Manipulating topological-insulator properties using quantum confinement

M Kotulla and U Zülicke
School of Chemical and Physical Sciences and macDiarmid Institute for Advanced Materials and Nanotechnology, Victoria University of Wellington, PO Box 600, Wellington 6140, New Zealand
E-mail: uli.zuelicke@vuw.ac.nz
Keywords: Dirac-like surface states, confinement-modulated band inversion, topological phases
Supplementary material for this article is available online

Abstract
Recent discoveries have spurred the theoretical prediction and experimental realization of novel materials that have topological properties arising from band inversion. Such topological insulators are insulating in the bulk but have conductive surface or edge states. Topological materials show various unusual physical properties and are surmised to enable the creation of exotic Majorana–fermion quasiparticles. How the signatures of topological behavior evolve when the system size is reduced is interesting from both a fundamental and an application-oriented point of view, as such understanding may form the basis for tailoring systems to be in specific topological phases. This work considers the specific case of quantum-well confinement defining two-dimensional layers. Based on the effective-Hamiltonian description of bulk topological insulators, and using a harmonic-oscillator potential as an example for a softer-than-hard-wall confinement, we have studied the interplay of band inversion and size quantization. Our model system provides a useful platform for systematic study of the transition between the normal and topological phases, including the development of band inversion and the formation of massless-Dirac-fermion surface states. The effects of bare size quantization, two-dimensional-subband mixing, and electron–hole asymmetry are disentangled and their respective physical consequences elucidated.

1. Introduction

Topological insulators [1–3] (TIs) have emerged as a new materials class that provides a testing ground for exploring ground-breaking new ideas [4, 5] about the properties of condensed matter. Realizations are possible in electronically two-dimensional [6–10] (2D) or three-dimensional [11] (3D) systems, and simulations using cold atoms subject to artificial gauge fields in optical lattices have been explored [12–16]. Quantum confinement was recognized early on as a mechanism for causing a transition between topological and normal phases [7, 8, 17], as well as a tool to manipulate the massless-Dirac-like surface (edge) states [17–22] of a 3D (2D) TI. Recently, the interplay between topological properties and quantum confinement has been studied as a way to disentangle various phenomena typically associated with topological materials (protected surface states, band inversion, a finite gap) in a controlled fashion [23].

We utilize an effective-model description of quantum wells made from TI materials to map changes in the size of the fundamental gap, the degree of band inversion, and the hybridization of surface states in a detailed way. As we augment the generic $k \cdot p$ Hamiltonian [24, 25] for 3D TIs with a harmonic confinement potential, our work is complementary to previous studies [17–21] where a hard-wall confinement was considered. Further motivation to investigate influences of the confinement strength on 3D-TI properties is provided by the observation that 2D TIs (realized in HgTe quantum wells or simulated by cold atoms, respectively) subjected to soft confinement experience a proliferation of edge modes that modifies topological properties [12, 13, 26, 27]. Our results serve to differentiate truly universal behavior from properties that depend on the material and/or
the dimensionality. We identify Bi\textsubscript{2}Te\textsubscript{3} and Sb\textsubscript{2}Te\textsubscript{3} as materials where effects due to size quantization are sufficiently prominent to be accessible experimentally, whereas Bi\textsubscript{2}Se\textsubscript{3} is typically not in a confinement-dominated regime. Strong mixing of 2D subbands occurs for all layered materials in the limit of thicker samples, resulting in the formation of the 2D-Dirac-like surface states. Electron–hole asymmetry can partly obscure the observability of topological effects. Unraveling the complexity of these various influences enables us to provide insights to guide future experimental study of confined TIs.

The remainder of this paper is organized as follows. In section 2, our effective-model description of a quantum-well-confined 3D TI is introduced and relevant notation established. For instruction and greater clarity, basic ramifications of bare subband quantization and confinement-induced subband mixing are discussed for the electron–hole-symmetric case in sections 3 and 4, before presenting results for the realistic situation with typically sizable electron–hole asymmetry in section 5. Signatures of topological behavior in the subband dispersions (section 3) and in bound-state properties such as the band-inversion-related pseudospin and the surface-state hybridization (section 4) are identified, also with respect to their observability in real (electron–hole-asymmetric) systems. Our results are summarized in section 6 and further illustrated by interactive simulations that are provided as supplementary data available online at stacks.iop.org/NJP/19/073025/mmmedia. Some more mathematical details about the methods used for calculating 3D-TI-layer subband dispersions and eigenstates are given in the appendix.

2. Model for quantum-confined 3D TIs

Our starting point is the effective low-energy model Hamiltonian \[ H = H_0 + H_1 + H_2, \]
with
\[
H_0(k_z, \mathbf{k}_c) = \begin{pmatrix} h_0(k_z) + h_{\parallel}(\mathbf{k}_c) & 0 \\ 0 & h_0(k_z) + h_{\parallel}^*(\mathbf{k}_c) \end{pmatrix},
\]
\[
H_1(k_z) = \begin{pmatrix} 0 & i B_0 k_z \tau_y \\ -i B_0 k_z \tau_y & 0 \end{pmatrix},
\]
\[
H_2(k_z, \mathbf{k}_c) = \begin{pmatrix} \epsilon(k_z, \mathbf{k}_c) \tau_0 & 0 \\ 0 & \epsilon(k_z, \mathbf{k}_c) \tau_0 \end{pmatrix}.
\]

The asterisk indicates complex conjugation. We introduced \( \tau_x, \tau_y, \tau_z \) and \( \tau_0 \) to denote, respectively, the Pauli matrices and the identity matrix acting in the pseudo-spin-1/2 subspaces spanned by the conduction and valence-band-edge states \( |P^{1+}_1, \sigma \rangle, |P^{2+}_2, -\sigma \rangle \) for fixed \( \sigma \), and used the abbreviations
\[
h_0(k_z) = (M_0 + M_1 k_z^2) \tau_z,
\]
\[
h_{\parallel}(\mathbf{k}_c) = M_2 k_c^2 \tau_z + A_0 (k_c \tau_x + k_y \tau_y),
\]
\[
\epsilon(k_z, \mathbf{k}_c) = C_0 + C_1 k_z^2 + C_2 k_c^2.
\]

Somewhat differing values for the band-structure parameters \( M_j, A_0, B_0, C_1 \) and \( C_2 \) applicable to currently available 3D-TI materials have been quoted in the literature \([20, 24, 25, 28, 29]\). To be specific, we use band-structure parameters from table 1 of \([29]\). Although the absolute values differ, basic materials-related trends found in our work generally emerge the same way when previously reported parameter sets are used, e.g., those from table 4 of \([25]\). The fact that \( M_0 < 0 \) embodies the band inversion responsible for the material’s topological properties.

The partition of \( H \) into its three parts \( H_0, H_1 \) and \( H_2 \) is motivated by the intention to discuss the interplay of various TI band-structure features with a confining potential in \( z \) direction. We choose the latter to be in the harmonic–oscillator form given by
\[
H_0(z) = \begin{pmatrix} V(z) \tau_z & 0 \\ 0 & V(z) \tau_0 \end{pmatrix},
\]
\[
V(z) = \frac{\hbar^2}{4M_1} \Omega^2 z^2.
\]

Although it will generally not be a fully realistic approximation for the electronic confinement in thin layers of TI materials, the harmonic–oscillator potential enables an illuminating theoretical treatment of quantum-confinement effects that yields universally applicable conclusions. Furthermore, investigating the features that are specific to a smooth confinement, in contrast to the typically considered abrupt hard-wall \([17–20, 22]\) or infinitely steep inverted-mass \([21]\) type, can yield interesting new insights into basic properties of topological protection \([26, 27]\).
To start with, we consider the basic confinement problem embodied in \( H_0(-i\partial_z, 0) + H_V(z) \), which can be reduced to a Schrödinger equation for the individual blocks representing pseudo-spin-1/2 subspaces that is given by

\[
[h_0(-i\partial_z) + V(z) \tau_z] \psi_0(z) = E_0 \psi_0(z).
\] (4)

It has the energy eigenvalues \( E^{(i)}_{\alpha 0} = \tau [M_0 + \hbar \Omega (n + 1/2)] \) with \( n = 0, 1, \ldots \), and the corresponding eigenstates can be written as \( \psi^{(i)}_{\alpha 0}(z) = \phi_\alpha(z) / l_0 \otimes |\tau_i\rangle \), where \( \phi_\alpha(z) \) are familiar harmonic-oscillator eigenfunctions, and \( |\tau_i\rangle \) denote the eigenstates of \( \tau_i \) with eigenvalues \( \tau = \pm 1 \) that are associated with the opposite-parity basis states. This establishes the characteristic energy scale \( \hbar \Omega \) for size quantization and \( l_0 \equiv \sqrt{2M_0 / (\hbar \Omega)} \) as a measure for system size in the confined dimension. Relating these scales to those relevant for real samples enables useful comparisons and detailed predictions based on our model. The form of the energy eigenvalues \( E^{(i)}_{\alpha 0} \) also foreshadows how the positive size-quantization energy competes with band inversion arising from a negative \( M_0 \).

3. Emergence of size-quantized subband structure and surface states

We now focus on deriving a general description for the in-plane motion of electrons in the confined 3D TI. To be able to express our results in terms of a minimal set of materials-dependent quantities, we introduce the system of interest are then also usefully measured in terms of \( E_i \), yielding the parameters

\[
\gamma_{M_0} = \frac{M_0}{E_1}, \quad \gamma_{\Omega} = \frac{\hbar \Omega}{E_1} = \frac{2M_0}{M_2} (q_z l_0)^2, \quad \gamma_l = \frac{B_q^2}{A_q^2} \frac{M_5}{2M_l}.
\] (5)

Confined motion in 2D direction and free motion in the xy plane are simultaneously contained in the Hamiltonian \( H_0(-i\partial_z, k_x) + H_V(z) \), which is block-diagonal. Focusing on the pseudo-spin-1/2 subspace spanned by the basis states \( |P_{1}^{\pm}, \frac{1}{2}\rangle \) yields the Schrödinger equation

\[
[h_0(-i\partial_z) + h_l(k_x) + V(z) \tau_z] \psi_{k}(z) = E_{k} \psi_{k}(z) .
\] (6)

Looking for its eigenstates in terms of the superposition \( \psi_{nk}(z) = \sum_{\tau=-\pm 1}^{\alpha_{nk}^{(i)}(z)} \psi^{(i)}_{\alpha n_0}(z) \) transforms (6) into an effective 2D Dirac equation for the subband dispersions,

\[
\left[ \frac{\Delta_{nk}}{E_l} \tau_z + \frac{k_x}{q_z} \tau_x + \frac{k_y}{q_z} \tau_y \right] \left\{ \begin{array}{c} a_{nk}^{(+)i} \\ a_{nk}^{(-)i} \end{array} \right\} = \frac{E_{nk}}{E_1} \left\{ \begin{array}{c} a_{nk}^{(+)i} \\ a_{nk}^{(-)i} \end{array} \right\} ,
\] (7a)

\[
\frac{\Delta_{nk}}{E_l} = \gamma_{M_0} + \gamma_{\Omega} \left( n + \frac{1}{2} \right) + \frac{k_y^2}{q_z^2} ,
\] (7b)

which has a gap parameter \( \Delta_{nk} \) that depends both on the confinement and on the wave vector \( k_x \) for in-plane motion. In particular, only the subbands with \( \Delta_{nk} < 0 \) still have inverted character close to their band edges, which implies \( n \leq n_c \) where

\[
n_c = \left\lfloor \frac{-\gamma_{M_0}}{\gamma_{\Omega}} - \frac{1}{2} \right\rfloor .
\] (8)

Thus similarly to the case of 2D TIs realized in semiconductor heterostructures [8, 10], quantum confinement drives transitions between topological and normal phases of individual subbands [17]. For \( \gamma_{\Omega} > -2\gamma_{M_0} \), all subbands are in the normal phase.

Diagonalization of (7a) is straightforward and yields the eigenvalues and eigenstate spinor components

\[
E_{nk, \alpha} = \alpha E_1 \sqrt{\left( \frac{\Delta_{nk}}{E_l} \right)^2 + \left( \frac{k_y}{q_z} \right)^2} ,
\] (9a)

\[
a_{nk, \alpha}^{(+)i} = \frac{i^{\alpha + \tau}}{\sqrt{k_x - \tau i k_y}} \frac{E_{nk, \alpha} + \tau \Delta_{nk}}{2 E_{nk, \alpha}} ,
\] (9b)

where \( \alpha = \pm 1 \) distinguishes positive-energy (conduction) and negative-energy (valence) subbands. The energy spectrum exhibits the axial symmetry of in-plane motion through the dependence of \( E_{nk, \alpha} \) on \( k_x \equiv \sqrt{k_x^2 + k_y^2} \). Due to the block-diagonal form of the Hamiltonian \( H_0(-i\partial_z, k_x) + H_V(z) \), its eigenvalues are those given by (9a); each of them doubly degenerate with corresponding eigenstates \( \psi_{nk, \alpha}(z) \otimes |+\rangle \) and \( \psi_{nk, \alpha}(z) \otimes |-\rangle \).

Here the \( |\sigma_i\rangle \) denote eigenstates of \( \sigma_i \) that distinguish the conduction and valence-band-edge subspaces spanned by \( \{ |P_{1}^{\pm}, \sigma_i \rangle, \sigma_i |P_{2}^{\pm}, -\sigma_i \rangle \} \) with opposite \( \sigma = \pm 1 \). The left panel in figure 1 shows the dispersions obtained for a situation with \( n_c = 1 \).
While the oscillator subbands obtained from diagonalizing \( H_0(-i\partial_z, k_z) + H_V(z) \) illustrate how confinement affects topological band inversion on a basic level, they can only yield a correct description when \( H_0(-i\partial_z) \) is a small perturbation, i.e., in the confinement-dominated limit where \( \gamma_1 \ll \gamma_0 \). InBi\(_2\)Se\(_3\), this is not the typical situation as \( \gamma_1 = 2.34 \) is larger than any reasonably possible value for \( \gamma_0 \), and the inclusion of \( H_0(-i\partial_z) \) significantly affects the subband dispersions. Formally, the coupling of individual oscillator subbands by \( H_0(-i\partial_z) \) as the 3D limit is approached has many similarities with the way models for 3D TIs have been constructed based on transversely coupled 2D TIs \cite{31}. Using the eigenstates of \( H_0(-i\partial_z, k_z) + H_V(z) \) as a basis, we are able to efficiently diagonalize the full Hamiltonian \( H_0(-i\partial_z, k_z) + H(-i\partial_z) + H_V(z) \). See the appendix for mathematical details and the right panel of figure 1 for illustrative results. Most crucially, the band mixing induced by \( H(-i\partial_z) \) conspires to create the surface-state dispersion expected for a confined 3D TI in the large-width limit \cite{17–21}. In contrast to Bi\(_2\)Se\(_3\), the parameter \( \gamma_1 \) is quite small for Bi\(_2\)Te\(_3\) (0.484) and Sb\(_2\)Te\(_3\) (0.117), as compared with the nominal maximum values of \( \gamma_0 \) (4 and 5) derived from equating \( \sqrt{2} l_{q1} \) with their quintuple-layer width \( \approx 1 \) nm. Hence, these materials should lend themselves to a more detailed exploration of the transition between the 3D-bulk and confinement-dominated regimes. The interactive simulations provided in the supplementary data for this article show the evolution of subband structure in confined TIs made from Bi\(_2\)Se\(_3\), Bi\(_2\)Te\(_3\), and Sb\(_2\)Te\(_3\).

The fundamental gap \( \Delta \) between the lowest conduction (sub-)band and the highest valence (sub-)band is a quantity of both conceptual and practical importance. It was shown to decrease exponentially as a function of system size in the confined direction both in 2D \cite{22} and 3D \cite{17–20} TIs. For 3D TIs \cite{18, 32, 33} and inversion-symmetry-breaking 2D TIs \cite{34}, the exponential decay of \( \Delta \) was observed to be modulated by an oscillation. Within our model, we observe the oscillatory behavior, but with interesting materials-dependent features. See figure 2. The more strongly confinement-dominated materials Bi\(_2\)Te\(_3\) and Sb\(_2\)Te\(_3\) (those with a rather small magnitude of the parameter \( \gamma_0 \)) show clear oscillations, and the condition for the bare subband crossings, i.e., vanishing effective Dirac gap given in (7b), quite accurately predicts the values of \( \gamma_0 \) at which \( \Delta \) is minimal. Thus the period of gap oscillations in these two materials is consistent with \( \delta(1/\gamma_0) \sim -1/\gamma_0 \), which corresponds to a period \( b\Delta \sim \sqrt{M_0/M_0} \) for oscillation of the gap as a function of the effective 2D-system width scale \( l_{q1} \). This estimate agrees with similar ones derived previously using a hard-wall confinement \cite{25} and a tight-binding model \cite{32}, respectively, but deviations from such effective-Hamiltonian-based results occur in few-layer samples \cite{29, 35, 36}. In contrast, no gap oscillations are discernible in our model with parameters applicable to Bi\(_2\)Se\(_3\), for which the mixing of bare harmonic-oscillator subbands is always important due to the large value of \( \gamma_0 \). Hence, even though transitions between normal and inverted phases occur typically for the 2D subbands in all three materials (see below), their heralding by gap oscillations depends sensitively on band-structure details

\[ 1 \text{ An order-of-magnitude estimate based on the naïve identification } \sqrt{2} l_{q1} \equiv 1 \text{ nm (approx. value of the Bi}_2\text{Se}_3\text{ quintuple-layer thickness} \text{ [30]) implies } \gamma_0 \leq 0.6 \text{ (using materials parameters from [29]).} \]
and, in contrast to expectation\cite{32}, is therefore not universal. The observed absence (presence) of gap oscillations in Bi$_2$Se$_3$ (Bi$_2$Te$_3$ and Sb$_2$Te$_3$) also suggests that the system is dominantly a strong (weak) TI\cite{38}.

4. Confinement tuning of band inversion and surface-state hybridization

The band inversion in TI materials can be measured directly by the pseudo-spin projection of energy eigenstates. Within our present notation, ordinary conduction (valence) band character of a state is quantified by the expectation value $z_{\text{ts}}$ being close to $+1 (-1)$. We use this measure to study the fate of band inversion in quantum-confined TI systems, complementing recent experimental studies\cite{23}. Within our formalism where a general state is given as a superposition

$$\Psi_k(z) = \sum_{n\alpha} [b_{nk}^{(\alpha+)} \psi_{nk,\alpha}(z) \otimes |+\rangle_n + b_{nk}^{(\alpha-)} \psi_{nk,\alpha}(z) \otimes |-\rangle_n],$$

its pseudo-spin expectation value is straightforwardly calculated as

$$\langle \tau_z \otimes \sigma_0 \rangle_k = \sum_{n\alpha,\beta} \sum_{\sigma} \frac{\tau}{2} \left( 1 + \frac{\Delta_{nk}}{E_{nk,\alpha}} \right) \left( 1 + \frac{\Delta_{nk}}{E_{nk,\beta}} \right) \text{Re} \left[ \langle \psi_{nk,\alpha}^{(\alpha\sigma)} | \psi_{nk,\beta}^{(\beta\sigma)} \rangle \right].$$

As the thickness of the TI layer decreases, more and more 2D subbands (those with band index $n \gg n_c$) show normal behavior, with conduction and valence band states having the ordinary pseudo-spin character. Low-energy subbands with $n \ll n_c$ continue to have band inversion, but the lowest-energy ones become increasingly dominated by confinement effects and less by the band mixing arising from $H_k$. In particular, the lowest subband oscillates between normal and inverted character in the strongly confinement-dominated regimes exhibited by Bi$_2$Te$_3$ and Sb$_2$Te$_3$. Figure 3 shows examples of such features using materials parameters applicable to Bi$_2$Se$_3$. More extensive simulations provided as supplementary data illustrate the evolution of band inversion as the confinement-related parameter $\gamma_3$ is varied continuously in the three materials of interest. Generally, an oscillating change between normal and inverted character of the lowest subband is observed to follow the cycle of closings for the bare gap parameter $\Delta_{nk}$ from equation (7b). To a lesser extent, this is the case.
for the not-confinement-dominated Bi$_2$Se$_3$. Generally, the magnitude of the gap appears to be more strongly affected by band mixing than the pseudo-spin character of the size-quantized subband states. Thus it can be misleading to use the measured sequence of gap minima to infer topological and normal regimes in a confined system, whereas the pseudo-spin always serves as a reliable identifier. The cycle of band inversions occurring in the bare harmonic–oscillator part of the model embodied in $H_0 = \frac{-\tilde{\epsilon}}{2} \pm \tilde{\varepsilon}_l - \frac{1}{2} \tilde{\epsilon} z^2$ generally coincides quite well with the succession of band inversions for the lowest 2D subbands obtained from diagonalizing the full model $H = H_0 + H_{\text{mix}}$, where $H_{\text{mix}}$ incorporates the effect of band mixing.

Whether a 2D subband is inverted or normal is conventionally determined by the character of states at the band edges. At finite in-plane wave vector $k_\perp$, inter-band coupling due to the term proportional to $A_0$ in (2b) induces a mixing of pseudo-spin states, and the term proportional to $M_2$ contributes to the effective band gap. As a result, a crossover from inverted to normal character generally occurs at large $k_\perp$, as can be seen for the first and second subbands shown in the left panel of figure 3. Interestingly, for low-lying 2D subbands and in the confinement-dominated regime, band inversion can also just occur within a region of finite $k_\perp$. See the right panel of figure 3 for a pertinent example, and the simulations provided as supplemental data for further illustration. Again, it should be emphasized that our results regarding the evolution of band inversions with quantum confinement are based on the effective-Hamiltonian approach from which we expect deviations to occur in few-layer samples [29,35,36].

In a thick sample, the lowest-lying subband corresponds to states localized at the system’s boundaries. As confinement is increased, i.e., sample thickness reduced, hybridization of states from opposite surfaces becomes important, and the states progressively loose their boundary-localized character. Eventually, even the lowest-lying subband-edge states are extended over the entire 2D-layer width. This general behavior is instructively demonstrated by the shape of probability-density distributions of lowest-subband-edge states shown as part of the interactive simulations provided in the supplementary data. To enable a more quantitative exploration of the transition between the boundary-localized and 2D-layer-extended regimes, we introduce $b_0 = |W_{b_0}(z)|^2$ for $k_\perp = 0$ and $z = 0$ as a measure for the hybridization-induced 2D-bulk character of subband-edge states. See figure 4. In Bi$_2$Se$_3$, small abrupt changes associated with band inversions occur on top of a systematic change to the strongly hybridized regime for $1/\gamma_1 \lesssim 3$. The practically vanishing hybridization for a broad range of confinement strengths is indicative of well-defined surface states. In contrast, the transition between surface-localized lowest-energy states to 2D-bulk-like behavior is sharper in the confinement-dominated case of Sb$_2$Te$_3$, but, at the same time, hybridization remains substantial even for the larger-thickness samples.
5. Effect of electron–hole asymmetry

So far, we have neglected the influence of asymmetries between conduction and valence bands that are embodied in the contribution $H_{\text{k}\text{k}'}$ to the TI-material Hamiltonian given in (1c). Comparing the magnitudes of $C_1$ and $C_2$ with those of $M_1$ and $M_2$, respectively, in real materials would suggest that electron–hole asymmetry is generally not a small effect. Here we investigate the latter’s influence on the modification of topological properties by quantum confinement.

The effect of the electron–hole-symmetry-breaking contribution $H_{\text{k}\text{k}'}$ on the subband dispersions generally turns out to be indeed quite large, as a comparison of results corresponding to Bi$_2$Se$_3$ shown in the left panels of figures 3 and 5 clearly demonstrates. Most importantly, the position of the neutrality point for the massless-Dirac-like surface states is shifted in energy, and the conduction and valence subbands acquire different effective masses, even leading to a mass inversion occurring in a valence subband. Significant qualitative differences in the subband structure for Sb$_2$Te$_3$, shown in the right panel of figure 5, are a consequence of the different magnitudes and opposite sign for the bulk-bandstructure parameters entering $H_{\text{k}\text{k}'}$.

Given the drastic changes exhibited in the lineshapes of subband spectra, one may ask whether, and if so how, any of the above-discussed features associated with the interplay of topological properties and quantum confinement survive in the presence of electron–hole asymmetry. We focus here specifically on the fundamental (i.e., lowest-subband energy) gap and the band inversion. Interactive simulations are included as supplemental.
data to visualize the situation. As it turns out, in the confinement-dominated regime, electron–hole asymmetry is found to lead only to marginal quantitative changes. Basic qualitative features such as gap oscillations and subband inversions occur very similarly to the situation where $H_k (k_x, k_y)$ is neglected, with the gap oscillations typically more significantly affected [37]. This is particularly clearly exhibited by the case of Sb$_2$Te$_3$ for not-too-weak confinement. But also for Bi$_2$Se$_3$ where mixing of bare-oscillator subbands is strong, the evolution of band inversions is basically identical with and without electron–hole asymmetry included, even for the lowest subband and weakest confinement strengths. Interestingly, for the system whose parameters correspond to Sb$_2$Te$_3$, the scrambling of low-energy subbands in the limit of very weak confinement tends to obscure the TI properties that are still unambiguously exhibited when electron–hole asymmetry is neglected.

### 6. Conclusions and outlook

We have used the effective-model description of bulk-TI band structures to investigate the effect of a soft, harmonic-oscillator-type, quantum confinement on physical properties that epitomize topological phases. Our study focuses explicitly on three materials systems (Bi$_2$Se$_3$, Bi$_2$Te$_3$ and Sb$_2$Te$_3$) that are representatives for quite different regions of the 3D-TI parameter space and, hence, exhibit distinctive features in the evolution of topological properties as the strength of the confinement is varied. Interactive simulations have been included as supplemental data to enable more detailed exploration of these differences. The results obtained here are also straightforwardly generalized to any materials systems whose band structure is described by the same type of $k \cdot p$ Hamiltonian that forms the basis for our theoretical approach, including Dirac semimetals [39–41] and other surmised TIs such as Bi$_2$Te$_3$L$_2$ [42].

The interplay of band inversion, size quantization, and band mixing is found to be governed by the relative magnitudes of unit-less parameters defined in equation (5). Characteristic features exhibited by the particular materials systems considered here can thus be rationalized in terms of the specific values of these parameters. Fundamentally, as the system size in the confined direction varies, the gaps of inverted bare-oscillator subbands (equation (7b)) are successively closing and reopening. Mixing of bare-oscillator subbands significantly modifies the bare-oscillator subband dispersions in the large-width regime ($\gamma_1 < \gamma_2$), establishing the vanishing-mass Dirac-like surface-state dispersion and eventually causing the disappearance of oscillations in the fundamental (lowest-subband) gap value. We track the evolution of band inversions by explicit consideration of the pseudo-spin value of eigenstates, establishing the robustness of TI phases with respect to band mixing and electron–hole asymmetry. In contrast, the occurrence of gap oscillations turns out to be a nonuniversal feature that is generally not a useful measure to monitor TI character in a quantum-confined system.

Our study considering 3D TIs subject to the harmonic-oscillator confining potential complements previous works that assumed hard-wall confinement. In contrast to the situation reported in single-particle analyses of 2D TIs [12, 15, 26], the basic features and overall trends associated with the effects of size quantization on 3D-TI properties do not drastically differ between the cases of hard-wall and soft-harmonic-oscillator potentials. Instead of the type of confining potential, differences in the 3D-TI bandstructure parameters are a major cause of significant variability in the exhibited behaviors.

The conclusions presented in this work have been reached based on an approach where certain aspects of real materials were ignored. One of these is the finite size of thin-film samples in the plane perpendicular to the quantum-confined direction and any effects arising from the conducting lateral surfaces. However, apart from the edge-state structure in strong perpendicular magnetic fields [43], systems with a large-enough aspect ratio are well-described by models assuming them to be infinite in the transverse directions. Another potentially relevant issue concerns the interplay of confinement and interactions. For example, the presence of long-range (Coulomb) interactions is known to induce a reconstruction of the boundary-electronic (edge-state) structure in 2D quantum-Hall [44–47] and quantum-spin-Hall [27, 48] systems. Although interaction effects often turn out to be less pronounced in higher spatial dimensions, it would be interesting for future studies to address the influence of Coulomb interactions on harmonically confined 3D TIs, especially those where strong correlations are expected to affect topological properties [49, 50]. The goal of such investigations would be to establish how, and where in parameter space, the purely single-particle-related confinement effects in 3D TIs discussed in the present work are modified due to interactions.

### Acknowledgments

Useful discussions with D Abergel, P Brydon, M Governale, K I Imura, and H Z Lu are gratefully acknowledged. We used a software tool available from [51] to create the interactive graphics animations of confined-3D-TI subband structures that are made available as supplementary data in conjunction with this article.
Here we provide some mathematical details for solving the Schrödinger problem for a quantum-confined 3D TI.

Making the ansatz (10) transforms (A.1) into a set of recursive relations

\[ \frac{E_{nk,β} - E_k}{E_k} b_{mk}^{(α)} + σ i \sum_α [C^{m-1, m}_{α/β, k} b_{m-1, k}^{(α-γ)} + C^{m+1, m}_{α, k} b_{m+1, k}^{(α-γ)\dagger}] = 0 \]

for the unknown coefficients \( b_{mk}^{(α)} \), with the matrix elements given explicitly by

\[ C^{n,m}_{α/β, k} = \frac{m! γ^3}{8 E_{nk, α} E_{nk, β}} \sqrt{(E_{mk, β} - Δ_{mk})(E_{nk, α} - Δ_{nk})} \]

We truncate this infinite set of equations by writing a matrix equation, \( \mathcal{H}_N b_N = E_N b_N \), with \( N \) large enough as required to achieve numerical accuracy, and the definitions

\[ b_N = (b_{0,0}^{(+)}, b_{0,0}^{(-)}, b_{1,1}^{(+)}, b_{1,1}^{(-)}, \ldots, b_{N-1,N-1}^{(-)})^T \]

\[ \mathcal{H}_N = \begin{pmatrix}
  \mathcal{E}_0' & \Gamma_0 & 0 & \cdots & 0 \\
  \Gamma_0^T & \mathcal{E}_1' & \Gamma_1 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \ddots & \vdots \\
  0 & \cdots & \cdots & \cdots & \Gamma_{N-1} \\
 \end{pmatrix}, \quad \Gamma_n = \begin{pmatrix} 0 & iC_n & 0 \\ 0 & -iC_n & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}, \quad \mathcal{E}_n' = \begin{pmatrix} C_{n+1,n}^{+} \\ C_{n+1,n}^{-} \\ \vdots \\ C_{N,n}^{+} \\ C_{N,n}^{-} \end{pmatrix} \]

For the case with electron–hole asymmetry included, the contribution \( H_r ( -i \partial_x, k_x ) \) from (1e) needs to be added to the Hamiltonian. As a result, the recursion relations (A.2) are modified and now read

\[ \frac{E_{mk,β} - E_k}{E_k} + γ \gamma \frac{m + 1}{2} + C_0 \frac{k^2}{M q^2} \]

\[ + σ i \sum_α [C^{m-1, m}_{α/β, k} b_{m-1, k}^{(α-γ)} + C^{m+1, m}_{α, k} b_{m+1, k}^{(α-γ)\dagger}] = 0, \]

with \( γ = \gamma \gamma / (2M) \) and

\[ D^{n,m}_{α/β, k} = \frac{\gamma \gamma (m + 1)}{4 E_{nk, α} E_{nk, β}} \sqrt{(E_{mk, β} - Δ_{mk})(E_{nk, α} + Δ_{nk})} \]

Consequently, this leads to an expression similar to (A.5),

\[ \mathcal{H}_N = \begin{pmatrix}
  \mathcal{E}_0 & \Gamma_0 & \Delta_0 & 0 & \cdots & 0 \\
  \Gamma_0^T & \mathcal{E}_1 & \Gamma_1 & \Delta_1 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
  \Delta_0^T & \Gamma_{N-1} & \mathcal{E}_{N-1} & \Gamma_{N-1} & \Delta_{N-1} & 0 \\
 \end{pmatrix}, \quad \mathcal{E}_n = \mathcal{E}_n' + E_k \text{diag} \left[ γ \frac{m + 1}{2} + C_0 \frac{k^2}{M q^2} \right], \]
\begin{equation}
\Delta_n = \begin{pmatrix} D_n & 0 \\ 0 & D_n \end{pmatrix}, \quad D_n = \begin{pmatrix} D_{n,n+2}^{n+1-k} & D_{n,n+2}^{n+1+k} \\ D_{n,n+2}^{n+1-k} & D_{n,n+2}^{n+1+k} \end{pmatrix}
\end{equation}

\textbf{ORCID}

M Kotulla \ ORCID \ https://orcid.org/0000-0003-2068-828X

U Zülicke \ ORCID \ https://orcid.org/0000-0001-5055-3330

\textbf{References}

[1] Hasan M Z and Kane C L 2010 Rev. Mod. Phys. \textbf{82} 3045–67
[2] Qi X L and Zhang S C 2011 Rev. Mod. Phys. \textbf{83} 1057–110
[3] Hasan M Z, Xu S Y and Bian G 2015 Phys. Scr. \textbf{T164} 014001
[4] Chiu C K, Teo I C Y, Schneyder A P and Ryu S 2016 Rev. Mod. Phys. \textbf{88} 035005
[5] Ludwig A W W 2016 Phys. Scr. \textbf{T168} 014001
[6] Kane C L and Mele E J 2005 Phys. Rev. Lett. \textbf{95} 146802
[7] Bernevig B A, Hughes T L and Zhang S 2006 Science \textbf{314} 1757–61
[8] König M, Wiedmann S, Brüne C, Roth A, Buhmann H, Molenkamp L W, Qi X L and Zhang S C 2007 Science \textbf{318} 766–70
[9] Liu C, Hughes T L, Qi X, Wang K and Zhang S 2008 Phys. Rev. Lett. \textbf{100} 236601
[10] Knez J, Du R and Sullivan G 2011 Phys. Rev. Lett. \textbf{107} 136603
[11] Hasan M Z and Moore E J 2011 Annu. Rev. Condens. Matter Phys. \textbf{2} 55–78
[12] Stanescu T D, Galitski V, Vaishnav J Y, Clark C W and Das Sarma S 2009 Phys. Rev. A \textbf{79} 053639
[13] Stanescu T D, Galitski V and Das Sarma S 2010 Phys. Rev. A \textbf{82} 013608
[14] Bermudez A, Mazza L, Rizzi M, Goldman N, Lewenstein M and Martin-Delgado M A 2010 Phys. Rev. Lett. \textbf{105} 190404
[15] Goldman N, Satija I, Nikolic P, Bermudez A, Martin-Delgado M A, Lewenstein M and Spielman I B 2010 Phys. Rev. Lett. \textbf{105} 235302
[16] Béri Band Cooper N R 2011 Phys. Rev. Lett. \textbf{107} 145301
[17] Liu C X, Zhang H, Yan B Q, Qi X L, Frauenheim T, Dai X, Fang Z and Zhang S C 2010 Phys. Rev. B \textbf{81} 041307
[18] Linder J, Yokoyama T and Sudbo A 2009 Phys. Rev. B \textbf{80} 205401
[19] Lu H Z, Shan W Y, Yao W, Niu Q and Shen S Q 2010 Phys. Rev. B \textbf{81} 115407
[20] Shan W Y, Lu H Z and Shen S Q 2010 New J. Phys. \textbf{12} 043048
[21] Zhang F, Kane C L and Mele E J 2012 Phys. Rev. B \textbf{86} 081303
[22] Zhou B, Lu H Z, Chu R L, Shen S Q and Niu Q 2008 Phys. Rev. Lett. \textbf{101} 246807
[23] Salah M, Shapourian H, Koirala N, Brahlek M J, Moon J and Oh S 2016 Nano Lett. \textbf{16} 5328–32
[24] Zhang H, Liu C X, Qi X L, Dai X, Fang Z and Zhang S C 2009 Nat. Phys. \textbf{5} 438–42
[25] Liu C X, Qi X L, Zhang H, Dai X, Fang Z and Zhang S C 2010 Phys. Rev. B \textbf{82} 045122
[26] Buchhold M, Cocke D and Hofstetter W 2012 Phys. Rev. A \textbf{86} 063614
[27] Wang J, Meir Y and Gefen Y 2017 Phys. Rev. Lett. \textbf{118} 046801
[28] Ortila M et al 2015 Phys. Rev. Lett. \textbf{114} 186401
[29] Nechaev I A and Krasovskii E E 2016 Phys. Rev. B \textbf{94} 201410
[30] Zhang G, Qin H, Teng J, Guo J, Guo Q, Dai X, Fang Z and Wu K 2009 Appl. Phys. Lett. \textbf{95} 053114
[31] Kobayashi K, Yoshimura Y, Imura K I and Ohtsuki T 2015 Phys. Rev. B \textbf{92} 235407
[32] Ozawa H, Yamakage A, Sato M and Tanaka Y 2014 Phys. Rev. B \textbf{90} 045309
[33] Betancourt J, Li S, Dang X, Burton I D, Tsymbal E Y and Velev J 2016 J. Phys.: Condens. Matter \textbf{28} 393501
[34] Takagaki Y 2014 Phys. Rev. B \textbf{90} 165305
[35] Förster T, Krüger P and Rohlfing M 2015 Phys. Rev. B \textbf{92} 201404
[36] Förster T, Krüger P and Rohlfing M 2016 Phys. Rev. B \textbf{93} 205442
[37] Okamoto M, Takekuma Y and Imura K I 2014 Phys. Rev. B \textbf{89} 125425
[38] Imura K I, Okamoto M, Yoshimura Y, Takekuma Y and Ohhitsu K T 2012 Phys. Rev. B \textbf{86} 245436
[39] Wang Z, Weng H, Wu Q, Dai X and Fang Z 2013 Phys. Rev. B \textbf{88} 125427
[40] Xiao Q, Yang S A, Liu Z, Li H and Zhou G 2015 Sci. Rep. \textbf{5} 7898
[41] Pan H, Wu M, Liu Y and Yang S A 2015 Sci. Rep. \textbf{5} 14639
[42] Nechaev I A, Eremin S V, Krasovskii E E, Echenique P M and Chulkov E V 2017 Sci. Rep. \textbf{7} 43666
[43] Liu Z, Jiang L and Zheng Y 2016 J. Phys.: Condens. Matter \textbf{28} 275501
[44] Dempsey J, Gelland B Y and Halperin B I 1993 Phys. Rev. B \textbf{47} 3639–42
[45] MacDonald A H, Yang S R E and Johnson M D 1993 Annu. J. Phys. \textbf{46} 345–58
[46] Channon C D and Wen X G 1994 Phys. Rev. B \textbf{49} 8227–41
[47] Barlas Y, Jogekar Y N and Yang K 2011 Phys. Rev. B \textbf{83} 205307
[48] Amaaricci A, Privitera L, Petocchi F, Capone M, Sangiovanni G and Trauzettel B 2017 Phys. Rev. B \textbf{95} 205120
[49] Hohenadler M and Assaad F F 2013 J. Phys.: Condens. Matter \textbf{25} 143201
[50] Amaaricci A, Budich J C, Capone M, Trauzettel B and Sangiovanni G 2016 Phys. Rev. B \textbf{93} 235112
[51] Nöckel U \ http://mathematica.stackexchange.com/questions/18257/how-to-translate-interactive-graphics-from-mathematica-to-standard-htmlsvg (Accessed: 21 November 2016)