

Modelling near-surface bound electron states in a 3D topological insulator: analytical and numerical approaches

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 J. Phys.: Condens. Matter 26 485003

(<http://iopscience.iop.org/0953-8984/26/48/485003>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 158.227.172.141

This content was downloaded on 24/11/2014 at 09:55

Please note that [terms and conditions apply](#).

Modelling near-surface bound electron states in a 3D topological insulator: analytical and numerical approaches

V N Men'shov^{1,2,3}, V V Tugushev^{1,2,3,4}, T V Menshchikova^{1,3},
S V Eremeev^{1,3,5}, P M Echenique^{1,6} and E V Chulkov^{1,6}

¹ Donostia International Physics Center (DIPC), P. de Manuel Lardizabal 4, 20018, San Sebastián, Basque Country, Spain

² NRC Kurchatov Institute, Kurchatov Sqr. 1, 123182 Moscow, Russia

³ Tomsk State University, prospekt Lenina, 40, 634050 Tomsk, Russia

⁴ Prokhorov General Physics Institute, Vavilov Str. 38, 119991 Moscow, Russia

⁵ Institute of Strength Physics and Materials Science, 634021 Tomsk, Russia

⁶ Departamento de Física de Materiales UPV/EHU, CFM—MPC UPV/EHU, 20080 San Sebastián/Donostia, Basque Country, Spain

E-mail: vnmenshov@mail.ru

Received 28 August 2014

Accepted for publication 23 September 2014

Published 23 October 2014

Abstract

We apply both analytical and *ab-initio* methods to explore heterostructures composed of a 3D topological insulator (3D TI) and an ultrathin normal insulator (NI) overlayer as a proving ground for the principles of topological phase engineering. Using the continual model of a semi-infinite 3D TI we study the surface potential (SP) effect caused by an attached ultrathin layer of 3D NI on the formation of topological bound states at the interface. The results reveal that the spatial profile and spectrum of these near-surface states strongly depend on both the sign and the strength of the SP. Using *ab-initio* band structure calculations to take the specificity of the materials into account, we investigate the NI/TI heterostructures formed by a single tetradymite-type quintuple or septuple layer block and the 3D TI substrate. The analytical continuum theory results relate the near-surface state evolution with the SP variation and are in good qualitative agreement with those obtained from density-functional theory (DFT) calculations. We also predict the appearance of the quasi-topological bound state on the 3D NI surface caused by a local band gap inversion induced by an overlayer.

Keywords: topological insulator, overlayer, interface states

(Some figures may appear in colour only in the online journal)

1. Introduction

It is generally recognized that there is one-to-one correspondence between the presence of nontrivial topological invariants characterizing the bulk electron states of a crystal and the appearance of specific electron modes localized at the crystal boundary [1–7]. This statement, which is known as the bulk-boundary correspondence theorem, reflects profound interrelation between the interior and exterior electron states of the truncated crystal. While the formulated assertion is based on general arguments of the topological concept for solids, in real materials and/or heterostructures, the effect of the pe-

culiar bulk properties on the surface may be quite intricate. This problem has been widely discussed in the context of the existence of topological bound states on the surface of a 3D topological insulator (3D TI) or at the interface between 3D TI and a topologically trivial material in various hybrid structures [4–7]. Angle-resolved photoemission spectra have provided evidence for the Dirac-like dispersion and the momentum-dependent spin texture of the 3D TI surface states in a family of Bi- and Sb-based narrow-gap semiconductors with strong spin-orbit coupling (SOC) [4–6, 8], which leads to the inverted band structure characterized by nontrivial topological invariants.

Although the topologically protected surface states are often considered as the most important and even decisive property of 3D TIs, in practice, the specific manifestations of the boundary-related electron properties in semiconductor materials with an inverted energy gap go far beyond the bulk-boundary correspondence paradigm. In other words, being formally correct, the topological arguments tell us too little about the characteristics of the topologically protected surface/interface states (e.g. the details of the dispersion, actual length scale, spin texture) or about any other in-gap electron states that might form at the real boundary of a 3D TI material. While rapid progress has been made in investigations of the vacuum-terminated TI surfaces [4–6, 8], it is still challenging to control the properties of the Dirac states under the surface modification in complex situations when a 3D TI is brought into contact with some substances. Recent experiments on 3D TIs have demonstrated that these states are notably sensitive to external perturbations, such as, chemical doping of the surface via deposition of both magnetic and nonmagnetic elements [9–11], oxidation of air-exposed samples [12], a change of the surface termination [8, 13], capping layers and interfaces with other materials [14, 15], applying an external gate voltage [16, 17], etc.

Perturbation of a bulk crystal potential exists naturally in truncated crystals and, in particular, 3D TIs and creates a surface potential (SP) affecting the electron properties of a crystal near/at the surface. In [18, 19] it was argued that a bulk-truncated surface of bismuth-chalcogenides can develop complex electronic structures in which the Dirac states coexist with the conventional states of the 2D electron gas (2DEG) in the quantum well, appearing near the 3D TI surface due to the band-bending effect. The *ab initio* calculations [20, 21] have shown that an expansion of the van-der-Waals (vdW) spacing in layered 3D TIs caused by intercalation of deposited atoms leads to a simultaneous emergence of 2DEG bands localized in the subsurface region. Moreover, the expansion of the vdW spacing also leads to a relocation of the Dirac topological states to the lower quintuple layers [21]. Wang *et al* have studied the effects of surface modification on the topological surface state in Bi_2Se_3 using first-principles calculations and have shown that Bi-capping and Se-removing can move the Dirac point upwards and slowly flatten the topological surface bands [22]. The short-range chemical forces related to dangling-bonds on the surface of thallium-based ternary chalcogenide TIs (TlBiTe_2 , TlBiSe_2 , TlSbTe_2 and TlSbSe_2) produce strong surface states [23] which can be removed by thallium adatoms [24]. In [25] it was suggested that the Dirac point of the helical surface states can be significantly shifted by applying uniaxial strain.

Many exciting physical properties of the Dirac helical quasiparticles (in particular, spin-dependent transport) are predicted to provide good opportunities for different spintronic applications [26–29]. To fully embody these promising ideas in devices, one requires multiple interfaces with topologically trivial materials rather than a single pristine surface of TI. Using density functional theory to design superlattice structures based on Bi_2Se_3 , it was shown that an interface state with an ideal Dirac cone is caused by alternating the layers of

3D TI and 3D normal insulator (NI) [30]. The authors of [31] have theoretically studied the $\text{Sb}_2\text{Se}_3/\text{Bi}_2\text{Se}_3$ heterostructures, the constituents of which possess the Bloch functions of the same symmetry. They found that the probability maximum of the Dirac state largely moves from the topologically nontrivial Bi_2Se_3 into the region of the topologically trivial Sb_2Se_3 . On the other hand, ARPES experiments [32] provide direct evidence that the surface state of the top surface of the heterostructure containing a single quintuple layer (QL) of Bi_2Se_3 on 19QLs of Bi_2Te_3 is similar to the surface state of Bi_2Se_3 . Moreover, the transport measurements [32] show that the studied heterostructure behaves more like Bi_2Se_3 even though there is only 1QL Bi_2Se_3 layer grown on 19QLs Bi_2Te_3 . In [33, 34] it was established that, as a result of depositing a 3D NI overlayer (conventional semiconductor ZnM , $\text{M} = \text{S}$, Se and Te) onto the 3D TI substrate (Bi_2Se_3 or Bi_2Te_3), the topological states can float to the top of the NI film, or stay put at the NI/TI interface, or are pushed down deeper into 3D TI. Recently, Berntsen and colleagues directly observed the Dirac states at the $\text{Bi}_2\text{Se}_3/\text{Si}(111)$ buried interface [15]. Another photoemission study in [35] revealed the existence of the interface topological states in the layered bulk crystal $(\text{PbSe})_5(\text{Bi}_2\text{Se}_3)_{3m}$, which forms a natural multilayer heterostructure composed of TI and NI. The evidence of a large shift of the Dirac point towards the conduction band edge relative to the case of the 3D TI/vacuum interface, due to the In_2Se_3 [14] or $\text{Sb}_2\text{Se}_2\text{Te}$ [36] capping layer on the epitaxial Bi_2Se_3 thin film, was reported, demonstrating the possibility of controlling the Dirac cone in 3D TI-based systems.

Thus, the experimental and theoretical data exhibit that the real 3D TI surface and 3D TI/NI interface possess very rich and diverse physics, in particular, they can hold both topological and non-topological (ordinary) in-gap states. It is well known that the non-topological bound states can be created or deleted or altered, depending on both the sign and strength of SP, when 3D NI is exposed to ambient conditions or when it comes into the contact with other materials. On the contrary, in 3D TI, the topological order itself is robust against such influences so that it can only be completely destroyed under a drastic perturbation [37, 38]. Nevertheless, the parameters of the topological states can undergo remarkable changes with even moderate external perturbations. The combination of the robustness of the topological states at the TI/NI interfaces with the tunability of their parameters to the external influence favours the design of the 3D TI/NI layered systems possessing suitable band structure, charge distribution and spin texture. The efficient design of the 3D TI/NI systems can be realized by combining analytic and numerical methods.

In the present work, we consider a special type of 3D TI/NI heterostructure of particular interest, which contain an ultrathin film of nonmagnetic NI (overlayer) artificially deposited on a relatively thick film of 3D TI (substrate). Due to a specific relation between electron affinities and band gap widths of the substrate and overlayer materials, significant modification of the spectrum and wave function of the Dirac states are expected as compared to the pristine 3D TI surface (i.e. the 3D TI/vacuum interface), thus making it possible to obtain the 3D TI-based heterostructure with tailor-made electron properties.

One assumes that the substrate film thickness is large enough to avoid sizeable hybridization of the bound states appearing at the opposite boundaries of the film. At the same time, the minimal thickness of the overlayer is formally limited by the condition of an electron motion quantization in the 3D NI material. Under these restrictions, to describe analytically the electron bound states near the surface of the truncated 3D TI covered by the 3D NI overlayer, a continual approach involving a method of effective surface potential (SP) was offered in [39]. In what follows, we will use the term ‘near-surface state’ (see also [39]) for identification of the in-gap electron state localized in the subsurface region and driven by the overlayer-induced SP. Just recently, in [36], several preliminary results concerning the near-surface states in the 3D TI/NI heterostructures with realistic material parameters were obtained within the numerical simulations based on density functional theory (DFT). Below, we employ the two complementary approaches—analytical and numerical—in order to thoroughly elucidate the important question of how a 3D NI overlayer affects the electron properties of 3D TI/NI heterostructures.

In the framework of the analytical approach, it is instructive to re-formulate this question in terms of the boundary conditions at the TI/NI interface for the wave function of the system. It is clear that, in systems composed of two topologically distinguishable materials, the characteristics of the near-surface state depend crucially on the choice of the boundary conditions, which still remains a highly disputable subject (for example, see [40–42]). The wave function at the ideal atomic interface between a pair of similar materials (e.g. 3D TI Bi_2Se_3 and 3D NI Sb_2Te_3 have the same crystal symmetry) satisfies the Ben-Daniel&Duke boundary conditions [43]. While matching the wave function at the contact of two dissimilar materials (e.g. such as Si and Bi_2Se_3) is complicated within the \mathbf{kp} formalism because the envelope functions (EF) on each side of the interface are defined using distinct orbital basis (see [44] and reference therein). However, such a complication was circumvented for the particular models describing different types of contacts within the effective interface potential concept [45–47].

Below, in the framework of the continual approach involving the SP scheme, we formulate general boundary conditions for the long-range envelope function of the truncated 3D TI and truncated 3D NI and find the solution for the bound near-surface states. We restrict ourselves to the situation when, solely, the orbital degree of freedom of electrons is manipulated by an external influence at the surface. We succeed in a general qualitative understanding of the dependence of the energy spectrum and spatial profile of the near-surface states on an effective SP. Furthermore, in order to elucidate the fine details of the near-surface state transformation induced by the overlayer, we employ the material-specific DFT calculations for the 3D TI/NI heterostructures. For conceptual reasons, within an effective SP scheme, we also discuss the near-surface states in other types of fictitious heterostructure, in which a substrate of a 3D NI close to the quantum transition into a 3D TI phase is covered with an overlayer of either NI or TI material.

The paper is organized as follows. In section 2, we discuss the effective SP concept, we propose the model for a truncated TI covered with an overlayer and we introduce the main ingredients and assumptions of the problem within the continual approach. In section 3, in the case of a spin-independent surface perturbation caused by an overlayer, we thoroughly investigate how the corresponding SP modifies the electron energy spectrum and the EF spatial profile of the Dirac-like near-surface state. In section 4, we analyze the main features of the near-surface state in a situation when a 3D NI substrate close to transition into a topological phase is covered with an ultrathin overlayer of either a NI material or a TI one. To corroborate the SP formalism results in section 5, we apply the DFT calculations and analyze the band structure and wave-function of the bound near-surface states for a set of heterostructures formed by a single tetradymite-type quintuple (QL) or septuple (SL) layer block and 3D TI substrates. Finally, the main conclusions are presented in section 6.

2. Surface potential concept and model Hamiltonian

Apart from the aforementioned simulations of the properties of 3D TIs based on the first principle calculations, various continual models have been discussed to describe relativistic fermions at the TI boundary [31, 40, 48]. There are theoretical studies of the TI properties, which are routinely based on the simple phenomenological 2D Hamiltonian for helical fermions with the linear Dirac-cone-like energy-momentum dispersion under an external influence [49]: $\mathbb{H}_s = -iv(\mathbf{e}_z[\boldsymbol{\sigma} \times \nabla]) + \mathbb{U}$, where \mathbf{e}_z is the unit vector normal to the surface, v is the Fermi velocity and $\boldsymbol{\sigma}$ is the vector composed of the Pauli matrices. It is generally thought that a controllable external field, \mathbb{U} , can be directly applied to the 3D TI surface to manage its electron states. For instance, a spin-independent term $\mathbb{U} = I_{2 \times 2} U$ could simulate the energy shift of the Dirac-cone point due to an electrostatic potential caused by a nonmagnetic overlayer. In turn, a spin-dependent term $\mathbb{U} \sim \sigma_z m$ could generate a gapped spin-polarized surface state through an exchange field proportional to the magnetization $\mathbf{m} = \mathbf{e}_z m$ applied along the normal to the surface of the 3D TI which is in contact with a ferromagnetic insulator [26, 50, 51]. In this manner, to take into account a perturbation arising from the external influence, the additional term \mathbb{U} is simply included in the 2D Hamiltonian \mathbb{H}_s , without a serious analysis of the microscopic origin of both \mathbb{H}_s and \mathbb{U} . The vast majority of theoretical works restrict to such the description and they predict many curious effects which can be realized in the 3D TI-based structures. However, the 2D Hamiltonian \mathbb{H}_s can be formally derived from a relevant 3D Hamiltonian only under free surface stipulation in the spirit of [31, 40]. In the framework of the consistent scheme, the bound near-surface states for the half-infinite 3D TI are composed of the eigenstates of the corresponding bulk 3D TI Hamiltonian. However, to date, nobody has written down the full set of the orthogonal wave-functions (including the bound and extended along \mathbf{e}_z -direction states) for the \mathbf{kp} Hamiltonian of 3D TI in half-infinite geometry, even under the free boundary conditions on the surface. Strictly speaking,

the bound states alone do not form the full basis set suited to the correct description of the \mathbb{U} field effect on 3D TI. The external surface perturbation excites electron density in the bulk 3D TI region of a nanoscopic scale adjacent to the surface. Both the bound and extended electron modes give rise to the response of 3D TI to this perturbation. Upon placing the 3D NI on 3D TI, besides the topological bound state, the so-called ordinary bound state [45] can arise near the interface due to the hybridization between the NI and TI atomic orbitals through the interface. Hence the near surface electron density perturbation of the 3D TI has very complicated spatial, orbital and spin configurations. The phenomenological 2D Hamiltonian \mathbb{H}_s could hardly serve as a starting point for the correct analysis of the configuration-dependent response of the 3D TI. So, to correctly take into account the effect of an external perturbation on the surface/interface electron states in the 3D TI, one must directly include the field \mathbb{U} into the ‘true’ 3D Hamiltonian of the system.

A basic idea to go beyond the scope of the 2D model is the use of the well-known $\mathbf{k}\mathbf{p}$ method [52]. To characterize the band electron states of a bulk semiconductor, $|n\mathbf{k}\rangle$ (\mathbf{k} is a wave vector, n is a band index) and in the region of the Brillouin zone around the point of band extrema \mathbf{k}_0 , the $\mathbf{k} \cdot \mathbf{p}$ method is reputed to be accurate enough. Under a smooth perturbation on the atomic scale, this method makes it possible to predict the evolution of the electron state wave function $\Psi_n(\mathbf{r})$ in terms of a product of a slowly varying envelope function (EF) $\theta_n(\mathbf{r})$ and the Bloch function of the unperturbed crystal $|n\mathbf{k}_0\rangle = \exp(i\mathbf{k}_0\mathbf{r})u_{n\mathbf{k}_0}(\mathbf{r})$ at the point \mathbf{k}_0 : $\Psi_n(\mathbf{r}) = \theta_n(\mathbf{r})|n\mathbf{k}_0\rangle$, $u_{n\mathbf{k}_0}(\mathbf{r})$ is the lattice periodic function. The EF concept may also be applied to the description of localized and resonant interface states in the semiconductor junctions of different types. However, a relevant choice of the boundary conditions for the function $\theta_n(\mathbf{r})$ remains an unsettled question in this concept, in particular for the TI based structures. The authors of [31, 40] impose the so-called ‘open’ boundary conditions fixing all EF components to zero at the crystal surface. This restriction formally simulates the effect of the vanishing quasiparticle wave function on the infinitely high SP barrier. Nevertheless, it should be pointed out that the zero constraint is not unique and other options for boundary conditions have been advocated in the literature. For instance, in [41] the problem is formulated in terms of an energy functional whose minimization yields the so-called ‘natural’ boundary conditions, intermixing the magnitudes and derivatives of different EF components. Both mentioned types of EF boundary conditions are extremely idealized and cannot adequately take into consideration the sensitivity of the 3D TI electron states to the surface modifications.

In this work we propose a formalism to directly incorporate the surface perturbation effect into the 3D TI Hamiltonian. We derive the appropriate EF boundary conditions through the construction of the effective semi-phenomenological SP localized at the 3D TI surface. The orbital and spin structure of the SP mimics induced fields resulting from a surface perturbation. As shown below, the structure and strength of the SP determine both the spatial and spectral features of the topological states.

The low energy and long wavelength bulk electron states of the prototypical TI, narrow-gap semiconductor of the Bi_2Se_3 -type, are described by the four band $\mathbf{k}\mathbf{p}$ Hamiltonian with strong SOC proposed in [31, 48]. Without a loss of generality, we make use of the simple version of this Hamiltonian in the form:

$$\mathbb{H}(\mathbf{k}) = \Xi(\mathbf{k})\tau_z \otimes \sigma_0 + A\tau_x \otimes (\boldsymbol{\sigma} \cdot \mathbf{k}), \quad (1)$$

where $\Xi(\mathbf{k}) = \Xi - Bk^2$, \mathbf{k} is the wave vector, $k = |\mathbf{k}|$, σ_α and τ_α ($\alpha = 0, x, y, z$) denote the Pauli matrices in the spin and orbital space, respectively. The Hamiltonian is written on the basis $u_{\mathbf{k}_0} = \{|+\uparrow\rangle, |-\uparrow\rangle, |+\downarrow\rangle, |-\downarrow\rangle\}$ of the four states at the Γ point of the Brillouin zone with $\mathbf{k}_0 = 0$. The superscripts \pm denote the even and odd parity states and the arrows $\uparrow\downarrow$ indicate the spin projection onto the z quantization axis. The Hamiltonian (1) captures the remarkable feature of the band structure: under the condition $\Xi B > 0$, the inverted order of the energy terms $|+\uparrow\rangle$ ($|\downarrow\rangle$) and $|-\uparrow\rangle$ ($|\downarrow\rangle$) around $\mathbf{k}_0 = 0$, which correctly characterizes the topological nature of the system due to a strong SOC. The Hamiltonian (1) is particle-hole symmetric and isotropic, which helps us to simplify the calculations.

We consider a semi-infinite 3D TI material, such as Bi_2Se_3 , occupying the region $z > 0$. The material boundary located at $z = 0$ is perfectly flat and displays translational symmetry in the (x, y) plane. The potential at the surface of a real 3D TI material is different from the bulk crystal potential, irrespective of whether the surface is kept in ultra-high vacuum or, for example, coated with an overlayer or interfaced with another material. To demonstrate the effect of the surface modification on the topological states within a conceptually simple scheme, we introduce the interaction of electrons with an external perturbation confined at the surface, implementing the effective SP $\mathbb{U}(\mathbf{r})$ into the EF calculation. Thus we write the full electron energy of the truncated 3D TI in the following form:

$$\Omega = \int_{z>0} d\mathbf{r} \Theta^\dagger(\mathbf{r}) [\mathbb{H}(-i\nabla) + \mathbb{U}(\mathbf{r})] \Theta(\mathbf{r}), \quad (2)$$

Here the operator $\mathbb{H}(-i\nabla)$ determined in equation (1) acts in the the spinor function space $\Theta(\mathbf{r}) = (\theta_1(\mathbf{r}), \theta_2(\mathbf{r}), \theta_3(\mathbf{r}), \theta_4(\mathbf{r}))^{\text{tr}}$, represented in the basis $u_{\mathbf{k}_0}$, the superscript tr denotes the transpose operation. The EF components $\theta_j(\mathbf{r})$ (the subscript j numbers the spinor components) are presumed to be smooth and continuous functions in the half-space $z > 0$, while the spatial symmetry and periodicity of the system are broken due to the existence of the TI surface. It is evident that the $\mathbf{k}\mathbf{p}$ approach cannot provide a correct description of the wavefunction behaviour near the surface, where large momenta are very important. To overcome this drawback we introduce the effective SP $\mathbb{U}(\mathbf{r})$, which affects the electron states of TI at the surface. The potential $\mathbb{U}(\mathbf{r})$ is nonzero in a small region d (of the order of a lattice parameter) around the geometrical boundary $z = 0$, where the validity of the $\mathbf{k}\mathbf{p}$ scheme is questionable. The introduction of the phenomenological SP in equation (2) enables us to correctly match the low-energy and long-range electronic states inside the truncated TI with

evanescent vacuum states through the boundary conditions for EF $\Theta(\mathbf{r})$. As long as the EF spatial variation of the sought state, $\Theta(\mathbf{r}) = \sum_{\kappa} \Theta(\kappa, z) \exp(i\kappa\rho)$ [$\rho = (x, y)$, $\kappa = (k_x, k_y)$], is sufficiently slow in the direction normal to the surface, one can adopt a local approximation for the SP. Namely one writes $\mathbb{U}(\mathbf{r}) = d\mathbb{U}(\rho)\delta(z+0)$, where the symbol $+0$ at the argument of the delta-function signifies that the sheet-like SP is placed inside the TI half-space but at an infinitesimally small distance from the boundary $z = 0$.

As a matter of course, an electron wave function has to be continuous at a crystal boundary. Nevertheless, in the system under consideration, since the Bloch factors of the wave function inside and outside TI do not coincide (in particular, they have distinct space symmetries), the long-range EF $\Theta(\kappa, z)$ can formally undergo a finite break (jump) across the boundary from $z = 0-$ to $z = 0+$ within the utilized **kp** method (we refer the reader to the detailed discussion in [41]). In the current work, we do not care how the wave-function behaves in the half-space $z < 0$ but next we make use of a functional

$$F\{\Theta^\dagger, \Theta\} = \int_0^\infty dz \Theta^\dagger(\kappa, z) [\mathbb{H}(\kappa, -i\partial_z) + d\mathbb{U}(\kappa)\delta(z+0) - \mathbb{I}E]\Theta(\kappa, z), \quad (3)$$

where the energy E plays the role of the Lagrange multiplier, \mathbb{I} is an unit 4×4 matrix, $\partial_z = \partial/\partial z$. The functional (3) is determined in the class of the smooth and continuous EFs in the TI half-space $z > 0$ and includes the effective surface potential $d\mathbb{U}(\kappa)\delta(z+0)$. Since, in a plane geometry, the wave-vector κ is a good quantum number, we determine the functional for each EF κ -mode, $\Theta(\kappa, z)$. Varying functional $F\{\Theta^\dagger, \Theta\}$ with respect to Θ^\dagger yields the Euler equations for the half-space $z > 0$ and the boundary conditions at the surface at $z = 0+$. The corresponding equations in the compact form are:

$$[\mathbb{H}(\kappa, -i\partial_z) - \mathbb{I}E]\Theta(\kappa, z) = 0, \quad (4)$$

$$i \frac{\delta \mathbb{H}(\kappa, -i\partial_z)}{\delta(-i\partial_z)} \Theta(\kappa, z)|_{z=0+} = 2d\mathbb{U}(\kappa)\Theta(\kappa, z)|_{z=0+}. \quad (5)$$

On the left side of equation (5) the current density operator acts on the EF spinor. Thus the right side associated with the surface perturbation plays the role of the external (with regard to the TI bulk) current source (sink). The equation (5) involves the surface potential parameters, in this sense it has something in common with the equation which was used to calculate the surface states of a crystal with a relativistic band structure in [53]. The solution of the boundary task, equations (4) and (5), answers the principal physical question; how the perturbation located just at the TI boundary affects the near-surface topological states.

In the half-space $z > 0$, the general solution of equation (4) for each EF spinor component obeying the condition $\theta_j(\kappa, z \rightarrow \infty) = 0$ can be represented as

$$\theta_j(\kappa, z) = \theta_j^0(\phi) \{ \alpha_j(\kappa, E) \exp[-q_1(\kappa, E)z] + \beta_j(\kappa, E) \times \exp[-q_2(\kappa, E)z] \}, \quad (6)$$

where

$$q_{1,2}(\kappa, E) = \sqrt{q_{1,2}^2(E) + \kappa^2}, \quad (7)$$

$$q_{1,2}^2(E) = \frac{A^2 - 2B\Xi \pm \sqrt{A^4 - 4B\Xi A^2 + 4B^2\Xi^2}}{2B^2}. \quad (8)$$

Here the phase factors $\theta_j^0(\phi)$ forming the spinor $\Theta^0(\phi) = (i, -\text{sgn}(A), \mp e^{i\phi}, \pm \text{sgn}(A)ie^{i\phi})^t$ depend only on the momentum polar angle, ϕ , $k_x \pm ik_y = \kappa \exp(\pm i\phi)$; the signs \pm relate to lower and upper spectral branches, respectively. The characteristic momenta $q_{1,2}(\kappa, E)$ are the solutions of the corresponding secular equation; $\kappa = |\kappa|$. The boundary conditions, equation (5), determine the coefficients $\alpha_j(\kappa, E)$ and $\beta_j(\kappa, E)$ as well as the dispersion relation for the near-surface states inside the bulk band gap, $|E(\kappa)| < \Xi$. The parameter $\lambda = A^2/4B\Xi$ is implied to be $\lambda \geq 1$.

3. Near-surface topological bound states

The potential \mathbb{U} in equation (5) is a 4×4 matrix specifying the internal properties of the TI surface and the matrix elements include different components of the scattering of the TI states on SP. In principle, choosing the structure of the matrix and the strength of its components allows us to tune the spatial and energy characteristics of the topological states. For example, as for the SP diagonal matrix elements, U_{jj} , the values $U_1 = (U_{11} + U_{33})/2$ and $U_2 = (U_{22} + U_{44})/2$ are proportional to the scattering intensity of particle and hole, respectively, on the spin-independent part of SP, while the quantities $Q_1 = (U_{11} - U_{33})/2$ and $Q_2 = (U_{22} - U_{44})/2$ are proportional to the scattering intensity of particle and hole, respectively, on the z -component of the exchange part of SP. The off-diagonal matrix elements $U_{jj'}$ with $j \neq j'$ result from the spin-orbit interaction at the surface, which, in general is different to the SOC in the TI bulk.

In this work we focus on the SP that preserves time reversal symmetry, i.e. $\mathbb{U} = \text{diag}\{U_1, U_2, U_1, U_2\}$, where $U_{11} = U_{33} = U_1, U_{22} = U_{44} = U_2$. Such that the SP structure in the basis $u_{\mathbf{k}_0}$ results in the following relations between the EF coefficients in equation (6): $\alpha_3 = \alpha_1, \beta_3 = \beta_1, \alpha_4 = \alpha_2, \beta_4 = \beta_2$. Moreover, we neglect the dependence of $U_{1,2}(\kappa)$ on κ in equation (5).

One can interpret the spin-independent scattering on the surface within the framework of a 'local band bending' scheme (which is quite reasonable for the contact of two insulators/semiconductors), where the band edge corresponding to the j -th spinor component is affected by the external perturbation confined at the surface: $\Xi \rightarrow \Xi \pm dU_j\delta(z)$. In the situation of TI covered with an overlayer, the intuitive idea is that the diagonal components of SP could be heuristically adjusted to the relative offsets between the corresponding energy levels (bands) of the TI substrate and the overlayer. In other words, the energy $dU_j\delta(z)$ mimics the local bending of the respective bands.

After some algebra the corresponding secular equation results in the implicit relation between the energy E and the

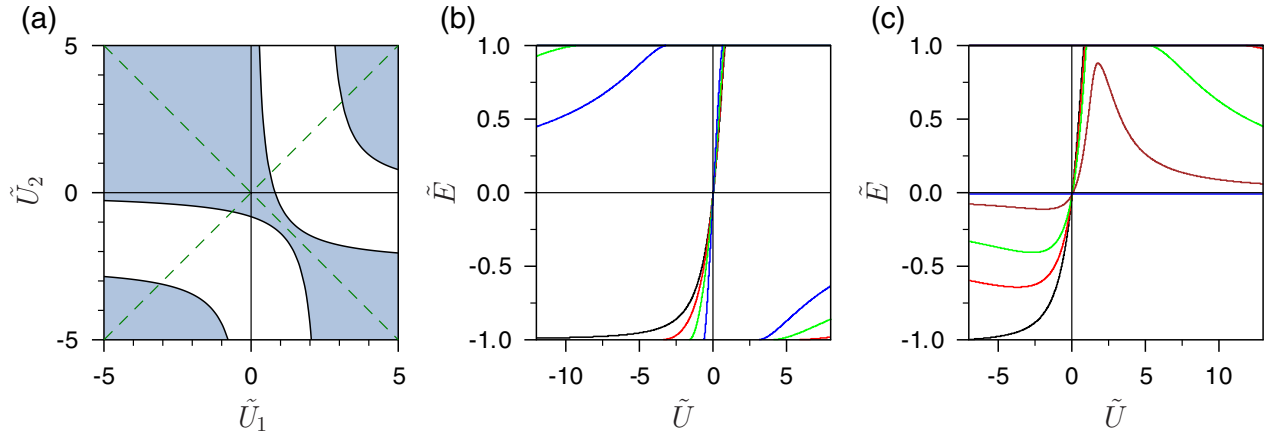


Figure 1. (a) Realm of the near-surface state Dirac point position in a 3D TI in terms of the matrix elements U_1 and U_2 . In the dashed areas the Dirac point is inside the bulk band gap. (b) Position of the Dirac point versus the SP strength when $U = U_1 = mU_2$, where $m = 0$ (black line), $m = 0.1$ (red line), $m = 0.3$ (green line), $m = 1.0$ (blue line). (c) Position of the Dirac point versus the SP strength when $U = U_1 = mU_2$, where $m = 0$ (black line), $m = -0.2$ (red line), $m = -0.4$ (green line), $m = -0.8$ (brown line), $m = -1.0$ (blue line coinciding with an abscissa axis). The units used are, $\tilde{U} = \frac{dU}{\sqrt{B\Xi}}$, $\tilde{E} = \frac{E}{\Xi}$ and $\lambda = \frac{A^2}{4B\Xi} = 2$.

in-plane momentum κ for the bound state at the TI surface:

$$\begin{aligned}
 & A^2[q_1 \pm \kappa][q_2 \pm \kappa] - [Bq_1^2 + \Xi(\kappa) - E][Bq_2^2 + \Xi(\kappa) - E] \\
 & + 2\left\{q_1q_2 - \frac{\Xi(\kappa) - E}{B} \pm \kappa[q_1 + q_2]\right\} \\
 & \times \left\{B^2q_1q_2 - dU_1dU_2 - \frac{A^2}{4}\right\} \\
 & - 2[\Xi(\kappa) - E][dU_2(q_1 + q_2) \pm \kappa(dU_1 + dU_2)] \\
 & - 2B\{dU_1q_1q_2(q_1 + q_2) \pm \kappa dU_1(q_1 + q_2)^2 \\
 & \mp \kappa(dU_1 + dU_2)q_1q_2\} = 0, \quad (9)
 \end{aligned}$$

where $q_{1,2} = q_{1,2}(\kappa, E)$ in accordance with equations (7) and (8), $\Xi(\kappa) = \Xi - B\kappa^2$. Note that equation (9) is invariant under the simultaneous permutations: $E \leftrightarrow -E$, $\kappa \leftrightarrow -\kappa$ and $U_1 \leftrightarrow -U_2$.

At the Γ point, equation (9) is reduced to the equation that determines the Dirac (node) point position, $E(\kappa = 0) = E_0(U_{1,2})$, as a function of the SP strength $U_{1,2}$:

$$\begin{aligned}
 & \left[\sqrt{1 + \frac{E}{\Xi}} - \sqrt{1 - \frac{E}{\Xi}}\right] \left[\lambda + \sqrt{1 - \frac{E^2}{\Xi^2}} - \frac{dU_1dU_2}{B\Xi}\right] \\
 & - \sqrt{\frac{2}{B\Xi}} \left[dU_1\sqrt{1 + \frac{E}{\Xi}} + dU_2\sqrt{1 - \frac{E}{\Xi}}\right] \\
 & \times \sqrt{2\lambda - 1 + \sqrt{1 - \frac{E^2}{\Xi^2}}} = 0. \quad (10)
 \end{aligned}$$

The shaded areas in figure 1(a) denote the realm of the near-surface bound state with $|E_0(U_{1,2})| < \Xi$ on the (U_1, U_2) -plane. The dependence of the node point position on the SP strength, obtained from equation (10) for several ratio values U_1/U_2 , is plotted in figure 1(b) (for $(U_1U_2) > 0$) and figure 1(c) (for $(U_1U_2) < 0$). One can see three different regions in these plots. At weak potential $d|U_{1,2}| \ll \sqrt{B\Xi}$, the Dirac point linearly shifts with respect to the TI bulk bands to either higher or lower binding energies depending on the SP strength sum,

$$E_0(U_1, U_2) = \frac{2\sqrt{\lambda}}{1 + \lambda} \sqrt{\frac{\Xi}{B}} (dU_1 + dU_2). \quad (11)$$

In case the SP strength is large, $d|U_1| + d|U_2| \gg \sqrt{B\Xi}$, the node point energy approaches zero as

$$E_0(U_1, U_2) = -2\sqrt{\lambda B\Xi^3} \left(\frac{1}{dU_1} + \frac{1}{dU_2} \right). \quad (12)$$

On the (U_1, U_2) -plane, there are regions (the unshaded areas in figure 1(a)) where the bound state is absent since the node point merges into the conduction or valence bulk band. For example, if $U_1 = U_2 = U$, the threshold values of the potential, at which the node point splits off the bulk band continuum, are $2dU_{\pm} = \sqrt{B\Xi}[\sqrt{2(4\lambda - 1)} \pm \sqrt{2(2\lambda - 1)}]$, so that $E_0(\pm U_{-}) = \pm\Xi$ and $E_0(\pm U_{+}) = \mp\Xi$.

If the energy $\varepsilon(\kappa)$ is a small deviation from the Dirac linear spectrum, $E^{(\pm)}(\kappa) = \pm A\kappa + \varepsilon(\kappa)$, $|\varepsilon(\kappa)| \ll \Xi$, the characteristic momenta, equation (8), are found as

$$q_{1,2}^{(\pm)}(\kappa, E) = q_{1,2}^0(\kappa) + \frac{(\pm\kappa)\varepsilon(\kappa)}{Bq_{1,2}^0(\kappa)[q_{1,2}^0(\kappa) - q_{2,1}^0(\kappa)]}, \quad (13)$$

where $q_{1,2}^0(\kappa)$ is given by equation (18). The deviation $\varepsilon(\kappa)$ appears to be small not only when $U_1 \simeq -U_2$ but also when SP is either weak or strong. Using the expression (13) one can obtain the spectrum and estimate the spatial distribution of the near-surface state in these limit situations.

So, for the extremely large potential, $|U_{1,2}| \rightarrow \infty$, the correction approaches zero, $\varepsilon(\kappa) \rightarrow 0$, in turn, the EF coordinate dependence is described by a difference of the exponents, $\Theta(\kappa, z) \sim \exp[-q_1^0(\kappa)z] - \exp[-q_2^0(\kappa)z]$, so that the maximum of the electron density, $|\Theta(z)|^2$, does not occur on the surface, where $\Theta(z = 0) = 0$, but rather near the point $z_0 = \ln(q_1^0/q_2^0)/(q_1^0 - q_2^0)$ (where $z_0 \lesssim \sqrt{\frac{B}{\Xi}} < (q_2^0)^{-1}$) that is distant from the surface. Such an EF distribution, together with the linear spectrum, was found under the free boundary conditions [40]. Our approach allows us to capture peculiarities of the surface state in 3D TI induced by the SP. If the SP strength is much greater than the characteristic energy, $d|U_{1,2}| \gg Bq_1^0$, within the perturbation theory, one obtains the

amendment to the linear dispersion law as

$$\varepsilon(\kappa) = -\frac{|A|\Xi^2}{\Xi(\kappa)} \left(\frac{1}{dU_1} + \frac{1}{dU_2} \right). \quad (14)$$

The surface state spectrum acquires a curvature and a shift of the node point, $E_0 = \varepsilon(0)$, which are inversely proportional to the potential, however the fermion group velocity $|A|$ near the node point does not change since the amendment $\varepsilon(\kappa)$ (14) does not contain a contribution linear in κ . In the lowest order in $(U_{1,2})^{-1}$, the relations between the coefficients in equation (6) are given by

$$\begin{aligned} \frac{\beta_1}{\alpha_1} &= -1 + \frac{\sqrt{A^2 - 4B\Xi(\kappa)}}{dU_1}, \\ \frac{\beta_2}{\alpha_2} &= -1 - \frac{\sqrt{A^2 - 4B\Xi(\kappa)}}{dU_2}, \end{aligned} \quad (15)$$

Thus, the electron density does not vanish on the TI surface, $|\Theta(z=0)|^2 \sim (U_{1,2})^{-2}$. However, under the SP influence, the EF components can vanish near the surface at $z = z_{1,2} < z_0$, namely, $\theta_{1,3}(0, z) = 0$ at $z = z_1 = B/dU_1$ when $U_1 > 0$ and $\theta_{2,4}(0, z) = 0$ at $z = z_2 = -B/dU_2$ when $U_2 < 0$. Besides, as seen from equation (13), the SP affects the decay length of the EF nonzero harmonics.

If the SP is formally absent, $U_{1,2} = 0$, one arrives at the solution obtained in [41] from using the natural boundary conditions: the surface state shows the linear spectrum $E^{(\pm)}(\kappa) = \pm A\kappa$ and the EF spatial profile in z -direction is merely a sum of the two exponents, $\Theta(\kappa, z) \sim \exp[-q_1^0(\kappa)z] + \exp[-q_2^0(\kappa)z]$, i.e. the probability density of the near-surface state is peaked on the boundary $z = 0$ and its tail penetrates into the TI bulk with the decay length $(q_2^0)^{-1}$. In the case of weak SP, $d|U_{1,2}| \ll \sqrt{B\Xi}$, the correction to the dispersion law is given by

$$\varepsilon(\kappa) = \frac{4|A|\Xi(\kappa)}{A^2 + 4B\Xi(\kappa)} (dU_1 + dU_2). \quad (16)$$

The spin-independent SP is seen to shift entirely and warp the energy-momentum dependence. Note, that the corrections (14) and (16) are the opposite in sign. Turning on the SP leads to the different contributions of the quick and slow exponents into EF: $\beta_j/\alpha_j = 1 + o(U_{1,2})$.

Let us consider thoroughly the specific case of the staggered alignment of the matrix elements, $U_1 = -U_2 = U$, when SP does not break the particle-hole symmetry. One can verify in equation (9) that in such a case the near-surface state maintains the ideal Dirac spectrum $E^{(\pm)}(\kappa) = \pm A\kappa$ regardless of the size and sign of U . While the spectrum is independent of the SP, the envelope function is strongly affected by it. The coordinate dependence of each component of the EF spinor is given by

$$\theta_j(\kappa, z) = \theta_j^0 \sqrt{\frac{\Xi}{B}} \frac{(1+v)\exp(-q_1^0 z) + (1-v)\exp(-q_2^0 z)}{\sqrt{\frac{(1+v)^2}{2}q_2^0 + \frac{(1-v)^2}{2}q_1^0 + (1-v^2)\frac{2\Xi(\kappa)}{|A|}}}, \quad (17)$$

where

$$\begin{aligned} q_{1,2}^0 &= q_{1,2}^0(\kappa) = q_{1,2}(\kappa, |E| = |A|\kappa) \\ &= \frac{|A| \pm \sqrt{A^2 - 4B\Xi(\kappa)}}{2B}. \end{aligned} \quad (18)$$

$$v = v(\kappa) = \frac{2dU}{\sqrt{A^2 - 4B\Xi(\kappa)}}. \quad (19)$$

The EF of equation (17) is normalized as $\int_0^\infty |\theta_j(\kappa, z)|^2 = 1$. The spatial behaviour of the EF zeroth harmonic $\theta_j(0, z)$ is illustrated in figures 2(a) and (b) for positive and negative U , respectively. With increasing SP strength the EF structure evolves from the sum of the exponents at $U = 0$ (black lines) to the difference at $|U| \rightarrow \infty$ (yellow lines). So, one sees a gradual change in the profile of the near-surface state such that its gravity centre moves from the surface to the TI interior. It is obvious the behaviour of the EF nonzero harmonics equation (17) is insignificantly different from what is plotted in figure 2; unless the tail at $\kappa \neq 0$ is slightly longer than that at $\kappa = 0$.

To study the modification of the near-surface states under the finite strength of SP we dwell at length on the situation $U_1 = U_2 = U$. As seen in figure 3, at a finite strength of U , $d|U| \simeq \sqrt{B\Xi}$, apart from the aforementioned shift of the node point, the form of the spectral dependence, $E(\kappa)$, alters (in comparison with the limiting cases $U = 0$ or $U \rightarrow \pm\infty$) under the SP influence. The group velocity of the surface topological excitations decreases from the quantity $|A|$ to zero when the SP strength $|U|$ either increases from zero to the threshold value U_- or decreases from infinity to the threshold value U_+ . A noticeable deviation from linearity can be seen for the strength $d|U| \simeq \sqrt{B\Xi}$. In the limit $|U| \rightarrow U_-$ the dispersion becomes parabolic at small κ and in the limit $|U| \rightarrow U_+$ the curve $E(\kappa)$ smoothly merges into $E = -\Xi$. The dependence $E(\kappa)$ acquires a curvature so that the relatively strong ($|U| > U_+$) and relatively weak ($|U| < U_-$) potentials provide the curvature with the opposite sign.

The spatial behaviour of the near-surface states for $U_1 = U_2 = U$ is shown in figure 4. When the SP is weak, $0 < |U| < U_-$, the probability density $\sim |\theta_j(z)|^2$ largely peaks near the surface. The strong SP, $|U| > U_+$, pushes the probability density towards the TI bulk. The EF is exponentially decaying away from the surface. We would like to emphasize that the EF decay lengths, $(q_{1,2})^{-1}$ (see equation (8)), are strongly influenced by the SP strength. For example, when the strength U varies either from 0 to $\pm U_-$ or from $\pm\infty$ to $\pm U_+$, the momentum $q_1(E)$ increases from $\sqrt{\frac{\Xi}{B}}(\sqrt{\lambda} + \sqrt{\lambda - 1})$ to $\sqrt{\frac{\Xi}{B}}\sqrt{2(2\lambda - 1)}$ (i.e. the 'long' exponent of EF (6) becomes longer) and the momentum $q_2(E)$ decreases from $\sqrt{\frac{\Xi}{B}}(\sqrt{\lambda} - \sqrt{\lambda - 1})$ to 0 (i.e. the 'short' exponent becomes shorter).

4. Quasi-topological bound states near the surface of a normal insulator

Next, we investigate the effect of the surface modification on the near-surface bound states when the bulk is a 3D normal (topologically trivial) insulator. Here we address the fundamental question of whether 3D NI responds to a localized surface perturbation in a different way than 3D TI. In order to describe 3D NI, one uses the same relativistic Hamiltonian (1) in which, however, we now have a normal (non-inverted)

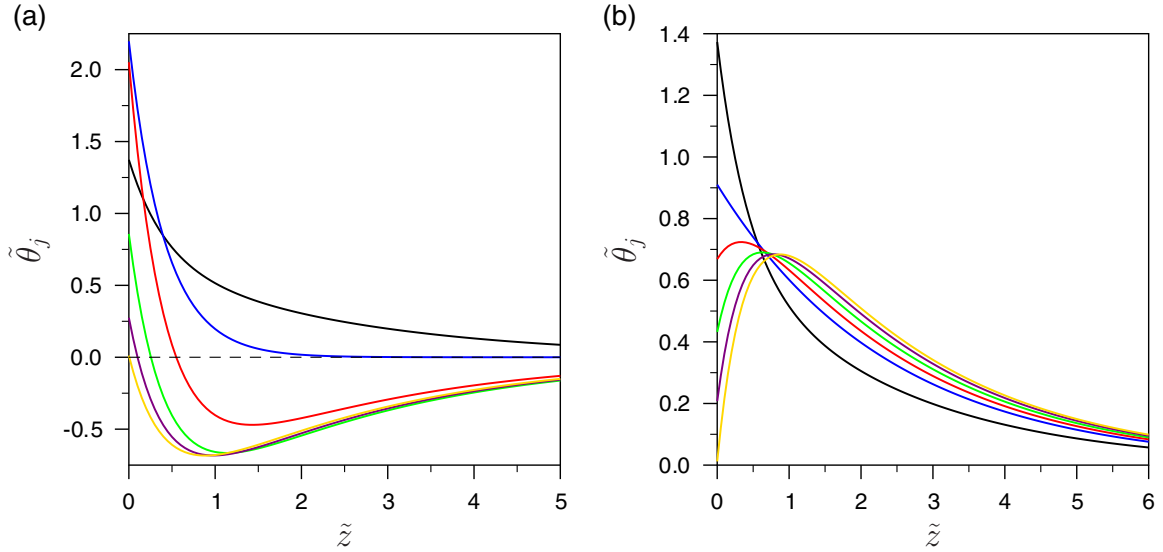


Figure 2. Space dependence of the zeroth harmonic ($\kappa = 0$) of the envelope function $\theta_j(\kappa, z)$ equation (17) in the case of $U_1 = -U_2 = U$ for several values of the potential: (a) $\tilde{U} = 0$ (black line), $\tilde{U} = 1$ (blue line), $\tilde{U} = 2$ (red line), $\tilde{U} = 4$ (green line), $\tilde{U} = 10$ (brown line), $\tilde{U} = \infty$ (yellow line); (b) $\tilde{U} = 0$ (black line), $\tilde{U} = -1$ (blue line), $\tilde{U} = -2$ (red line), $\tilde{U} = -4$ (green line), $\tilde{U} = -10$ (brown line), $\tilde{U} = -\infty$ (yellow line); where $\tilde{U} = \frac{dU}{\sqrt{B\xi}}$, $\tilde{z} = \sqrt{\frac{\xi}{B}}z$, $\tilde{\theta}_j = \sqrt{\frac{B}{\xi}}\theta_j$, $\lambda = \frac{A^2}{4B\xi} = 2$. The envelope function is normalized as $\int_0^\infty dz |\theta_j(\kappa, z)|^2 = 1$.

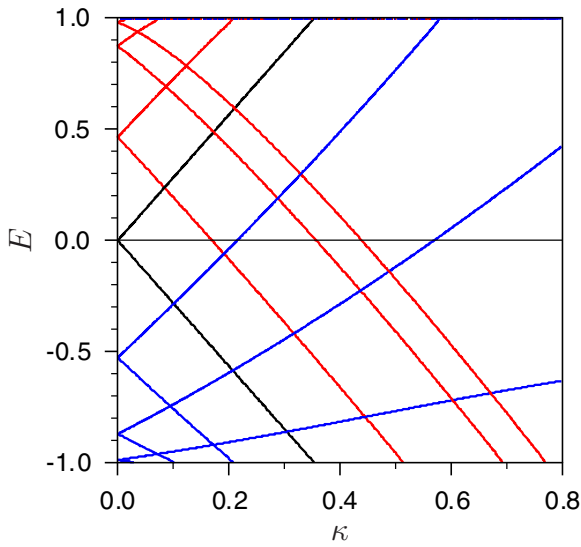


Figure 3. Spectrum of surface states $E(\kappa)$ at $U_1 = U_2 = U$ for several values of SP: $\tilde{U} = 0$ (black line), $\tilde{U} = 0.25, 0.5, 0.6$ (red lines), $\tilde{U} = 3.5, 5.0, 10.0$ (blue lines); $\tilde{U} = \frac{dU}{\sqrt{B\xi}}$, $\tilde{E} = \frac{E}{\xi}$, $\tilde{\kappa} = \sqrt{\frac{B}{\xi}}\kappa$, $\lambda = \frac{A^2}{4B\xi} = 2$.

alignment of the energy terms of different parity $|+\uparrow(\downarrow)\rangle$ and $|-\uparrow(\downarrow)\rangle$ around $\mathbf{k}_0 = 0$ is implied, i.e. $\Xi < 0$ and $\Xi B < 0$. For instance, in the case of the In_2Se_3 crystal, which shares the same crystal structure with Bi_2Se_3 , the SOC is not strong enough to provide the inversion between two p_z orbitals with opposite parity at the Γ point [31]. In the limit $A \rightarrow 0$, when SOC is negligibly small, equation (1) defines merely the semiconductor with a simple (nonrelativistic) two-band spectrum.

It is evident that equations (2)–(5) and the relevant sentences are valid regardless of the Ξ sign. Therefore, analogous to what was done in the previous sections (the details are omitted), one can obtain the characteristics of the bound electron states on the NI surface subjected to the external spin-independent influence. The existence of these states is determined by real solutions of the corresponding secular equation within the bulk gap, $E(0)$, which is given by equation (9) at $\kappa = 0$ and $\Xi = -|\Xi|$. Figure 5(a) shows the existence realm of the near-surface bound state, i.e. the area on the (U_1, U_2) -plane where $|E_0| < |\Xi|$. The energy $E(0)$ as a function of the SP strength for several ratio values U_1/U_2 is represented in figure 5(b) (for $U_1 U_2 > 0$) and figure 5(c) (for $U_1 U_2 < 0$). As one can see, except for the quadrant ($U_1 < 0, U_2 > 0$), one may choose the ratio U_1/U_2 to match the SP which induces the bound electron state on the NI surface.

In what follows, we consider thoroughly only two particular cases: $U_1 = -U_2$ and $U_1 = U_2$. When the SP matrix elements are in staggered rows, $U_1 = -U_2 = U$, the bound state exists at $U > 0$, which virtually diminishes the 3D NI bulk gap on the surface. Because of the presence of such a SP, the particle-hole symmetry of the system is preserved. The relations between the energy and momentum is given by

$$E(\kappa) = \pm \left\{ \Omega^2(\kappa) + A^2 \kappa^2 - \left[dU \sqrt{\frac{A^2}{2B^2} + \frac{2\Omega(\kappa)}{B} - \frac{d^2 U^2}{B^2} - \frac{A^2}{4B}} \right]^2 \right\}^{1/2}, \quad (20)$$

where $\Omega(\kappa) = |\Xi| + B\kappa^2$, $\pm\Omega(\kappa)$ is the projection of the bulk spectrum onto the surface. In figure 6, the spectral dependence $E(\kappa)$ is illustrated for several choices of the SP strength U . So, the system exhibits a non-linearly dispersing surface state which is specified by the energy gap $2E(0)$. The crossing

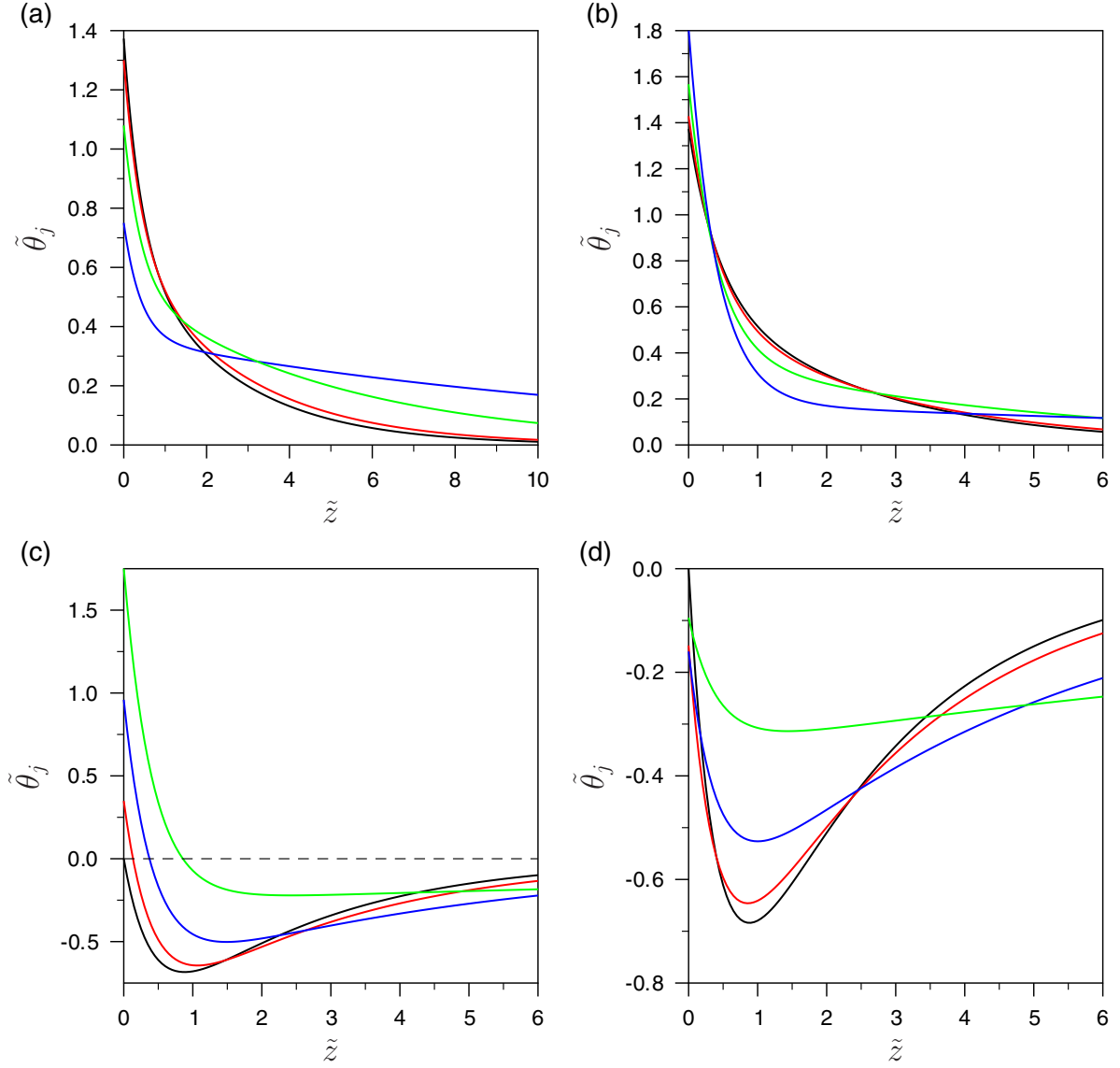


Figure 4. Space dependence of the zeroth harmonic ($\kappa = 0$) of the envelope function $\theta_j(\kappa, z)$ in the case of $U_1 = U_2 = U$ for several values of (a) the relative weak positive potential with $0 < U < U_-$ [$\tilde{U} = 0$ (black line), $\tilde{U} = 0.25$ (red line), $\tilde{U} = 0.5$ (green line), $\tilde{U} = 0.6$ (blue line)], (b) the relative weak negative potential with $-U_- < U < 0$ [$\tilde{U} = 0$ (black line), $\tilde{U} = -0.25$ (red line), $\tilde{U} = -0.5$ (green line), $\tilde{U} = -0.6$ (blue line)], (c) the relative large positive potential with $U_+ < U < \infty$ [$\tilde{U} = \infty$ (black line), $\tilde{U} = 10$ (red line), $\tilde{U} = 5$ (blue line), $\tilde{U} = 3.5$ (green line)], (d) the relative large negative potential with $-\infty < U < -U_+$ [$\tilde{U} = -\infty$ (black line), $\tilde{U} = -10$ (red line), $\tilde{U} = -5$ (blue line), $\tilde{U} = -3.5$ (green line)], where $\tilde{U} = \frac{dU}{\sqrt{B|\Xi|}}$, $\tilde{z} = \sqrt{\frac{\Xi}{B}}z$, $\tilde{\theta}_j = \sqrt{\frac{4}{B|\Xi|}}\theta_j$, $\lambda = \frac{A^2}{4B|\Xi|} = 2$. The envelope function is normalized as $\int_0^\infty dz |\theta_j(\kappa, z)|^2 = 1$.

black lines in figure 5(c) show the half gap $E(0)$ as a function of U . One can see in figures 5(a) and (c) that the near-surface state stays in the bulk band gap, $|E(\kappa)| < |\Xi|$, when the SP strength is restricted by the interval $U_+ > U > U_-$, where $\tilde{U}_\pm = (\sqrt{1 + 2|\lambda|} \pm 1)/\sqrt{2}$. While, outside this interval, it is buried in the bulk band continuum. In the close vicinity of a band-crossing point \tilde{U}_0 , where $|E(\kappa)| \ll |\Xi|$, the dispersion relation is given by $E(\kappa) = \pm \sqrt{4\Xi^2(\tilde{U} - \tilde{U}_0)^2 + A^2\kappa^2}$, so that a degeneracy at the band-crossing point U_0 is lifted due to gapping $\sim |U - U_0|$. It is convenient to measure the SP strength in the dimensionless units $\tilde{U} = \frac{dU}{\sqrt{B|\Xi|}}$, then $\tilde{U}_0 = \sqrt{1 + |\lambda|}$, where $|\lambda| = \frac{A^2}{4B|\Xi|}$. In turn, in a small energy window near the bulk band edges, where $|\Xi| - |E(\kappa)| \ll |\Xi|$, the dependence

(20) becomes

$$E(\kappa) = \pm \left\{ \Omega^2(\kappa) + A^2\kappa^2 - \frac{1}{4\lambda^2} \left[|\Xi|(\tilde{U}_+^2 - \tilde{U}^2)(\tilde{U}^2 - \tilde{U}_-^2) + 2\tilde{U}^2 B\kappa^2 \right]^2 \right\}^{1/2}. \quad (21)$$

Note, in the case $A = 0$, equation (20) reduces to $E(\kappa) = \pm[\Omega(\kappa) - d^2U^2/B]$; in other words, in the 3D NI under finite value of SOC, a finite strength of SP, $U_- \sim A^2$, is required to split off the in-gap near-surface state from the 3D bulk continuum.

Such a behaviour of the bound state location in the energy axis with an increasing SP strength can intuitively be explained

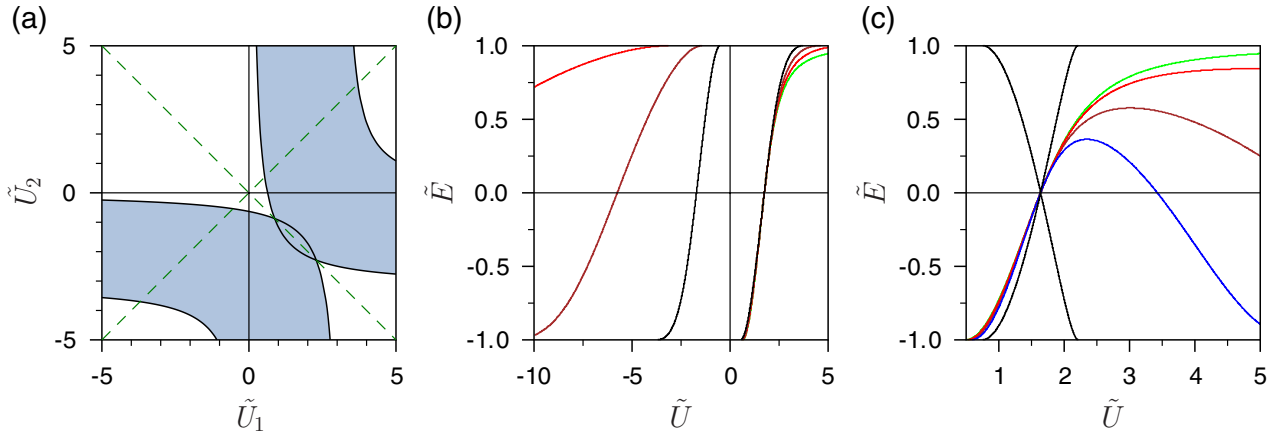


Figure 5. (a) Realm of the near-surface state existence in 3D NI under SP with the matrix elements U_1 and U_2 . In the dashed (fill, painted) areas (domains) on the (U_1, U_2) -plane, the point $E(0)$ lives inside the bulk band gap. (b) Position of the band edge energy of the near-surface state versus the SP strengths of the same sign under the stipulation $U_2 = nU_1$ and $U_1 = U$, where $n = 0$ (green line), $n = 0.1$ (red line), $n = 0.3$ (brown line), $n = 1.0$ (black line). (c) Position of the band edge energy of the near-surface state versus the SP strengths of opposite signs under the stipulation $U_2 = nU_1$ and $U_1 = U$, where $n = 0$ (green line), $n = -0.1$ (red line), $n = -0.3$ (brown line), $n = -0.5$ (blue line), $n = -1.0$ (black line). Dimensionless units are used, $\tilde{U} = \frac{dU}{\sqrt{|B||\Xi|}}$, $\tilde{E} = \frac{E}{|\Xi|}$ and $|\lambda| = \frac{A^2}{4B|\Xi|} = 2$.

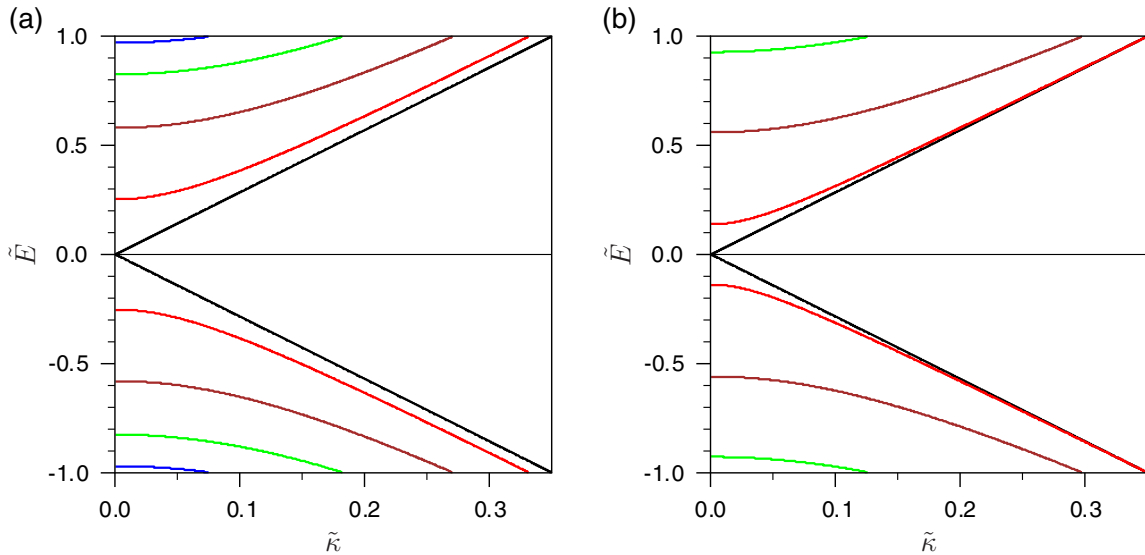


Figure 6. Spectrum of the near-surface state $E(\kappa)$ at $U_1 = -U_2 = U$ for several values of the SP strength: (a) $U \leq U_0$ [$\tilde{U} = \sqrt{3}$ (black line), $\tilde{U} = 1.6$ (red line), $\tilde{U} = 1.4$ (brown line), $\tilde{U} = 1.2$ (green line), $\tilde{U} = 1.0$ (blue line)]; (b) $U \geq U_0$ [$\tilde{U} = 1.8$ (red line), $\tilde{U} = 2.0$ (brown line), $\tilde{U} = 2.2$ (green line), $\tilde{U} = \sqrt{3}$ (black line)], where $\tilde{U} = \frac{dU}{\sqrt{|B||\Xi|}}$, $\tilde{E} = \frac{E}{|\Xi|}$, $\tilde{\kappa} = \sqrt{\frac{B}{|\Xi|}}\kappa$, $|\lambda| = \frac{A^2}{4B|\Xi|} = 2$.

in the language of the ‘local band bending’ scheme proposed in the previous section. When 3D NI is brought into contact with a thin dielectric overlayer, a relative weak positive SP, $U \gtrsim U_-$, splits off states from both the conduction bulk band and the valence one due to a local narrowing of the gap, $|\Xi| - dU\delta(z)$. At $U \simeq U_0$, the local band bending is so steep that the bulk band edges of 3D NI cross over the energy levels of an adjusted overlayer with opposite parity, as would be the case if the 3D NI surface was in contact with a TI overlayer. A further increase of the strength U above the critical value U_+ pushes the near-surface state into the bulk continuum.

Figure 7 visualizes the effect of the external potential with $U_1 = -U_2 = U$, on the space profile of the in-gap state of the truncated 3D NI. It is of interest to note that, while crossing the value U_0 , the form of the space

dependence of the EF zeroth harmonic, $\theta_j(\kappa = 0, z)$, switches over from monotonically decreasing, figure 7(b), (in the situation of $U > U_0$, which mimics the 3D NI/TI-overlayer heterostructure) to nonmonotonically decreasing with a minimum at $z \approx \sqrt{B/|\Xi|}$, figure 7(a), (in the situation of $U < U_0$, which mimics the 3D NI/NI-overlayer heterostructure). If $U \simeq U_0$, the EF can be approximated as:

$$\theta_j(0, z) \sim \exp[-q_1(0, 0)z] + \text{sgn}(U_0 - U)(\sqrt{1 + |\lambda|} - \sqrt{|\lambda|}) \times \exp[-q_2(0, 0)z], \quad (22)$$

$$q_{1,2}(0, 0) = \sqrt{\frac{B}{|\Xi|}}(\sqrt{1 + |\lambda|} \pm \sqrt{|\lambda|}). \quad (23)$$

When the strength $U \rightarrow U_{\pm}$, i.e. $|E| \rightarrow |\Xi|$, the EF (6) is the superposition of the slow exponent with relatively

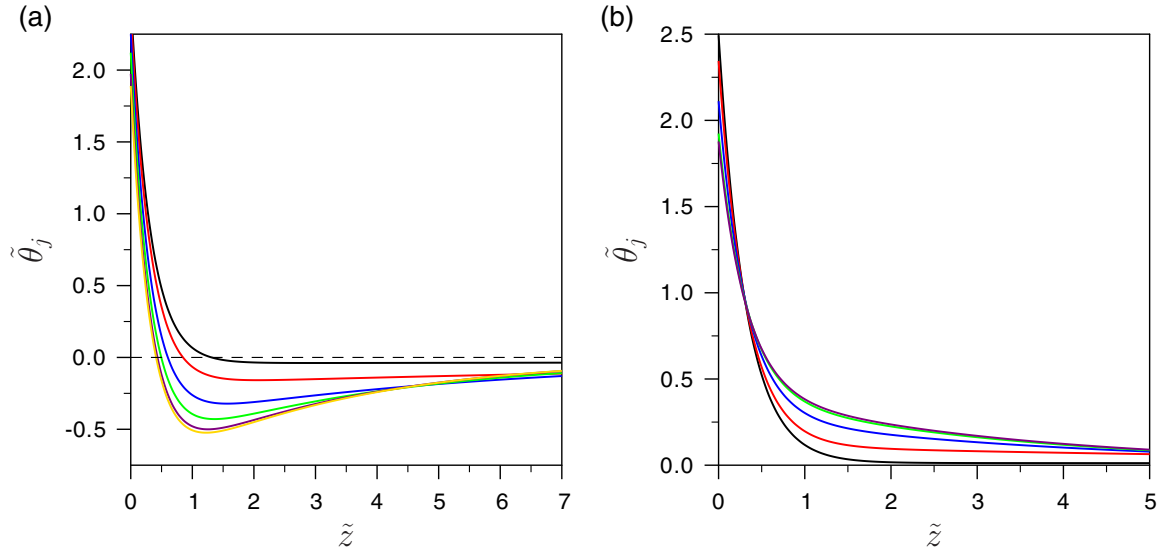


Figure 7. Space dependence of the zeroth harmonic ($\kappa = 0$) of the NI envelope function $\theta_j(\kappa, z)$ in the case of $U_1 = -U_2 = U$ for several values of the SP strength with (a) $U_- < U < U_0$ [$\tilde{U} = 0.9$ (black line), $\tilde{U} = 1.0$ (red line), $\tilde{U} = 1.2$ (blue line), $\tilde{U} = 1.4$ (green line), $\tilde{U} = 1.6$ (brown line), $\tilde{U} = 1.7$ (yellow line)], (b) $U_+ > U > U_0$ [$\tilde{U} = 2.28$ (black line), $\tilde{U} = 2.2$ (red line), $\tilde{U} = 2.0$ (blue line), $\tilde{U} = 1.8$ (green line), $\tilde{U} = 1.75$ (brown line)], where $\tilde{U} = \frac{dU}{\sqrt{B|\Xi|}}$, $\tilde{z} = \sqrt{\frac{|\Xi|}{B}}z$, $\tilde{\theta}_j = \sqrt{\frac{B}{|\Xi|}}\frac{\theta_j}{\sigma_0^0}$, $|\lambda| = \frac{A^2}{4B|\Xi|} = 2$. The envelope function is normalized as $\int_0^\infty dz |\theta_j(\kappa, z)|^2 = 1$.

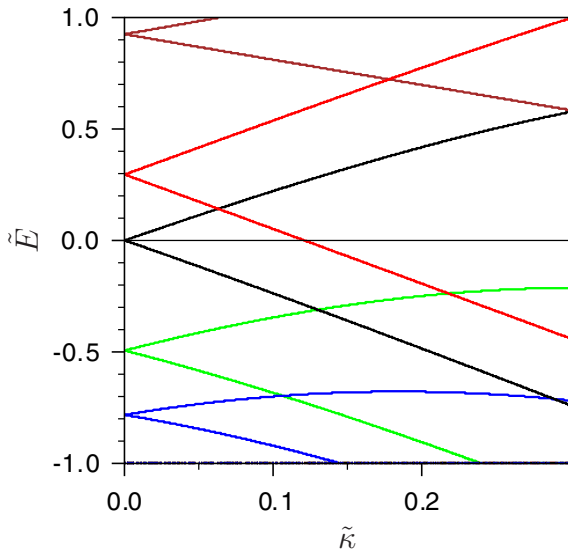


Figure 8. Spectrum of the near-surface state $E(\kappa)$ at $U_1 = U_2 = U$ for several values of the SP strength: $\tilde{U} = 2.0$ (red line), $\tilde{U} = 3.0$ (brown line), $\tilde{U} = 1.3$ (green line), $\tilde{U} = 1.0$ (blue line), $\tilde{U} = \sqrt{3}$ (black line), where $\tilde{U} = \frac{dU}{\sqrt{B|\Xi|}}$, $\tilde{E} = \frac{E}{|\Xi|}$, $\tilde{\kappa} = \sqrt{\frac{B}{|\Xi|}}\kappa$, $|\lambda| = \frac{A^2}{4B|\Xi|} = 2$.

small weight and the quick one with relatively large weight: $q_2/q_1 \simeq \sqrt{1 - E^2/\Xi^2}/(2 + 4|\lambda|)$. This situation is depicted with the black and red curves in figure 7.

Let us now draw attention to the case when the surface perturbation has the spinor structure $U_1 = U_2 = U$ answering to the surface electrostatic potential. Figure 8 shows the dispersion law for several values of the SP strength U . In the vicinity of the points $U = \pm U_0 = \pm\sqrt{1 + |\lambda|}$, where

$|E| \ll |\Xi|$, the relations between the energy and the in-plane momentum (in the leading order in κ) is given by

$$\frac{E(\kappa)}{|\Xi|} = \text{sgn}(U) \left(\frac{\tilde{U}^2}{\tilde{U}_0^2} - 1 \right) \pm \frac{2|\lambda|\tilde{\kappa}}{\tilde{U}_0}, \quad (24)$$

where $\tilde{\kappa} = \kappa\sqrt{B/|\Xi|}$. Thus, it is clear, given U belonging to the interval(s) $W_- < |U| < W_+$ (where $W_\pm = (\sqrt{1 + 4|\lambda|} \pm \sqrt{1 + 2|\lambda|})/\sqrt{2}$), the surface state consists of a single Dirac cone. Within the framework of a heuristical ‘local band bending’ scheme, the variation of the strength U is linked with the relative movement of the energy levels of the 3D NI substrate and the NI overlayer. When the SP strength value exceeds the threshold quantity, $|U| > W_-$, the band structure of this system (which consists of the two materials with a normal gap band alignment) is inverted, i.e. either the substrate conduction band is lower than the overlayer valence band or the substrate valence band is higher than the overlayer conduction band. As a result, if the strength is in the interval $W_- < |U| < W_+$, the near-surface state of the NI covered by the normal overlayer can display a linear dispersion dependence of the Dirac-cone form, $E(\kappa) = E(0) \pm v\kappa$, where the node point location $E(0)$ and propagation velocity v are the functions of the strength U and band structure parameter $|\lambda|$. The energy $E(0)$ is inside the NI bulk gap, $|E(0)| < |\Xi|$. The quasi-topological bound state is also specified by the space distribution, which is shown in figure 9. The corresponding EF decays exponentially away from the surface. When the strength $|U|$ attains the quantity W_\pm , the near-surface state merges into the bulk continuum states, in turn the decay length $\sim q_2^{-1}$ becomes large (the black and green curves). At $|U| \simeq W_+$, the probability density is concentrated close to the surface, in the other case, it is rather smeared.

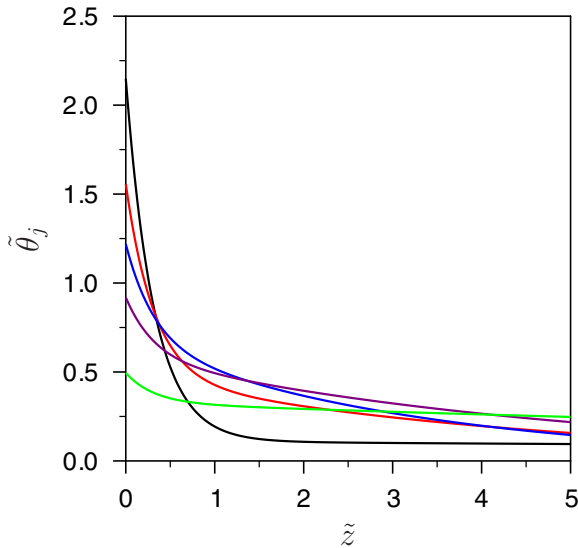


Figure 9. Space dependence of the zeroth harmonic ($\kappa = 0$) of the NI envelope function $\theta_j(\kappa, z)$ in the case of $U_1 = U_2 = U$ for several values of the SP strength: $\tilde{U} = 3.5$ (black line), $\tilde{U} = 2.5$ (red line), $\tilde{U} = 1.5$ (blue line), $\tilde{U} = 1.0$ (brown line), $\tilde{U} = 0.65$ (green line), where $\tilde{U} = \frac{dU}{\sqrt{B|\Theta|}}$, $\tilde{z} = \sqrt{\frac{|\Theta|}{B}}z$, $\tilde{\theta}_j = \sqrt{\frac{4|\Theta|}{B}}\frac{\theta_j}{\theta_j^0}$, $|\lambda| = \frac{A^2}{4B|\Theta|} = 2$. The envelope function is normalized as $\int_0^\infty dz |\theta_j(\kappa, z)|^2 = 1$.

5. *Ab initio* calculations

The proposed continual approach gives a transparent physical explanation for the evolution of the near-surface state in both momentum and real spaces with the SP superimposed on the TI or NI boundary. This approach describes fairly well the electron density distribution of the corresponding Dirac-cone-like states on a scale exceeding the lattice spacing through EF(s) as the superposition $\theta_j(\kappa, z)$ (see equation (6), where the exponents $q_{1,2}(\kappa, E)$ and the coefficients $\alpha_j(\kappa, E)$, $\beta_j(\kappa, E)$, are functions of the SP components and the bulk band structure parameters. However, within the SP scheme, we are unable to elucidate the electron density features on a scale of the order of the SP spacing $d \ll q_{1,2}^{-1}$, in particular, to capture the fine effect of the the topological state relocation within near-surface layers [33, 54]. Below, in order to provide a closer look at the wave function of the bound near-surface state and to accurately reproduce its band structure over the whole Brillouin zone, we present *ab initio* density functional theory calculation results for some systems representing the topological insulator substrate covered by the insulator ultrathin film.

Electronic structure calculations were carried out within the density functional theory using the projector augmented-wave method [55] as implemented in the VASP code [56, 57]. The exchange-correlation energy was treated using the generalized gradient approximation [58]. The Hamiltonian contained the scalar relativistic corrections and the spin-orbit coupling was taken into account by the second variation method [59]. In order to take into account the effect of dispersion interactions, we use the van der Waals nonlocal correlation functional within the DFT-D2 approach [60].

The thin film NI/TI heterostructures were simulated within a model of repeating slabs separated by a vacuum spacing of

10 Å. The overlayers were symmetrically attached to both sides of the substrate slab to preserve the inversion symmetry.

The following substrates were chosen; 3D TIs with tetradymite-like layered structures $\text{Bi}_2\text{Te}_2\text{S}$, $\text{Sb}_2\text{Te}_2\text{Se}$ [20, 21] composed of quintuple layer (QL) blocks and GeBi_2Te_4 [8, 61] composed of septuple layer (SL) blocks. The substrates were simulated by 6 QL (5 SL) slabs. We used single QL(SL) films of $\text{Bi}_2\text{Te}_2\text{S}$, $\text{Sb}_2\text{Te}_2\text{S}$ and GeBi_2Te_4 , which have gapped noninverted spectra, as the thin insulating overlayers. The interface systems under investigation are provided in table 1. As can be seen in table 1, the lattice mismatch between the overlayer and the TI substrate in the considered heterostructures does not exceed 0.5% providing very good epitaxial compatibility. The in-plane lattice parameters of the heterostructures were fixed to the experimental ones of the substrate slab [62–64]. The interlayer distances within the overlayer and the TI block, closest to the interface, were optimized.

Figure 10(a) shows spectra of the free-standing GeBi_2Te_4 overlayer and the $\text{Bi}_2\text{Te}_2\text{S}$ substrate. In the spectrum of $\text{Bi}_2\text{Te}_2\text{S}$ the topological surface state (TSS) with the Dirac point lying in the $\bar{\Gamma}$ valley of the valence bulk states propagates across the bulk energy gap. The spectrum of the free-standing overlayer has a 440 meV gap. The energies of the overlayer states are matched to the substrate spectrum in accordance with work functions Φ_1 and Φ_2 given in table 1. Thus the highest occupied state of the overlayer lies ~ 30 meV above the top of the valence band of the substrate. The attaching of the topologically trivial insulating overlayer to the TI substrate keeps the gapless topological surface state, however, it leads to a strong modification of the spectrum (figure 10(b)): the DP shifts towards the bottom of the conduction band of the substrate so that above the DP it propagates as a resonant state, mixed with bulk-like states of the $\text{Bi}_2\text{Te}_2\text{S}$ slab. In the real space, the TSS in the heterostructure is almost completely relocated into the overlayer and its probability maximum lies near the $\text{GeBi}_2\text{Te}_4/\text{Bi}_2\text{Te}_2\text{S}$ interface plane (figures 10(c) and (d)). Such behaviour of the topological state can be explained by the potential change upon the interface formation. In figure 10(e) the change in the electrostatic potential of the heterostructure with respect to potentials in the free-standing substrate and overlayer is shown. One can see that due to hybridization between orbitals of the overlayer and the substrate the potential within the TI substrate is smoothly bent towards the interface plane while the potential within the overlayer undergoes more noticeable changes and as a whole it shifts down by ~ 120 meV with respect to its position in the free-standing overlayer. In spite of the downward shift of the potential the resulting topological state has a Dirac point position higher than in the TSS of the pristine TI surface. The change in the DP position is related to the fact that in the heterostructure the orbitals of GeBi_2Te_4 contribute more to the TSS than to the orbitals of $\text{Bi}_2\text{Te}_2\text{S}$. Thus the modification of the topological state in the considered system is qualitatively similar to the behaviour found in the continual model with $U_1 \approx U_2$ where for negative U_1 beneath a critical value solutions arise with a positive shift of the Dirac point (see figure 1(b)).

Table 1. The structural and energetic characteristics of the NI/TI Overlayer/Substrate pairs given in the X_1/X_2 format (except for the lattice mismatch, Δ), with X_i being the experimental lattice parameter, a_i , or the calculated work function, Φ_i , with an indication of the thickness of the films, or the calculated band gap, E_i ($i = 1, 2$). Note that the band gap values, E_i , are given for slab and bulk in the NI overlayer ($i = 1$) and TI substrate ($i = 2$) cases, respectively.

NI/TI	a_1 (Å)/ a_2 (Å)	Δ (%)	Φ_1 (eV)/ Φ_2 (eV)	E_1 (eV)/ E_2 (eV)
[GeBi ₂ Te ₄] _{1SL} /Bi ₂ Te ₂ S	4.3225/4.316	+0.15	4.99/5.04	0.44/0.27
[Sb ₂ Te ₂ S] _{1QL} /Sb ₂ Te ₂ Se	4.17/4.188	-0.43	4.99/4.62	0.47/0.30
[Bi ₂ Te ₂ S] _{1QL} /GeBi ₂ Te ₄	4.316/4.3225	-0.15	5.38/4.76	0.33/0.08

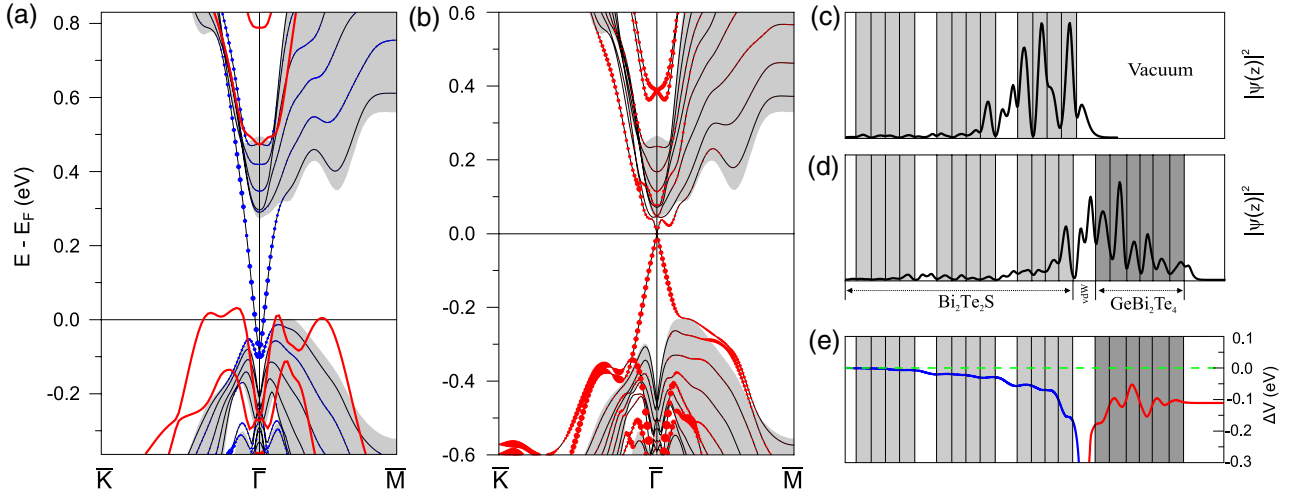


Figure 10. (a) Band spectra of the Bi₂Te₂S TI substrate and the free-standing septuple layer block of GeBi₂Te₄ (red lines). Blue markers show the weight of the states in near-surface QL of the substrate. The shaded areas indicate projection of the bulk states of Bi₂Te₂S TI onto the surface Brillouin zone; (b) Electronic spectrum of GeBi₂Te₄/Bi₂Te₂S heterostructure. The size of the red circles reflects the weight of the states in the GeBi₂Te₄ overlayer; (c) Spatial distribution of the TSS charge density integrated over the (x, y) plane for Bi₂Te₂S (c) and GeBi₂Te₄/Bi₂Te₂S (d); (e) the change in electrostatic potential ΔV of GeBi₂Te₄/Bi₂Te₂S with respect to potentials in the free-standing Bi₂Te₂S TI substrate and GeBi₂Te₄ overlayer.

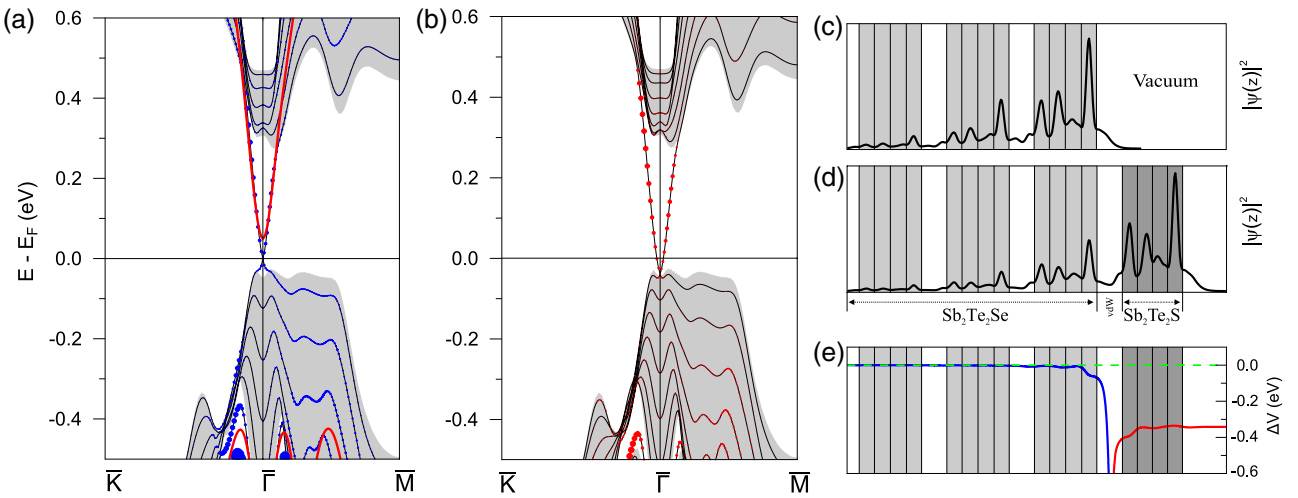


Figure 11. (a) Band spectra of the Sb₂Te₂Se substrate and free-standing Sb₂Te₂S QL (red lines). Blue markers show the weight of the states in the near-surface QL of the TI substrate; (b) electronic spectrum of Sb₂Te₂S/Sb₂Te₂Se heterostructure; (c) Spatial distribution of the TSS charge density for pristine Sb₂Te₂Se surface (c) and Sb₂Te₂S/Sb₂Te₂Se heterostructure (d); (e) the change in electrostatic potential ΔV of Sb₂Te₂S/Sb₂Te₂Se with respect to potentials in the free-standing Sb₂Te₂Se substrate and Sb₂Te₂S QL.

In the Sb₂Te₂S/Sb₂Te₂Se heterostructure (figure 11) the TSS dispersion remains almost unchanged with respect to that on the pristine Sb₂Te₂Se substrate being shifted towards the bulk valence band by ~ 40 meV. Along with this, the maximum of the charge density of the topological state

relocates into the Sb₂Te₂S QL (figures 11(c) and (d)) owing to a downward shift of the overlayer potential (figure 11(e)) and demonstrates a resonance like behaviour due to the DP proximity to the bulk continuum. In the model this scenario is realized when moderate negative potential at the

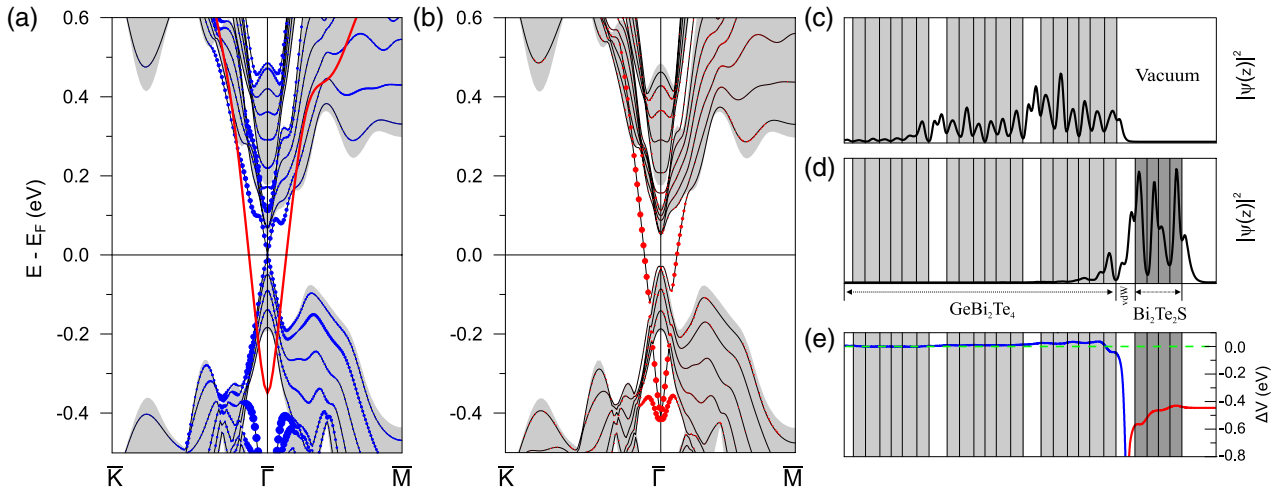


Figure 12. Electronic structure of GeBi_2Te_4 slab and free-standing $\text{Bi}_2\text{Te}_2\text{S}$ QL (red lines). Blue markers show the weight of the states in near-surface SL of GeBi_2Te_4 ; (b) electronic spectrum of the $\text{Bi}_2\text{Te}_2\text{S}/\text{GeBi}_2\text{Te}_4$ heterostructure; spatial distribution of the TSS charge density for pristine substrate surface (c) and the heterostructure (d); (e) the change in electrostatic potential ΔV of the heterostructure with respect to potentials in free-standing GeBi_2Te_4 slab and $\text{Bi}_2\text{Te}_2\text{S}$ QL.

interface does not lead to substantial change in the energy of the TSS.

A rather distinct type of the TSS modification demonstrates the $\text{Bi}_2\text{Te}_2\text{S}/\text{GeBi}_2\text{Te}_4$ system (figure 12). The work function of the free standing overlayer is 0.7 eV larger than that of the GeBi_2Te_4 substrate slab (see table 1). This means that the gap-edge states of the overlayer are far below the bulk gap of the TI substrate (figure 12(a)). In contrast to the previous cases, where when attaching the overlayer, the topological state remains within the bulk gap of the substrate, the TSS in the $\text{Bi}_2\text{Te}_2\text{S}/\text{GeBi}_2\text{Te}_4$ heterostructure occurs at -0.4 eV, in the local $\bar{\Gamma}$ gap of the bulk valence band of GeBi_2Te_4 . This state lies in a deep and abrupt potential well of ~ 500 meV depth that causes its strong localization within $\text{Bi}_2\text{Te}_2\text{S}$ QL. Thus the modification of the TSS in the $\text{Bi}_2\text{Te}_2\text{S}/\text{GeBi}_2\text{Te}_4$ heterostructure is qualitatively similar to the case in the continual model where a large negative value of the potential results in a huge downward shift of the Dirac point from its position in the initial system.

6. Summary and concluding remarks

In this work, we have shown that the energy spectrum and spatial profile of the near-surface electron states in the 3D TI substrate/NI overlayer type of heterostructure can be controlled by the overlayer induced SP. By choosing an appropriate overlayer it is possible to adjust the Dirac point to the required position in the band gap. The DFT calculations clearly demonstrate how the overlayer-induced shift of the Dirac point from its energy position on the pristine surface is associated with the parameters of the band structure of the 3D TI substrate and the quintuple/septuple overlayer. These results are in good qualitative agreement with the tendency predicted from the analytic continual scheme, in which the overlayer-induced change in the energy spectrum is determined by the SP. Indeed the SP matrix elements can be intuitively associated with the relative energy offsets between the relevant band edges of the

substrate and the overlayer as $U_1 \sim \Phi_2 - \Phi_1 + E_1 - E_2$ and $U_2 \sim \Phi_2 - \Phi_1$, where $E_{1,2}$ and $\Phi_{1,2}$ are bandgaps and work functions of an overlayer (subscribe 1) and a TI substrate (subscribe 2). Then taking into account the data in table 1 one can roughly estimate these values: $U_1 \sim 0.2$ eV and $U_2 \sim 0$ eV for $[\text{GeBi}_2\text{Te}_4]_{\text{ISL}}/\text{Bi}_2\text{Te}_2\text{S}$; $U_1 \sim -0.2$ eV and $U_2 \sim -0.4$ eV for $[\text{Sb}_2\text{Te}_2\text{S}]_{\text{IQL}}/\text{Sb}_2\text{Te}_2\text{Se}$; $U_1 \sim -0.4$ eV and $U_2 \sim -0.6$ eV for $[\text{Bi}_2\text{Te}_2\text{S}]_{\text{IQL}}/\text{GeBi}_2\text{Te}_4$. When the SP is relatively weak, one can see that the Dirac point shifts slightly towards the upper (figure 10) or lower (figure 11) binding energies depending on the sign of $U_{1,2}$. When the SP is relatively large, the Dirac point leaves the gap merging deeply into the bulk valence band continuum (figure 12). Such a modification of the spectrum is in qualitative agreement with the behaviour of $E_0(U_{1,2})$ represented in figure 1(b).

In the framework of the continual approach, we have succeeded in formulating general boundary conditions for the long-range EF and in finding the solution for electron bound states at the TI/NI interface. We have obtained analytical expressions for the energy spectrum and the EF for different types and values of SP. Our results are strictly consistent with the limiting cases of the zero and infinite surface potentials, which were previously studied in [40, 41]. The boundary conditions of equation (5) involves the SP parameters, in this sense the represented approach has something in common with the one used to calculate the surface states of a crystal with the relativistic band structure [53]. Note that in [45], the bound states at the interface between 3D TI and NI have been explored on the basis of the functional defined in the entire space. In that work, the explicit expressions for the matrix elements of the interface pseudo-potential, which affect electrons on the TI side of the interface, were analytically derived.

From a theoretical point of view, one cannot suggest a universal recipe to impose the restrictions on the EF behaviour at the 3D TI boundary for all types of heterostructures containing 3D TIs. Note, in particular, that the boundary conditions of equation (5) imposed upon the near-surface states are different from those obtained in [45], where, for an

interface between 3D TI and a topologically trivial insulator, the authors succeeded in a formulation of the EF boundary task, the solution of which provides an insight into the electron states at the interface. Although the EF approach may be adapted for the description of the interface states in many semiconductor junctions, by taking into account the general Hermiticity and symmetry requirements [44], this traditional description is not always able to capture the principal features of the topological states in systems containing narrow-gap semiconductors with inverted band structure and sometimes yields rather dubious results [65]. As for the above-stated conception, the appropriate EF boundary conditions are derived within the framework of the formalism of the sheet-like SP. In such an approach, some information (for example, about the effects of electron–electron interactions) which is evidently lost is compensated by a relative calculation simplicity and a transparent physical interpretation. The method, conceptually presented in our work for the study of the truncated 3D TI covered with an atomically thin non-magnetic insulating overlayer, can be straightforwardly extended to solve a wide range of problems related to the behaviour of the near-surface topological states under the surface perturbations listed in the Introduction.

Indeed, as was shown in section 4, the area of applicability for the SP method significantly oversteps the formal limits of the 3D TI substrate/NI overlayer systems. We have determined the existence or absence of the Dirac-like near-surface modes in hypothetical situations when the truncated 3D NI (close to transition into a topological phase) is brought into contact with the ultrathin overlayer of the 3D NI or TI material. We predict that the near-surface ‘topological-like’ mode can appear in the 3D NI substrate/NI overlayer, i.e. in the system composed only of two topologically trivial materials. This unusual item has something in common with the appearance of a 2D TI state in an InAs/GaSb Type-II semiconductor quantum well (QW) [66, 67]. The unique feature of InAs/GaSb QW is that the conduction band minimum of InAs has lower energy than the valence band maximum of GaSb (due to the large band-offset). Consequently, when the QW thickness is large enough, the first electron sub-band of the InAs layer lies below the first hole sub-band of the GaSb layer, i.e. an inverted band alignment, similar to that in HgTe QWs [4, 5], occurs. The experimental study of low temperature electronic transport has shown strong evidence for the existence of helical edge modes in the hybridization gap of inverted InAs/GaSb QWs [67].

In view of the aforementioned, one raises the following questions concerning the different strategies to create the topological near-surface/interface states and manage their electronic properties: (i) How does a 3D NI overlayer influence a 3D TI? (ii) When one puts an overlayer of one 3D TI on a substrate of another 3D TI, what new property comes into being? (iii) How about a hypothetical 3D TI thin film on the surface states of 3D NI? (iv) Finally, one could ask how to construct a quasi-topological bound state at the boundary between two topologically trivial insulators? Our work partly answers these questions. The approach proposed above unveils the physics of the near-surface states in the semiconductor heterostructures containing 3D TIs. At the same time, it

might be considered as a good guidebook for the qualitative interpretation and forecast of the topological phase behaviour tendencies in 3D TIs under surface perturbations. The obtained results should open new opportunities to design various combinations of topological and conventional materials for electronic/spintronic applications.

Acknowledgments

We acknowledge the partial support from the Basque Country Government, Departamento de Educación, Universidades e Investigación (Grant No. IT-756-13), the Spanish Ministerio de Ciencia e Innovación (Grant No. FIS2010-19609-C02-01), the Ministry of Education and Science of the Russian Federation (No. 2.8575.2013) and the Russian Foundation for Basic Researches (grant 13-02-00016). Calculations were performed on the SKIF-Cyberia supercomputer of Tomsk State University.

References

- [1] Fu L and Kane C L 2007 *Phys. Rev. B* **76** 045302
- [2] Moore J E and Balents L 2007 *Phys. Rev. B* **75** 121306
- [3] Essin A M and Gurarie V 2011 *Phys. Rev. B* **84** 125132
- [4] Hasan M Z and Kane C L 2010 *Rev. Mod. Phys.* **82** 3045
- [5] Qi X L and Zhang S C 2011 *Rev. Mod. Phys.* **83** 1057
- [6] Okuda T and Kimura A 2013 *J. Phys. Soc. Japan* **82** 021002
- [7] Ando Y 2013 *J. Phys. Soc. Japan* **82** 102001
- [8] Ereemeev S V *et al* 2012 *Nat. Commun.* **3** 635
- [9] Wray L A *et al* 2011 *Nat. Phys.* **7** 32
- [10] Scholz M R *et al* 2012 *Phys. Rev. Lett.* **108** 256810
- [11] Valla T *et al* 2012 *Phys. Rev. Lett.* **108** 117601
- [12] Kong D *et al* 2011 *ACS Nano* **5** 4698
- [13] Miao L *et al* 2013 *Proc. Natl Acad. Sci.* **110** 2758
- [14] Jenkins G S *et al* 2013 *Phys. Rev. B* **87** 155126
- [15] Berntsen M H, Götberg O and Tjernberg O 2013 *Phys. Rev. B* **88** 195132
- [16] Chen J *et al* 2010 *Phys. Rev. Lett.* **105** 176602
- [17] Checkelsky J G *et al* 2011 *Phys. Rev. Lett.* **106** 196801
- [18] Bianchi M *et al* 2011 *Phys. Rev. Lett.* **107** 086802
- [19] Bahrany M S *et al* 2012 *Nat. Commun.* **3** 1159
- [20] Menshchikova T V *et al* 2011 *JETP Lett.* **94** 106
- [21] Ereemeev S V *et al* 2012 *New J. Phys.* **14** 113030
- [22] Wang X *et al* 2012 *Phys. Lett. A* **376** 768
- [23] Ereemeev S V *et al* 2011 *Phys. Rev. B* **83** 205129
- [24] Kuroda K *et al* 2013 *Phys. Rev. B* **88** 245308
- [25] Zhao L *et al* 2012 *Appl. Phys. Lett.* **100** 131602
- [26] Garate I and Franz M 2010 *Phys. Rev. Lett.* **104** 146802
- [27] Yu R *et al* 2010 *Science* **329** 61
- [28] Fujita T, Jalil M B A and Tan S G 2011 *Appl. Phys. Express* **4** 094201
- [29] Pesin D and Macdonald A H 2012 *Nat. Mater.* **11** 409
- [30] Song J-H, Jin H and Freeman A J 2010 *Phys. Rev. Lett.* **105** 096403
- [31] Zhang Q *et al* 2012 *ASC NANO* **6** 2345
- [32] Zhao Y *et al* 2013 *Sci. Rep.* **3** 3060
- [33] Wu G *et al* 2013 *Sci. Rep.* **3** 1233
- [34] Li X *et al* 2013 *Chin. Phys. B* **22** 097306
- [35] Nakayama K *et al* 2012 *Phys. Rev. Lett.* **109** 236804
- [36] Menshchikova T V *et al* 2013 *Nano Lett.* **13** 6064
- [37] Ereemeev S V, Koroteev Yu M and Chulkov E V 2010 *JETP Lett.* **91** 387
- [38] Chena C *et al* 2012 *Proc. Natl Acad. Sci.* **109** 3694
- [39] Men'shov V N, Tugushev V V and Chulkov E V 2013 *JETP Lett.* **98** 603

- [40] Shan W-Y, Lu H-Z and Shen S-Q 2010 *New J. Phys.* **12** 043048
- [41] Medhi A and Shenoy V B 2012 *J. Phys.: Condens. Matter* **24** 355001
- [42] Michetti P *et al* 2012 *Semicond. Sci. Technol.* **27** 124007
- [43] BenDaniel D J and Duke C B 1966 *Phys. Rev.* **152** 683
- [44] Tokatly I V, Tsibizov A G and Gorbatsevich A A 2002 *Phys. Rev. B* **65** 165328
- [45] Men'shov V N, Tugushev V V and Chulkov E V 2013 *JETP Lett.* **97** 258
- [46] Men'shov V N *et al* 2010 *Phys. Rev. B* **81** 235212
- [47] Men'shov V N *et al* 2013 *Phys. Rev. B* **88** 224401
- [48] Liu C X *et al* 2010 *Phys. Rev. B* **82** 045122
- [49] Fu L 2009 *Phys. Rev. Lett.* **103** 266801
- [50] Tserkovnyak Y and Loss D 2012 *Phys. Rev. Lett.* **108** 187201
- [51] Yokoyama T, Zang J and Nagaosa N 2010 *Phys. Rev. B* **81** 241410
- [52] Bir G L and Pikus G E 1974 *Symmetry, Strain-Induced Effects in Semiconductors* (New York: Wiley)
- [53] Volkov V A and Pinsker T N 1981 *Sov. Phys. Solid State* **23** 1022
- Enaldiev V V, Zagorodnev I V and Volkov V A 2014 arXiv:1407.0945v1
- [54] Ereemeev S V *et al* 2013 *Phys. Rev. B* **88** 144430
- [55] Blöchl P E 1994 *Phys. Rev. B* **50** 17953
- [56] Kresse G and Furthmüller J 1996 *Phys. Rev. B* **54** 11169
- [57] Kresse G and Joubert D 1999 *Phys. Rev. B* **59** 1758
- [58] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
- [59] Koelling D D and Harmon B N 1977 *J. Phys. C: Solid State Phys.* **10** 3107
- [60] Grimme S 2006 *J. Comput. Chem.* **27** 1787
- [61] Menshchikova T V *et al* 2011 *JETP Lett.* **93** 15
- [62] Hulliger F 1976 *Structural Chemistry of Layer-Type Phases* (Boston: Reidel)
- [63] Grauer D C *et al* 2009 *Mater. Res. Bull.* **44** 1926
- [64] Karpinsky O G *et al* 1998 *J. Alloys Compounds* **265** 170
- [65] De Beule C and Partoens B 2013 *Phys. Rev. B* **87** 115113
- [66] Liu C *et al* 2008 *Phys. Rev. Lett.* **100** 236601
- [67] Knez I, Du R-R and Sullivan G 2011 *Phys. Rev. Lett.* **107** 136603