coupling between opposite surfaces is seen. We have used the parameters of the model without strain from [12].

Table 4. Experimental measurements of the surface state gap, induced by coupling between opposite surfaces. First row is from [24], second row is from [25]. In the second row the gap increases from 2 to 3 QL, signifying an oscillatory dependence of the gap on the thickness.

L	2 QL	3 QL	4 QL	5 QL
Δ(eV)	0.25	0.138	0.07	0.041
Δ(eV)	0.28	0.34		

$|\alpha_3|^2 + |\beta_3|^2 < \frac{4\tilde{M} \mid D_+ D_- \mid}{B_1}$ the solutions to equation (22), λ_i, are complex giving an oscillatory spatial dependence of the SS. Oscillatory SS give oscillations in the SS gap as a function of L, which has been theorized to signify topological phase transitions between a trivial 2D insulator and a quantum spin hall phase [20–22]. We note here that these oscillations are present only using the parameters of [1], not the parameters of [12] as shown in [23]. Experimental evidence of oscillations in the gap remains inconclusive, since both a simple decay of the gap for 2–5 QL has been reported [24] as well as an increase in the gap from 2 to 3 QL [25] signifying a topological phase transition between 2 and 3 QL. The observed gaps are summarized in table 4.

We have demonstrated that the gap of the SS can be tuned by strain. However, quantitative predictions require access to the parameters relating the strain tensor to $\sqrt{|\alpha_3|^2 + |\beta_3|^2}$. Tuning the band gap of the SS is important for both applications and fundamental research. One particular experimental challenge is that the layered structure of Bi_2Se_3 makes it easy to fabricate integer numbers of QL. Thus the gap dependence on the layer thickness has only been experimentally studied with few data points, since the gap becomes too small to be measured already at 6 QL [24]. With strain, it is be possible to increase the SS gap and investigate the thickness dependence of the gap further. Another intriguing prospect would be tuning the gap via strain, so as to pass through the topological transition in a controlled manner.

6. Conclusion

We have derived the most general Hamiltonian for the Bi_2Se_3-class of materials including terms to third order in the wave vector, first order in electric and magnetic fields, first order in strain and first order in both strain and wave vector. We show that this model provides a description of a range of different effects of strain on the electronic structure of these materials. Specifically we have analytically derived the spectrum of the SS for a semi-infinite topological insulator, showing qualitatively the effects of strain on both the spectrum and the spin structure. The terms first order in both wave vector and strain break the full rotational symmetry close to $k_{||} = 0$, leading to an anisotropy in the Dirac cone otherwise absent for the simple (111) surface termination. The spin structure is altered as well, and for some strain configurations a spin component perpendicular to the surface arises. In an analysis of the finite size effect we show that increasing the bulk band gap by virtue of strain, decreases the induced gap in the SS. Strain-tuning the SS gap gives new possibilities for experimental investigation of the thickness dependence of the SS gap and control of transport and optical properties.

Acknowledgments

MRB and MW gratefully acknowledge financial support from the Danish Council of Independent Research (Natural Sciences) grant no.: DFF-4181-00182. AML gratefully acknowledges financial support from the Carlsberg Foundation. The Center for Quantum Devices is funded by the Danish National Research Foundation.

Appendix. Effective masses

Here we give the expressions for the bulk effective masses including both contributions from strain and electric fields. Upper (lower) signs refer to the conduction (valence) band.

$$\begin{aligned}\frac{\hbar^2}{m_x} = 2D_2 \pm \frac{2}{\Delta_\Gamma}&(-2(M + M_1\epsilon_{zz} + M_2\epsilon_{||})B_2 + (A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} + Y_3\epsilon_{zy} + Y_4(\epsilon_{xx} - \epsilon_{yy}))^2 \\ &+ (Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy})^2 + (X_1\epsilon_{zx} + 2X_2\epsilon_{xy})^2 + (-X_3\epsilon_{zy} + X_4(\epsilon_{xx} - \epsilon_{yy}))^2 \\ &- \frac{4}{\Delta_\Gamma^2}(4E_z^2W_z^2((A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} + Y_3\epsilon_{zy} + Y_4(\epsilon_{xx} - \epsilon_{yy}))^2 + (Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy})^2) \\ &+ 4W_{||}^2E_{||}^2((X_1\epsilon_{zx} + 2X_2\epsilon_{xy})^2 + (-X_3\epsilon_{zy} + X_4(\epsilon_{xx} - \epsilon_{yy}))^2) \\ &+ (W_{||}E_x(A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} + Y_3\epsilon_{zy} + Y_4(\epsilon_{xx} - \epsilon_{yy})) - W_{||}E_y(Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy}) \\ &+ W_zE_z(X_1\epsilon_{zx} + 2X_2\epsilon_{xy}))^2 \\ &+ (W_{||}E_x(Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy}) + W_{||}E_y(A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} + Y_3\epsilon_{zy} + Y_4(\epsilon_{xx} - \epsilon_{yy})) \\ &+ W_zE_z(-X_3\epsilon_{zy} + X_4(\epsilon_{xx} - \epsilon_{yy})))^2 \\ &- 2(W_{||}E_x(A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} + Y_3\epsilon_{zy} + Y_4(\epsilon_{xx} - \epsilon_{yy})) - W_{||}E_y(Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy}) \\ &- W_zE_z(X_1\epsilon_{zx} + 2X_2\epsilon_{xy}))^2 \\ &+ 2(W_{||}E_x(Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy}) + W_{||}E_y(A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} + Y_3\epsilon_{zy} + Y_4(\epsilon_{xx} - \epsilon_{yy})) \\ &- W_zE_z(-X_3\epsilon_{zy} + X_4(\epsilon_{xx} - \epsilon_{yy})))^2)),\end{aligned} \tag{A1}$$

$$\begin{aligned}\frac{\hbar^2}{m_y} = {} & 2D_2 \pm \frac{2}{\Delta_\Gamma}(-2(M + M_1\epsilon_{zz} + M_2\epsilon_{||})B_2 + (Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy})^2 \\ & + (A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} - Y_3\epsilon_{zy} - Y_4(\epsilon_{xx} - \epsilon_{yy}))^2 + (X_1\epsilon_{zy} + X_2(\epsilon_{xx} - \epsilon_{yy}))^2 + (X_3\epsilon_{zx} - 2X_4\epsilon_{xy})^2 \\ & - \frac{4}{\Delta_\Gamma^2}(4E_z^2W_z^2((Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy})^2 + (A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} - Y_3\epsilon_{zy} - Y_4(\epsilon_{xx} - \epsilon_{yy}))^2) \\ & + 4W_{||}^2E_{||}^2((X_1\epsilon_{zy} + X_2(\epsilon_{xx} - \epsilon_{yy}))^2 + (X_3\epsilon_{zx} - 2X_4\epsilon_{xy})^2) \\ & + (W_{||}E_x(Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy}) - W_{||}E_y(A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} - Y_3\epsilon_{zy} - Y_4(\epsilon_{xx} - \epsilon_{yy})) \\ & + W_zE_z(X_1\epsilon_{zy} + X_2(\epsilon_{xx} - \epsilon_{yy})))^2 \\ & + (W_{||}E_x(A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} - Y_3\epsilon_{zy} - Y_4(\epsilon_{xx} - \epsilon_{yy})) + W_{||}E_y(Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy}) \\ & + W_zE_z(X_3\epsilon_{zx} - 2X_4\epsilon_{xy}))^2 \\ & - 2(W_{||}E_x(Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy}) - W_{||}E_y(A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} - Y_3\epsilon_{zy} - Y_4(\epsilon_{xx} - \epsilon_{yy})) \\ & - W_zE_z(X_1\epsilon_{zy} + X_2(\epsilon_{xx} - \epsilon_{yy})))^2 \\ & + 2(W_{||}E_x(A_2 + A_{21}\epsilon_{zz} + A_{22}\epsilon_{||} - Y_3\epsilon_{zy} - Y_4(\epsilon_{xx} - \epsilon_{yy})) + W_{||}E_y(Y_3\epsilon_{zx} + 2Y_4\epsilon_{xy}) \\ & - W_zE_z(X_3\epsilon_{zx} - 2X_4\epsilon_{xy}))^2)),\end{aligned} \tag{A2}$$

$$\begin{aligned}\frac{\hbar^2}{m_z} = {} & 2D_1 \pm \frac{2}{\Delta_\Gamma}(-2(M + M_1\epsilon_{zz} + M_2\epsilon_{||})B_1 + (Y_1\epsilon_{zx} + 2Y_2\epsilon_{xy})^2 + (Y_1\epsilon_{zy} + Y_2(\epsilon_{xx} - \epsilon_{yy}))^2 \\ & + (A_1 + A_{11}\epsilon_{zz} + A_{12}\epsilon_{||})^2 \\ & - \frac{4}{\Delta_\Gamma^2}(4E_z^2W_z^2((Y_1\epsilon_{zx} + 2Y_2\epsilon_{xy})^2 + (Y_1\epsilon_{zy} + Y_2(\epsilon_{xx} - \epsilon_{yy}))^2) + 4W_{||}^2E_{||}^2(A_1 + A_{11}\epsilon_{zz} + A_{12}\epsilon_{||}) \\ & + (W_{||}E_x(Y_1\epsilon_{zx} + 2Y_2\epsilon_{xy}) - W_{||}E_y(Y_1\epsilon_{zy} + Y_2(\epsilon_{xx} - \epsilon_{yy})) + W_zE_z(A_1 + A_{11}\epsilon_{zz} + A_{12}\epsilon_{||}))^2 \\ & + (W_{||}E_x(Y_1\epsilon_{zy} + Y_2(\epsilon_{xx} - \epsilon_{yy})) + W_{||}E_y(Y_1\epsilon_{zx} + 2Y_2\epsilon_{xy}))^2 \\ & - 2(W_{||}E_x(Y_1\epsilon_{zx} + 2Y_2\epsilon_{xy}) - W_{||}E_y(Y_1\epsilon_{zy} + Y_2(\epsilon_{xx} - \epsilon_{yy})) - W_zE_z(A_1 + A_{11}\epsilon_{zz} + A_{12}\epsilon_{||}))^2 \\ & + 2(W_{||}E_x(Y_1\epsilon_{zy} + Y_2(\epsilon_{xx} - \epsilon_{yy})) + W_{||}E_y(Y_1\epsilon_{zx} + 2Y_2\epsilon_{xy}))^2)).\end{aligned} \tag{A3}$$

References

[1] Zhang H, Liu C, Qi X, Dai X, Fang Z and Zhang S 2009 *Nat. Phys.* **5** 438
[2] Fu L 2009 *Phys. Rev. Lett.* **103** 266801
[3] Moon C-Y, Han J, Lee H and Choi H J 2011 *Phys. Rev.* B **84** 195425
[4] Trescher M, Sbierski B, Brouwer P W and Bergholtz E J 2015 *Phys. Rev.* B **91** 115135
[5] Feng Y *et al* 2014 *Sci. Rep.* **4** 5385
[6] Lu H-Y, Cuamba A S, Lin S-Y, Hao L, Wang R, Li H, Zhao Y Y and Ting C S 2016 *Phys. Rev.* B **94** 195423
[7] Luo X, Sullivan M B and Quek S Y 2012 *Phys. Rev.* B **86** 184111
[8] Young S M, Chowdhury S, Walter E J, Mele E J, Kane C L and Rappe A M 2011 *Phys. Rev.* B **84** 085106
[9] Liu W, Peng X, Tang C, Sun L, Zhang K and Zhong J 2011 *Phys. Rev.* B **84** 245105
[10] Li B, Chen W, Guo X, Ho W, Dai X, Jia J and Xie M 2017 *Appl. Surf. Sci.* **396** 1825
[11] Lew L C, Voon Y and Willatzen M 2009 *The k.p Method* (Berlin: Springer)
[12] Liu C-X, Qi X-L, Zhang H J, Dai X, Fang Z and Zhang S-C 2010 *Phys. Rev.* B **82** 045122
[13] Brems M R, Paaske J, Lunde A M and Willatzen M 2018 *Phys. Rev.* B **97** 081402
[14] Silvestrov P G, Brouwer P W and Mishchenko E G 2012 *Phys. Rev.* B **86** 075302
[15] Bir G L and Pikus G E 1974 *Symmetry and Strain-Induced Effects in Semiconductors* (New York: Wiley)
[16] Koster G F, Dimmock J O, Wheeler R G and Statz H 1963 *Properties of the Thirty-Two Point Groups* (Cambridge, MA: MIT Press)
[17] Nechaev I A and Krasovskii E E 2016 *Phys. Rev.* B **94** 201410
[18] Souma S, Kosaka K, Sato T, Komatsu M, Takayama A, Takahashi T, Kriener M, Segawa K and Ando Y 2011 *Phys. Rev. Lett.* **106** 216803
[19] Shan W-Y, Lu H-Z and Shen S-Q 2010 *New J. Phys.* **12** 043048
[20] Lu H-Z, Shan W-Y, Yao W, Niu Q and Shen S-Q 2010 *Phys. Rev.* B **81** 115407
[21] Liu C-X, Zhang H J, Yan B, Qi X-L, Frauenheim T, Dai X, Fang Z and Zhang S-C 2010 *Phys. Rev.* B **81** 041307
[22] Linder J, Yokoyama T and Sudbø A 2009 *Phys. Rev.* B **80** 205401
[23] Jensen M R 2014 Understanding the low energy physics of bismuth selenide: a three-dimensional topological insulator *Master's Thesis* University of Copenhagen
[24] Zhang Y *et al* 2010 *Nat. Phys.* **6** 712
[25] Sakamoto Y, Hirahara T, Miyazaki H, Kimura S I and Hasegawa S 2010 *Phys. Rev.* B **81** 165432