Heavy Dirac fermions in a graphene/topological insulator hetero-junction

Wendong Cao¹, Rui-Xing Zhang¹, Peizhe Tang¹, Gang Yang², Jorge Sofo¹, Wenhui Duan¹,³,⁴ and Chao-Xing Liu¹

¹ State Key Laboratory of Low-Dimensional Quantum Physics, Department of Physics, Tsinghua University, Beijing 100084, People’s Republic of China
² Department of Physics, The Pennsylvania State University, University Park, PA 16802-6300, USA
³ Department of Physics, McCullough Building, Stanford University, Stanford, CA 94305-4045, USA
⁴ Institute for Advanced Study, Tsinghua University, Beijing 100084, People’s Republic of China

E-mail: cxl56@psu.edu

Keywords: graphene, topological insulators, quantum anomalous Hall effect

Abstract

The low energy physics of both graphene and surface states of three-dimensional topological insulators (TIs) is described by gapless Dirac fermions with linear dispersion. In this work, we predict the emergence of a ‘heavy’ Dirac fermion in a graphene/TI hetero-junction, where the linear term almost vanishes and the corresponding energy dispersion becomes highly nonlinear. By combining ab initio calculations and an effective low-energy model, we show explicitly how strong hybridization between Dirac fermions in graphene and the surface states of TIs can reduce the Fermi velocity of Dirac fermions. Due to the negligible linear term, interaction effects will be greatly enhanced and can drive ‘heavy’ Dirac fermion states into the half quantum Hall state with non-zero Hall conductance.

1. Introduction

Two-dimensional Dirac physics has aroused great interest in condensed matter physics ever since the discovery of graphene [1–3] and topological insulators (TIs) [4–6] due to its importance in both fundamental physics and device applications. In graphene, gapless Dirac cones exist at the momenta \( K \) and \( K’ \) point in the Brillouin zone (BZ) with a large Fermi velocity \( \sim 1 \times 10^6 \text{ m s}^{-1} \), which is about 1/300 of the speed of light [7], and results in a high mobility for electron transport in graphene. Due to its high mobility, graphene is believed to possess the potential in the applications involving fast speed electronic devices [8–10]. On the other hand, when a gap (mass) is opened for Dirac cones, the so-called ‘parity anomaly’, which was first known in high energy physics [11], can occur [12] and lead to a large variety of topological states, including the quantum anomalous Hall (QAH) effect [12] and the quantum spin Hall (QSH) effect [13], in graphene. Furthermore, it was predicted that these topological states in graphene can be spontaneously induced by interaction, leading to the so-called ‘topological Mott insulators’ [14]. However, due to the large Fermi velocity, the density of states (DOS) vanishes rapidly near the Dirac cone, and the critical interaction strength is relatively large. Thus, the reduction of Fermi velocity in Dirac fermions (or equivalently a ‘heavy’ Dirac fermion) is valuable for exploring new topological states in the Dirac systems.

In this work, we explore Dirac physics in a hetero-junction with graphene on top of a TI. The QSH effect with a large energy gap has been predicted for graphene coupled to TI thin films [15, 16]. Here we focus on the case of graphene coupled to a single surface state of a three-dimensional TI, which can be realized in experiments when the thickness of the TI is large enough so that the topological surface states on opposite surfaces are decoupled. In this case, we find that due to the strong hybridization between the four Dirac cones from graphene (two valleys and two spins) and one Dirac cone from the topological surface state, four of the resulting states will be gapped while one remains gapless with the Fermi velocity significantly reduced. By constructing an effective model of this system based on first principles calculations, we explore the underlying physical reason why the linear dispersion relation term is so small. As a consequence, the ‘heavy’ Dirac fermion has a much larger DOS near the Dirac point, and a relatively small
interaction can drive the system into a gapped topological phase with non-zero Hall conductance (parity anomaly). Below, we will first construct an effective model for this hetero-junction that reproduces the physics observed in our first-principles calculations, and then study the effects of interactions at the mean field level.

2. Effective model for graphene/TI hetero-junctions

Firstly, we perform ab initio calculations for a hetero-junction with graphene on a Sb$_2$Te$_3$ film with eight quintuple (QL) layers, as shown in figure 1(a). Due to the lattice mismatch between graphene and Sb$_2$Te$_3$ films (substrate), the slab model used in this calculation contains one unit cell of Sb$_2$Te$_3$ films [17], whose in-plane lattice constant corresponds to that of the $\sqrt{3} \times \sqrt{3}$ supercell of graphene with 0.3% compression. The top view of the most stable configuration is shown in figure 1(b) in which the topmost Te atom located at the center of the hexagonal ring of graphene [16]. From our calculations, it is found that the equilibrium distance between graphene and Sb$_2$Te$_3$ is 3.456 Å and the binding energy is 41.9 meV per carbon atom. Since the $\sqrt{3} \times \sqrt{3}$ supercell of graphene is used in our simulation, the BZ of the hetero-junction is folded, and reduced to one third of the original BZ of the intrinsic graphene (the original and folded BZs are shown in figure 1(c)), and thus the $K_0$ and $K'_0$ in the original BZ are mapped into $\Gamma$ in the folded BZ. Consequently, the Dirac cones from both graphene and TI appear at the same momentum $\Gamma$, and we only need to focus on the low energy physics around the $\Gamma$ point. From the ab initio calculations, the band structure of the hetero-junction is shown in figure 1(d) and its low energy bands are zoomed in (dotted lines) in figure 1(e). Spin–orbit coupling (SOC) is included in the calculations and more details about ab initio calculation method can be found in the appendix A. It can be seen in figure 1(d) that there is a large hole pocket on the $\Gamma$–$M$ line. However, since the hole pocket is far away from the Dirac cone, it will not influence surface energy dispersion around the $\Gamma$ point in the junction system. Also, in principle, one can consider applying a top gate to eliminate the hole pocket by tuning the chemical potential of this junction structure. In figure 1(e), via analyzing the wave functions at the $\Gamma$ point around the Fermi level, we find that the bottom surface states of TIs has no hybridization with graphene and contribute to the gapless Dirac cone labelled by blue-dotted lines in figure 1(e). The bands labelled by green-dotted lines originate from the hybridization of energy bands in graphene and top surface states of TIs. Therefore, we focus on these energy bands (marked by the green-dotted lines) and unveil underlying physics for the ‘heavy’ Dirac Fermion.

To understand the hybridization between these Dirac fermions, we first neglect the topological surface states of Sb$_2$Te$_3$ films and consider only the $p_z$ orbitals of graphene under the environment of the Sb$_2$Te$_3$ substrate. For intrinsic graphene, there are two sub-lattice sites in one unit cell and the low energy physics can be effectively described by a two-dimensional Dirac type of Hamiltonian around the $K_0$ and $K'_0$ points in the original BZ [3]. In contrast, due to the influence of the substrate, one unit cell of this hetero-junction contains a hexagon (benzene ring) with six equivalent carbon sites and the hopping terms between them could be described by three hopping parameters, denoted as $t_1$, $t_2$ and $t_3$ in figure 1(b). Therefore, a $6 \times 6$ tight-binding model [3] on the basis of $(p_z, n)$, $n = 1, \ldots, 6$ is used to describe this system (see equation (B.1) in the appendix B) in which, for the basis $(p_z, n)$, $p_z$ is the atomic orbital and $n$ represents the site index.

For convenience, we perform a unitary transformation to change the basis into eigenstates of the rotation operator. We consider the six-fold rotation operation $C_6$ and the corresponding rotation symmetric basis are denoted as $C_6 |L_z'\rangle = \exp\left(-iL'_z\frac{\pi}{6}\right) |L_z\rangle$, where $L'_z = \pm 2, \pm 1, 0, 3$ labels six eigen states for $C_6$ rotation. (The detailed form of unitary transformation, as well as the basis $|L_z'\rangle$, is given by equation (B.2) in the appendix B and the form of the effective Hamiltonian expanded around $k = 0$ is given by equation (B.3) in the appendix B). We emphasize that six-fold rotation symmetry only exists in pristine graphene ($t_1 = t_2 = t_3$), in which two gapless Dirac fermions appear under the basis $|L'_z = \pm 2\rangle$ and $|L'_z = \pm 1\rangle$ and the states labelled by $|L'_z = 0\rangle$ and $|L'_z = 3\rangle$ are gapped. This is because the states $|L'_z = \pm 2\rangle$ and $|L'_z = \pm 1\rangle$ originate from the states in the $K_0$ and $K'_0$ while the states $|L'_z = 0\rangle$ and $|L'_z = 3\rangle$ from those at $\Gamma_0$ in the original BZ. Therefore, in the following discussion, we only focus on the basis $|L'_z = \pm 2\rangle$ and $|L'_z = \pm 1\rangle$. Due to SOC, the orbital angular momentum is not a good quantum number and the total angular momentum should be considered. Thus, these eigenstates can be labeled by $|J'_{\pm}\rangle = \frac{3}{2}, \frac{1}{2}\rangle$, $\frac{1}{2}, \frac{3}{2}\rangle$, $\frac{1}{2}, \frac{1}{2}\rangle$, $\frac{5}{2}, \frac{1}{2}\rangle$, $\frac{5}{2}, \frac{3}{2}\rangle$, $\frac{1}{2}, \frac{5}{2}\rangle$ and $\frac{3}{2}, \frac{3}{2}\rangle$.

Due to the hybridizations with the Sb$_2$Te$_3$ substrate, the six-fold rotation symmetry $C_6$ is broken down to three-fold rotation $C_3$, which is known as Kekulé modulation [18]. In this case, we may re-label our basis states $|L'_z = \pm 2\rangle$ and $|L'_z = \pm 1\rangle$ by the eigenvalues of $C_3$ rotation, denoted as $|L_z, \eta\rangle$, where $C_3 |L_z, \eta\rangle = \exp\left(-iL_z\frac{2\pi}{3}\right) |L_z, \eta\rangle$ with $L_z = \pm 1, \eta = \pm 1$. The index $\eta$ is introduced to distinguish two degenerate states with the same $L_z$. The explicit form of $|L_z, \eta\rangle$ in the basis set of $(p_z, n)$ and more details about the additional index $\eta$ can be found in the appendix C. Under the basis $|+1, +1\rangle, |-1, -1\rangle, |-1, +1\rangle$ (or equivalently $|L'_z = -2\rangle, |\Gamma\rangle, |2\rangle, |1\rangle$), the effective Hamiltonian is written as
\[
H_{0.4 \times 4} = \begin{pmatrix}
-\Delta \cos \theta & i \hbar v_F^1 k_+ & 0 & \Delta \sin \theta \\
\hbar v_F^1 k_- & +\Delta \cos \theta & \Delta \sin \theta & 0 \\
0 & -\Delta \sin \theta & -\Delta \cos \theta & -i \hbar v_F^2 k_- \\
\Delta \sin \theta & 0 & -i \hbar v_F^2 k_+ & +\Delta \cos \theta
\end{pmatrix}
\]

where \(v_F^1\) is the Fermi velocity of graphene, \(k_\pm = k_x \pm i k_y\), \(\Delta = \sqrt{t_1^2 + t_2^2 + t_3^2 - t_1 t_2 - t_2 t_3 - t_3 t_1}\), and \(\tan \theta = \sqrt{3} (t_1 - t_2)/(t_1 + t_2 - 2t_3)\). The Kekulé modulation terms, described by two independent parameters \(\Delta\) and \(\theta\), can lead to a gap opening with the size of \(2\Delta\) for both Dirac cones.

When SOC is considered, the basis set \(\{|L_z, \eta\}\) should be enlarged and \(L_z\) is replaced by the total angular momentum along the \(z\)-direction \(L_z = L_z + L_z\). The new basis states for graphene take the form of

\[
\begin{pmatrix}
L_z = \frac{3}{2}, \eta = +1 \\
L_z = \frac{1}{2}, \eta = +1 \\
L_z = \frac{1}{2}, \eta = -1
\end{pmatrix}
\]
is the Fermi velocity of top surface states, and $\mu^S$ denotes the corresponding chemical potential. The basis states for topological surface states can also be labelled by their eigenvalues of $I_\sigma \left( \left| J^S = \pm \frac{1}{2} \right\rangle \right)$.

Furthermore, we can construct the hybridization Hamiltonian $H_{GS}$ between graphene and TIs directly from the hopping process between the $p_z$ orbitals of carbon atoms and all the $p$ orbitals of Se and Te atoms. The explicit form of $H_{GS}$ is shown in equation (B.7) in the appendix B.

With the effective Hamiltonian $H_{\text{full}}$ in equation (2), we can write the energy spectrum of graphene/TIs hetero-structure, as shown by red lines in figure 1(e). The corresponding parameters for $H_{\text{full}}$ are listed in the table B.1 of the appendix B. The effective model reproduces well the band structure from first-principles calculations (dotted-green lines), especially for the bands labeled by $\Phi_{\text{Kekulé}}^1$, $\Phi_{\text{Kekulé}}^2$, $\Phi_{\text{Rashba}}^1$, and $\Phi_{\text{Rashba}}^2$ in figure 1(e). The bands $\Phi_{\text{Rashba}}^1$ and $\Phi_{\text{Rashba}}^2$ mainly come from the hybridization between the topological surface states $\left\{ \left| J^S = \pm \frac{1}{2} \right\rangle \right\}$ and graphene states $\left\{ \left| -\frac{1}{2}, \pm 1 \right\rangle \right\}$ and $\left\{ \left| -\frac{1}{2}, \pm 1 \right\rangle \right\}$ (or equivalently $\left| J' = -\frac{1}{2}, \uparrow \right\rangle \right\}$ and $\left\{ \left| J' = \frac{1}{2}, -\downarrow \right\rangle \right\}$). Due to strong SOC in TIs, these bands reveal strong Rashba type of spin splitting and thus are labeled by ‘Rashba’ bands. Both $\Phi_{\text{Kekulé}}^1$ and $\Phi_{\text{Kekulé}}^2$ originate from four graphene states $\left\{ \left| -\frac{1}{2}, \pm 1 \right\rangle \right\}$ and $\left\{ \left| -\frac{1}{2}, \pm 1 \right\rangle \right\}$ (or equivalently $\left\{ \left| J' = \frac{1}{2}, -\downarrow \right\rangle \right\}$). It should be noted that all these four states can hybridize with each other since the $C_6$ rotation symmetry is broken to the $C_3$ rotation symmetry due to the Kekulé modulation (thus dubbed as ‘Kekulé’ bands). The gap between $\Phi_{\text{Kekulé}}^1$ and $\Phi_{\text{Kekulé}}^2$ requires the Kekulé modulation and thus is smaller than that between $\Phi_{\text{Rashba}}^1$ and $\Phi_{\text{Rashba}}^2$, which only depends on the coupling strength between graphene and TIs.

Within the gap between $\Phi_{\text{Kekulé}}^1$ and $\Phi_{\text{Kekulé}}^2$, there are other two bands, labeled by $\Phi_{\text{Dirac}}$ which are dominated by the states $\left\{ \left| -\frac{1}{2}, -1 \right\rangle \right\}$ and $\left\{ \left| -\frac{1}{2}, -1 \right\rangle \right\}$ (or correspondingly $\left\{ \left| J' = -\frac{1}{2}, \downarrow \right\rangle \right\}$ and $\left\{ \left| J' = -\frac{1}{2}, \downarrow \right\rangle \right\}$). For these two bands, the effective model only recovers the energy dispersion close to the $\Gamma$ point. For larger momentum, a highly nonlinear behavior can be observed from the first-principles calculations and suggests that higher momentum terms are dominant for these two bands. To get an effective description for these two bands, we apply the Löwdin perturbation theory to the full Hamiltonian $H_{\text{full}}$ (2) and project it into the low-energy subspace of $\Phi_{\text{Dirac}}$. Up to the third order in the momentum $k$, we obtain the following two-band effective Hamiltonian:

$$H_{\text{eff}} \left( k \right) = (A_{1}k_{x}^{2} + A_{2})\hat{\sigma} \cdot \left( \hat{\sigma} \times k \right) - \frac{A_{3}}{2}(k_{x}^{2} + k_{y}^{2})\sigma_{2} + (\epsilon_{0} - A_{0}k_{z}^{2})\hat{I}_{x2},$$

where $\epsilon_{0} = 0.01734$ eV, $A_{0} = 1.7 \times 10^{3}$ eV Å$^{-2}$, $A_{1} = 1.5 \times 10^{6}$ eV Å$^{2}$, $A_{2} = 2.7 \times 10^{6}$ eV Å$^{2}$, $A_{3} = 0.04079$ eV Å are obtained by fitting to the energy dispersion. $\hat{\sigma} = (\sigma_{x}, \sigma_{y}, \sigma_{z})$ are the Pauli matrices. For further discussions, we can ignore the identity terms and re-write the effective Hamiltonian in a nice form: $H_{\text{eff}} = \sum_{x,y,z} d_{x,y,z}(k)\sigma_{x}$. The parameter $A_{3}$ describes the linear dispersion and corresponds to the Fermi velocity around $6 \times 10^{3}$ m s$^{-1}$, which is greatly reduced compared to its original value $3 \times 10^{5}$ m s$^{-1}$ (the Fermi velocity of TI surface states). This estimate quantitatively reveals that the Fermi velocity of the bands $\Phi_{\text{Dirac}}$ is quite small and the dominating terms in equation (4) are the cubic terms.

The heavy Dirac fermion of $\Phi_{\text{Dirac}}$ can be physically understood from the form of their wave functions. One notices that the spin components of $\left\{ \left| \frac{1}{2}, +1 \right\rangle \right\}$ and $\left\{ \left| \frac{1}{2}, -1 \right\rangle \right\}$ are opposite, so the coupling between them has to involve a spin–flip process. Since SOC in graphene is negligible [19], this coupling can only originate from the hybridization with the TI surface states. Therefore, the interactions between $\left\{ \left| \frac{1}{2}, -1 \right\rangle \right\}$ and $\left\{ \left| \frac{1}{2}, +1 \right\rangle \right\}$ in graphene can only be mediated by the interlayer coupling between graphene and the topological surface states.
through second order (or higher order) perturbations. A detailed analysis of the possible coupling processes is shown in the appendix D. We conclude that the weak coupling between \( \frac{1}{2}, +1 \) and \( \frac{1}{2}, -1 \) makes the resulting Dirac fermion much heavier than that of the TI surface states.

### 3. Interaction effect in graphene/TI hetero-junctions

Based on the above low-energy effective Hamiltonian (4), we will next discuss interaction effect in this system. We consider the Hubbard repulsion interaction and write the whole Hamiltonian in the real space as

\[
H = \sum_{(r',r)} \Psi^\dagger (r') H_0 (r, r') \Psi (r') + U \sum_r \psi_1^\dagger (r) \psi_1 (r) \psi_2^\dagger (r) \psi_2 (r),
\]

(5)

where the spinor \( \Psi (r) = (\psi_1 (r), \psi_2 (r))^T \) is written under the basis \( \frac{1}{2}, +1 \) and \( \frac{1}{2}, -1 \). Based on the mean field approximation (details are shown in the appendix E), the above interaction Hamiltonian can be decomposed into the form

\[
H = \sum_k \Psi^\dagger (k) [H_{\text{eff}} (k) - (\Delta_y \sigma_y + \Delta_z \sigma_z)] \Psi (k) + \frac{L^2}{U} (\Delta_y^2 + \Delta_z^2),
\]

(6)

where \( L^2 \) is the surface area of the sample and the order parameters \( \Delta_y, \Delta_z \) are defined as follows:

\[
\Delta_y = \frac{U}{2} \langle \Psi^\dagger \sigma_y \Psi \rangle, \quad \Delta_z = \frac{U}{2} \langle \Psi^\dagger \sigma_z \Psi \rangle.
\]

(7)

Thus, we find two types of order parameters leading to the gap opening of heavy Dirac fermions in the heterostructure. These order parameters \( \Delta_y \) and \( \Delta_z \) both break time reversal symmetry, which leads to spontaneous half quantum Hall effect of the Dirac fermions. In addition, \( \Delta_z \) also breaks three-fold in-plane rotational symmetry and can be recognized as a type of nematic phases.

By minimizing the free energy, we obtain the self-consistent equations for order parameters \( \Delta_i (i = y, z) \)

\[
\frac{1}{U} \Delta_i = \frac{1}{2L^2} \sum_k \Delta_i - d_i (k) \sqrt{d_i (k)^2 + (d_y (k) - \Delta_y)^2 + (d_z (k) - \Delta_z)^2}.
\]

(8)

We emphasize that when the coefficient of linear term \( A_3 \) vanishes, DOS in the above equation (8) is divergent at the Dirac point, thus leading to an instability for heavy Dirac fermions. An infinitesimal interaction \( U \) will yield a gap and drive the system into the ordered phase. With a finite but small \( A_3 \), the DOS vanishes at the Dirac point, and thus, a finite interaction is needed to open a gap at the Dirac point. The phase boundary is characterized by the critical interaction strength \( U_c \), under which the order parameters take the limit \( \lim_{U \to U_c} \Delta_i = 0^+ \). The analytical expression of \( U_c \) is given by (see appendix F for a detailed derivation)

\[
\frac{1}{U_c} = \frac{1}{2L^2} \sum_k \frac{1}{\sqrt{d_i (k)^2 + d_y (k)^2 + d_z (k)^2}} - \frac{d_i (k)^2}{[d_i (k)^2 + d_y (k)^2 + d_z (k)^2]^2}.
\]

(9)

Here we arrived at two inequivalent self-consistent equations for \( U_c (i = y, z) \) and the physical phase boundary happens for \( U = \min \{U_c^y, U_c^z\} \).

The dependence of the critical interaction strength \( U_c \) as a function of \( A_3 \) is shown in figure 2. We find that \( U_c^y \) (plotted in a solid blue line) is always smaller than \( U_c^z \) (plotted in an orange dashed line). Therefore, the blue line in figure 2 determines the phase transition line in the real system. This numerical calculation also verifies our previous expectation that (i) when \( A_3 \) is zero, the required \( U_c \) is also zero, and (ii) as \( A_3 \) increases, \( U_c \) also increases from zero and a finite interaction strength is required to drive the system into the ordered phase. In figure 2, we also show the \( U_c \) for both an isolated TI surface state (\( A_3 \approx 1.8 \text{ eV } \text{Å} \)) and a graphene/TI junction (\( A_3 \approx 0.04 \text{ eV } \text{Å} \)) in red dots. One can see that the required critical interaction for ordered phases is almost one order of magnitude smaller in graphene/TI junctions, compared to that in pristine graphene.

### 4. Discussion and conclusion

In this work, we demonstrate the existence of a heavy Dirac fermion with a small Fermi velocity and highly nonlinear energy dispersion in the hetero-junction of graphene and a TI film, and identify the strong coupling between the graphene bands and the topological surface states of TI films as the underlying physical origins of heavy Dirac fermion, by combining ab initio calculations with the low energy effective models. Due to the significant
reduction of Fermi velocity, the low-energy states in graphene/TI junctions are more unstable in the presence of interactions. One may notice that the energy dispersion from first-principles calculations shows a larger nonlinearity compared to that for the effective model in figure 1. This suggests that an even larger DOS appears near the Dirac cone and the corresponding critical interaction strength would be further reduced. The decreased Fermi velocity and the enhanced DOS of the Dirac fermion can be verified in experiments of angle-resolved photoemission spectroscopy [20–22] and scanning tunneling spectroscopy [23, 24]. If the interaction effect in this system is strong enough, topological electromagnetic effect [25, 26] can be spontaneously realized in this heterojunction and one dimensional chiral fermions are expected to exist at the domain wall of this gaped phase. We notice that in bilayer or trilayer graphene, the energy dispersion becomes quadratic or cubic dispersing due to the vanishing of linear term. Interaction will also play an important role in multi-layer graphene and can lead to different interacting states [27, 28]. The main difference between our system and multi-layer graphene is that strong SOC exists in our system while weak SOC is found in multi-layer graphene. As a consequence, only a single heavy Dirac electron state appears in our graphene/TI junction system while even numbers of quadratic or cubic dispersing electron states exist in multi-layer graphene. In our case, both mass terms will lead to the quantum anomalous Hall states, while other types of Hall states, including quantum valley Hall state and quantum spin anomalous Hall states, while other types of Hall states, are possible in multi-layer graphene [27, 28].

Acknowledgments

We would like to thank Nitin Samarth and Jun Zhu for the helpful discussions. C-XL acknowledges the support from Office of Naval Research (Grant No. N00014-15-1-2675). WC and WD acknowledge the support from the Ministry of Science and Technology of China (Grant No. 2016YFA0301001) and the National Natural Science Foundation of China (Grant No. 11334006).

Appendix A. Details of ab initio calculations

The calculations were carried out by using density-functional theory (DFT) with a plane wave basis set and the frozen projector augmented wave method for the treatment of the core electrons [29, 30], as implemented in the Vienna ab initio simulation package [31]. A plane wave basis set with a kinetic energy cutoff of 450 eV was used. Before the band structure calculation, graphene and the four topmost atomic layers of the substrate are fully relaxed until the residual forces are less than $1 \times 10^{-3}$ eV Å$^{-1}$. During the relaxation process, the van der Waals correction of DFT-D2 method of Grimme is included [32]. With SOC included, the band structure was calculated along high symmetry lines near the $\Gamma$ point. In both geometry optimization and electronic calculation, the Monkhorst–Pack k points are 15 × 15 × 1.

Appendix B. Explicit form of the effective Hamiltonian $H$

With $p_n$ orbitals of the six carbon atoms being the basis states $| p_n, n = 1, \ldots, 6 \rangle$, the spinless tight-binding Hamiltonian of graphene is

$$H_{tb} = \begin{pmatrix}
0 & t_2 & 0 & t_1 e^{i(k_\perp a_x-\alpha)} & 0 & t_1 \\
0 & t_1 & 0 & t_1 e^{-i\alpha} & 0 & 0 \\
0 & t_2 & 0 & t_1 e^{-i\alpha} & 0 & t_1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
h.c. & 0 & t_2 & 0 & 0 & \end{pmatrix},$$

(B.1)

where the index $n$ for carbon atoms can be seen in figure 1. To reveal the low-energy Dirac physics, we transform the basis states to $| L_\perp^\pm = \pm 2, \pm 1, 0, 3 \rangle$, where $C_6 | L_\perp^\pm \rangle = \omega^{-2L_\perp} | L_\perp^\pm \rangle$, $C_6$ is the six-fold rotation operator and $\omega = \exp\left(\frac{2\pi i}{6}\right)$. The basis transformation matrix $U_{6 \times 6}$ is

$$U_{6 \times 6} = \frac{1}{\sqrt{6}} \begin{pmatrix}
-2 & -1 & 0 & 1 & 2 & 3 \\
1 & \omega^{-3} & \omega^3 & 1 & 1 \\
\omega^{-4} & \omega^{-5} & \omega^5 & \omega^{-4} & 1 \\
1 & \omega^3 & \omega^{-3} & 1 & 1 \\
\omega^{-4} & \omega & \omega^{-1} & \omega^4 & 1 \\
\omega^4 & \omega^{-1} & 1 & \omega^4 & \omega^{-4} & \end{pmatrix},$$

(B.2)

After the basis transformation, we can expand the spinless Hamiltonian around the $\Gamma$ point

$$\begin{pmatrix}
| -2 \rangle & | -1 \rangle & | 0 \rangle & | 1 \rangle & | 2 \rangle & | 3 \rangle \\
-\Delta \cos \theta & h_{\kappa_+} & 0 & 0 & \Delta \sin \theta & 0 \\
h_{\kappa_-} & + \Delta \cos \theta & 0 & \Delta \sin \theta & 0 & 0 \\
0 & 0 & t_1 + t_2 + t_3 & 0 & 0 & 0 \\
0 & \Delta \sin \theta & 0 & - \Delta \cos \theta & h_{\kappa_-} & 0 \\
\Delta \sin \theta & 0 & 0 & - h_{\kappa_+} & + \Delta \cos \theta & 0 \\
0 & 0 & 0 & 0 & 0 & t_1 - t_2 - t_3 \\
\end{pmatrix},$$

(B.3)
The physical meaning of those parameters have been explained in the main context.

With the spin degree of freedom, the Hamiltonian of graphene becomes

\[
H_{th} = \begin{pmatrix}
H_0 & 0 \\
0 & H_0
\end{pmatrix}
\] (B.4)

and the basis states expands to \{ |p_2, 1, \uparrow\rangle, |p_2, 2, \uparrow\rangle, |p_2, 3, \uparrow\rangle, |p_2, 4, \uparrow\rangle, |p_2, 5, \uparrow\rangle, |p_2, 1, \downarrow\rangle, |p_2, 2, \downarrow\rangle, |p_2, 3, \downarrow\rangle, |p_2, 4, \downarrow\rangle, |p_2, 5, \downarrow\rangle, |p_2, 6, \downarrow\rangle \}.

Next, a unitary transformation \( U \) is applied to the basis states and transforms the Hamiltonian as

\[
H_{th} \rightarrow U^+ H_{th} U.
\] (B.5)

The 12 \times 12 transformation matrix \( U \) is chosen to be

\[
U = \frac{1}{\sqrt{6}} \begin{pmatrix}
1 & \omega^{-3} & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & \omega^3 \\
\omega^{-1} & \omega^{-3} & 0 & 0 & 1 & 0 & 0 & 0 & -1 & \omega^4 & 0 & \omega^5 \\
\omega^4 & \omega^5 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & \omega^{-4} & 0 & \omega^{-5} \\
\omega^4 & \omega^{-1} & 0 & 0 & 1 & 0 & 0 & 0 & -1 & \omega^{-4} & 0 & \omega \\
\omega^{-4} & \omega & 0 & 0 & 1 & 0 & 0 & 0 & 1 & \omega^4 & 0 & \omega^{-1} \\
1 & \omega^3 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & \omega^{-4} & 0 & \omega \\
\omega^{-4} & \omega & 0 & 0 & 1 & 0 & 0 & 0 & 1 & \omega^4 & 0 & \omega^{-1} \\
\omega & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & \omega^{-4} & 0 & \omega \\
0 & 0 & 1 & 1 & 0 & \omega^{-3} & 1 & \omega^3 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & \omega^{-4} & 0 & \omega^{-3} & \omega^3 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 1 & \omega^4 & 0 & \omega^3 & \omega^{-4} & \omega^{-5} & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & \omega^3 & 1 & \omega^{-3} & 0 & 0 & -1 & 0 \\
0 & 0 & 1 & \omega^{-4} & 0 & \omega & \omega^4 & \omega^{-1} & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & \omega^4 & 0 & \omega^{-1} & \omega^{-4} & \omega & 0 & 0 & -1 & 0
\end{pmatrix}
\] (B.6)

Expanding the transformed Hamiltonian around \( k = 0 \) and excluding the energy levels far away from \( \mu \) gives

\[
H_{GG} = \mu \tilde{c}
\]

\[
\begin{pmatrix}
\frac{3}{2}, +1 & -\frac{1}{2}, +1 & \frac{1}{2}, +1 & -\frac{3}{2}, +1 & -\frac{5}{2}, +1 & \frac{1}{2}, -1 & -\frac{1}{2}, -1 & \frac{3}{2}, -1
\end{pmatrix}
\]

\[
\begin{pmatrix}
-\Delta \cos \theta & \hbar v_f k_+ & 0 & 0 & 0 & 0 & 0 & \Delta \sin \theta \\
\hbar v_f k_- & +\Delta \cos \theta & 0 & 0 & 0 & 0 & \Delta \sin \theta & 0 \\
0 & 0 & -\Delta \cos \theta & \hbar v_f k_+ & 0 & \Delta \sin \theta & 0 & 0 \\
0 & 0 & \hbar v_f k_- & +\Delta \cos \theta & \Delta \sin \theta & 0 & 0 & 0 \\
0 & 0 & \Delta \sin \theta & -\Delta \cos \theta & -\hbar v_f k_- & 0 & 0 & 0 \\
0 & 0 & \Delta \sin \theta & 0 & -\hbar v_f k_+ & +\Delta \cos \theta & 0 & 0 \\
\Delta \sin \theta & 0 & 0 & 0 & 0 & -\Delta \cos \theta & -\hbar v_f k_- & +\Delta \cos \theta
\end{pmatrix}
\] (B.7)
and
\[
\begin{vmatrix}
| p_z, 1 \\ 
| p_z, 2 \\ 
| p_z, 3 \\ 
| p_z, 4 \\ 
| p_z, 5 \\ 
| p_z, 6 \\
\end{vmatrix}
\times
\begin{pmatrix}
1 & \omega^{-4} & \omega^{-5} & \omega^{-3} & \omega^{-1} & \omega \\
\omega & \omega^{-4} & \omega^{-5} & \omega^{-3} & \omega^{-1} & \omega \\
\omega & \omega^{-4} & \omega^{-5} & \omega^{-3} & \omega^{-1} & \omega \\
\omega & \omega^{-4} & \omega^{-5} & \omega^{-3} & \omega^{-1} & \omega \\
\omega & \omega^{-4} & \omega^{-5} & \omega^{-3} & \omega^{-1} & \omega \\
\omega & \omega^{-4} & \omega^{-5} & \omega^{-3} & \omega^{-1} & \omega \\
\end{pmatrix}
\]

of \( k \):
\[-\left\{ -\frac{3}{2}, -1 \right\} \frac{1}{2} = \left\{ 3 \right\} \frac{1}{2} - 1 = 0 \]
\[-\left\{ \frac{3}{2}, -1 \right\} \frac{1}{2} = \left\{ 3 \right\} \frac{1}{2} - 1 = 0 \]
\[-\left\{ \frac{1}{2}, -1 \right\} \frac{1}{2} = \left\{ \frac{1}{2} \right\} \frac{1}{2} + 1 = 0 \]
\[-\left\{ \frac{3}{2}, -1 \right\} \frac{1}{2} = \left\{ 3 \right\} \frac{1}{2} - 1 = 0 \]
\[-\left\{ \frac{1}{2}, -1 \right\} \frac{1}{2} = \left\{ \frac{1}{2} \right\} \frac{1}{2} + 1 = 0 \]
\[-\left\{ \frac{3}{2}, -1 \right\} \frac{1}{2} = \left\{ 3 \right\} \frac{1}{2} - 1 = 0 \]
\[-\left\{ \frac{1}{2}, -1 \right\} \frac{1}{2} = \left\{ \frac{1}{2} \right\} \frac{1}{2} + 1 = 0 \]
\[-\left\{ \frac{3}{2}, -1 \right\} \frac{1}{2} = \left\{ 3 \right\} \frac{1}{2} - 1 = 0 \]
\[-\left\{ \frac{1}{2}, -1 \right\} \frac{1}{2} = \left\{ \frac{1}{2} \right\} \frac{1}{2} + 1 = 0 \]

\( J_z, L_z \) are the total angular momentum and orbital angular momentum of \( \mathcal{C}_3 \) rotation in z-direction respectively. \( \mathcal{C}_3 \mid L_z, \eta \rangle = \omega^{-L_z} \mid L_z, \eta \rangle \). The original \( K_{\sigma} K'_{\sigma} \) being folded onto the \( \Gamma ' \) point gives rise to the additional index \( \eta \). The details about the definition of the index \( \eta \) are in appendix C. The four states \( \mid L_z = 0, \eta = \pm 1, \uparrow ( \downarrow ) \rangle \) are excluded because their energy levels are much further away from the Fermi level.

The surface states are described by \( H_{SS} \):
\[
\begin{pmatrix}
\mu^+ + i \hbar \gamma k \\
- i \hbar \gamma k \\
\mu^+ + i \hbar \gamma k \\
\end{pmatrix}
\]
with the basis states:
\[
| \pm \frac{1}{2} \rangle = \sum_{\alpha = \text{Sb, Te}} u_{\alpha}(\alpha, p_z, \uparrow ( \downarrow )) + v_{\alpha}(\alpha, p_z, \downarrow ( \uparrow ))
\]
where \( | p_z \rangle = \frac{1}{\sqrt{2}} (| p_z \rangle \pm i | p_z \rangle ) \) and \( u_{\alpha}, v_{\alpha} \) are assumed to be real.

The hybridization between \( p \) orbitals can be decoupled to two types \( (V_{pp}, V_{pp}) \). In this case, the hybridization between \( p_z \) orbital of carbon atoms and \( p_z/p_x \) orbital from \( \alpha \), \( \alpha = \text{Te, Sb} \) is proportional to \( V_{pp} \) and we have the following decompositions:
\[
V_{pp}^\alpha = V_{pp} \sin^2 \theta^\alpha + V_{pp}^\alpha \cos^2 \theta^\alpha
\]
\[
V_{pp}^\alpha = \pm V_{pp} e^{i \varphi^\alpha}
\]
where \( \theta^\alpha \) stands for the angle between the \( z \)-axis and \( L^\alpha \), \( \varphi^\alpha \) denotes the angle between the positive \( x \)-axis and the projection of \( L^\alpha \) on the \( x-y \)-plane. Here \( L^\alpha \) denotes the vector connecting carbon atom and \( \alpha \) atom and pointing to the carbon atom. For the hybridization submatrix \( H_{SS} \), the sixteen terms have been deduced from real space tight-binding method (up to the order...
and
\[
\left\{ \frac{3}{2}, +1 | H | - \frac{1}{2} \right\} = \left\langle L_z = 1, \eta = 1, \uparrow | H | - \frac{1}{2} \right\rangle = \sum_{n} a_{n} (p_{z}, n, \uparrow )
\]

\[
eq +1, \uparrow | H | (\alpha, p_{z}, \uparrow )
\]

\[
= \sum_{n} c_{n} (p_{z}, n, \uparrow )
\]

\[
= v_{Te} V_{Te} (1 \cdot \omega^{6} + \omega^{4} \cdot \omega^{4} + \omega^{4} \cdot \omega^{2} + 1 + \omega^{4} \cdot \omega^{-2} + \omega^{4} \cdot \omega^{-4} / \sqrt{6} + \omega^{4} \cdot \omega^{2} + \omega^{-4} \cdot \omega^{4} ) / \sqrt{6} = 0,
\]

(B.15)

where the hybridization is limited between states with the same spin and the transformation \( \left\{ \frac{3}{2}, +1 \right\} = \sum_{n} a_{n} (p_{z}, n, \uparrow ) \) in equations (B.8) and (B.9) is used. The complete Hamiltonian \( H_{SS} \) for the coupling between graphene and topological surface states of Sb_{2}Te_{3} films is written as

\[
\begin{pmatrix}
\frac{\sqrt{2}}{4} V_{s} V_{b_{s}} a_{1} \mid k_{-} \ 
\frac{\sqrt{2}}{4} V_{b_{s}} a_{1} \mid k_{+} \ 
\frac{\sqrt{2}}{4} V_{s} V_{b_{s}} a_{1} \mid k_{-} \ 
\frac{\sqrt{2}}{4} V_{b_{s}} a_{1} \mid k_{+} \ 
0 \ 
\end{pmatrix} =
\begin{pmatrix}
0 \ 
- \sqrt{6} i V_{Te} V_{b_{s}} a_{1} \mid k_{-} \ 
\frac{\sqrt{2}}{4} V_{s} V_{b_{s}} a_{1} \mid k_{-} \ 
\frac{\sqrt{2}}{4} V_{b_{s}} a_{1} \mid k_{+} \ 
- \sqrt{6} i V_{Te} V_{b_{s}} a_{1} \mid k_{+} \ 
\end{pmatrix}.
\]

(B.16)

Now all the terms in \( H_{G0}, H_{SS} \) and \( H_{GS} \) have been explicitly shown and the parameters in them can be determined by fitting the the ab initio band structure, as listed in the table B.1.

**Appendix C. Definition of the index \( \eta \)**

For a pristine graphene with atom 1 and 2 in the primitive cell in figure 1(b), the eigenstates of the Hamiltonian at the Dirac point \( | K \rangle \) and \( | K' \rangle \) should take the form of

\[
\begin{pmatrix}
\frac{\sqrt{2}}{4} V_{s} V_{b_{s}} a_{1} \mid k_{-} \ 
\frac{\sqrt{2}}{4} V_{b_{s}} a_{1} \mid k_{+} \ 
\frac{\sqrt{2}}{4} V_{s} V_{b_{s}} a_{1} \mid k_{-} \ 
\frac{\sqrt{2}}{4} V_{b_{s}} a_{1} \mid k_{+} \ 
0 \ 
\end{pmatrix} =
\begin{pmatrix}
0 \ 
- \sqrt{6} i V_{Te} V_{b_{s}} a_{1} \mid k_{-} \ 
\frac{\sqrt{2}}{4} V_{s} V_{b_{s}} a_{1} \mid k_{-} \ 
\frac{\sqrt{2}}{4} V_{b_{s}} a_{1} \mid k_{+} \ 
- \sqrt{6} i V_{Te} V_{b_{s}} a_{1} \mid k_{+} \ 
\end{pmatrix}.
\]

(B.16)

projections of \( | K \rangle \) on the sublattice containing atom 1 and \( | K' \rangle \) on the sublattice containing atom 2. To classify the doubly degenerate states at \( K, K' \) further, we introduce \( C_{6} \) the six-fold rotation operator. And we find

\[
C_{6} (| K, 1, 0 \rangle + \omega^{-4} | K', 1, 0 \rangle ) = \omega^{4} (| K, 1, 0 \rangle + \omega^{-4} | K', 1, 0 \rangle )
\]

\[
= \omega^{4} C_{6} L_{z} = -2
\]

\[
C_{6} (\omega^{4} | K, 1, 0 \rangle + \omega^{4} | K', 1, 0 \rangle ) = \omega^{-2} (\omega^{4} | K, 1, 0 \rangle + \omega^{4} | K', 1, 0 \rangle )
\]

\[
= \omega^{-2} C_{6} L_{z} = 1
\]

(C.4)
In the graphene/Sb$_2$Te$_3$ hetero-junction, the C$_3$ rotation symmetry is preserved but C$_6$ rotation symmetry is broken. So to take advantage of the classifications above, we introduce an additional index $\eta$ and set

$$|L_2 = 1, \eta = +1\rangle = |C_{\eta \eta}, L'_2 = -2\rangle, \quad |L_2 = 1, \eta = -1\rangle = |C_{\eta \eta}, L'_2 = 1\rangle. \quad (C.5)$$

The situation of the states $|K (K'), 0, 1\rangle$ is similar.

$$C_6 (\omega^{-4}|K, 0, 1\rangle + \omega^4|K', 0, 1\rangle) = \omega^2 (C_{\eta \eta}, L'_2 = -1)$$

$$C_6 (\omega^{-4}|K, 0, 1\rangle + \omega^4|K', 0, 1\rangle) = \omega^{-4} (C_{\eta \eta}, L'_2 = 2)$$

$$|L_2 = -1, \eta = +1\rangle = |C_{\eta \eta}, L'_2 = -1\rangle, \quad |L_2 = -1, \eta = -1\rangle = |C_{\eta \eta}, L'_2 = 2\rangle. \quad (C.6)$$

It is easy to check that the definitions of $|L_2, \eta\rangle$ here in equations (C.5) and (C.6) are consistent with those in equation (B.9). To sum up, the electron state $|L_2, \eta\rangle$ is an eigenstate of $C_6$ rotation operator and also a linear supercomposition of $|K\rangle$ projected on one sublattice and $|K'\rangle$ projected on the other sublattice. According to the definitions above, the time-reversal operation changes the sign of both $L_2$ and $\eta$.

**Appendix D. Possible coupling processes between $|\frac{-1}{2}, -1\rangle$ and $|\frac{1}{2}, +1\rangle$**

The spin–flip coupling process should include the coupling between these graphene states with the topological surface states $\pm |\frac{1}{2}\rangle$, which can be decomposed into $p$ orbitals of Sb and Te atoms as in equation (B.1). Bearing the decompositions in mind, we have the following perturbative terms that produce linear momentum dependence (the energy difference in the denominator is assumed to be constant and ignored)

$$\left\langle \frac{1}{2}, 1 | H | \alpha, p_s, 1 \right\rangle \left\langle \alpha, p_s, 1 | H | \frac{1}{2}, -1 \right\rangle \propto k_-$$

$$\left\langle \frac{1}{2}, -1 | H | \alpha, p_s, 1 \right\rangle \left\langle \alpha, p_s, 1 | H | \frac{1}{2}, -1 \right\rangle \propto k_-$$

$$\left\langle \frac{1}{2}, 1 | H | \alpha, p_s, 1 \right\rangle \left\langle \alpha, p_s, 1 | H | \frac{1}{2}, -1 \right\rangle \propto k_-$$

$$\left\langle \frac{1}{2}, 1 | H | \alpha, p_s, 1 \right\rangle \left\langle \alpha, p_s, 1 | H | \frac{1}{2}, -1 \right\rangle \propto k_-$$

$$\left\langle \frac{1}{2}, 1 | H | \alpha, p_s, 1 \right\rangle \left\langle \alpha, p_s, 1 | H | \frac{1}{2}, -1 \right\rangle \propto k_-$$

$$\left\langle \frac{1}{2}, 1 | H | \alpha, p_s, 1 \right\rangle \left\langle \alpha, p_s, 1 | H | \frac{1}{2}, -1 \right\rangle \propto k_-$$

$$\left\langle \frac{1}{2}, 1 | H | \alpha, p_s, 1 \right\rangle \left\langle \alpha, p_s, 1 | H | \frac{1}{2}, -1 \right\rangle \propto k_-$$

with $\alpha = \text{Te, Sb}$ for TIs. Here the first two terms are from $\left\langle \frac{1}{2}, 1 | H | \frac{1}{2}, 1 \right\rangle = \frac{1}{2} \left\langle \frac{1}{2}, 1 | H | \frac{1}{2}, 1 \right\rangle$ and the last two are from $\left\langle \frac{1}{2}, 1 | H | \frac{1}{2}, 1 \right\rangle = \frac{1}{2} \left\langle \frac{1}{2}, 1 | H | \frac{1}{2}, 1 \right\rangle$.

\[ \left\{ \tilde{H}_S = \frac{1}{2} | H | \frac{1}{2}, 1 \right\} \]

The underlined terms involve the spin–flip hopping between graphene and TI. Therefore, only second order (or higher order) perturbation processes can give rise to the linear momentum dependence and that is why the Fermi velocity of the resulting state is much smaller than that of the TI surface states. Based on the above perturbation theory, it is easy to estimate the effective Fermi velocity $A_3$:

$$A_3 \sim \frac{V_{\text{eff}}^{\text{Sb}} \times V_{\text{eff}}^{\text{Sb}}}{\Delta E} \sim \mathcal{O}(10^{-2} \text{ eV} \ \text{Å}) \quad (D.2)$$

where $\Delta E$ is the energy difference between the graphene states and TI surface states at $\Gamma$ point.

**Appendix E. Mean field theory for ‘heavy’ Dirac fermions with interaction**

Based on the third order Löwdin perturbation theory [33], we obtain the effective Hamiltonian near the $\Gamma$ point

$$H_{D3} = \left( \tilde{A}_k k_x k_y - A_k k_x + iA_k k_y \right)$$

$$+ (e_0 - A_k k_x) L_x \zeta$$

$$= (A k_x + A_3) \zeta \cdot (\sigma \times k) - \frac{A_3}{2} (k_x^2 + k_y^3) \zeta_2$$

$$+ (e_0 - A_k k_x) L_x \zeta$$

$$\left( E_k = \pm \left[ (A k_x^2 + A_3) k_y + m_x \right] \right. \right.$$}

$$\left. + (m_y - (A_k k_x + A_3) k_y) \zeta_2 \right. \right. \right.$$}

$$\left. + (A_k k_x^2 - A_k k_x + m_x)^2 \zeta_2 \right. \right.$$

where the identity term $A_k k_x^2 L_x \zeta_2$ is ignored. When $m_x = m_y = 0$, $m_x \sigma_2$ is obviously a mass term. When
\( m_y = m_x = 0 \), being gapless requires \( k_x = 0 \) and the energy dispersion is

\[
E^x_d = 2(A_2 k_y^2 + A_1)k_y + m_x. \tag{E.3}
\]

So we could always find a possible \( k_y \), where \( E^x_d = 0 \) when \( k_x = 0 \). When \( m_x = m_y = 0 \), being gapless requires \( k_x = 0 \) and the corresponding eigen energy becomes

\[
E^y_d = 2\sqrt{(m_y - (A_2 k_y^2 + A_1)k_y)^2 + (A_1 k_y^4)} \tag{E.4}
\]

In this case, we cannot find a gapless point in the band structure. Thus, the general form for the mass terms should be \( M = m_y \sigma_y + m_z \sigma_z \).

We consider an onsite Hubbard repulsion

\[
H = \sum_{\langle r, r' \rangle} \Psi^\dagger(r')H_0(r, r')\Psi(r') + U\sum_{\langle r, r' \rangle} \psi^\dagger_1(r)\psi_1(r)\psi^\dagger_2(r)\psi_2(r). \tag{E.5}
\]

Here we have defined real space basis \( \Psi^\dagger(r) = (\psi^\dagger_1(r), \psi^\dagger_2(r)) \), and a real space Hamiltonian \( H(r, r') = H(r - r') \) related to our effective Hamiltonian via Fourier transformation

\[
H_0 = \sum_{\langle r, r' \rangle} c_{r'}^\dagger H(r, r')c_r = \sum_k c_{k}^\dagger c_k H_{\text{eff}}(k). \tag{E.6}
\]

Based on previous mass terms analysis, we expect interaction-induced mass terms to be proportional to \( \sigma_y \) and \( \sigma_z \). Such mass terms will naturally arise if we consider the following decomposition scheme to the four-fermion interaction:

\[
\psi_1^\dagger \psi_1^\dagger \psi_2 \psi_2 = (\psi_1^\dagger \psi_1)(\psi_2^\dagger \psi_2) - (\psi_1^\dagger \psi_2)(\psi_2^\dagger \psi_1) = \{, - \}
\]

We define the following order parameters:

\[
\delta_0 = \langle \psi_1^\dagger \psi_2 \rangle = \delta e^{i\theta}
\]

\[
\delta^*_0 = \langle \psi_2^\dagger \psi_1 \rangle = \delta e^{-i\theta}
\]

\[
\delta_1 = \langle \psi_1^\dagger \psi_1 \rangle
\]

\[
\delta_2 = \langle \psi_2^\dagger \psi_2 \rangle
\]

and we could perform the mean field theory based on the above order parameters

\[
\{ = (\psi_1^\dagger \psi_1 - \delta_1 + \delta_0)(\psi_2^\dagger \psi_2 - \delta_2 + \delta_0) = -\delta_1\delta_2 + \psi^\dagger(\begin{bmatrix} \delta_2 & 0 \\ 0 & \delta_1 \end{bmatrix})\psi
\]

\[
= \frac{m_z^2 - m_x^2}{4} + \psi^\dagger\begin{bmatrix} \frac{m_z - m_x}{2} & 0 \\ 0 & \frac{m_z + m_x}{2} \end{bmatrix}\psi
\]

\[
= \left(\frac{m_z^2}{4} - \frac{m_z^2}\Psi^\dagger_1\Psi\right) - \frac{m_0^2}{4} - \frac{m_0^2}\Psi^\dagger_0\Psi. \tag{E.9}
\]

Here we have introduced

\[
m_z = \delta_1 - \delta_2 = \langle \Psi^\dagger_1\sigma_z\Psi \rangle
\]

\[
m_0 = \delta_1 + \delta_2 = \langle \Psi^\dagger_0\sigma_0\Psi \rangle. \tag{E.10}
\]

Similarly

\[
\{ = (\psi_1^\dagger \psi_2 - \delta_0 + \delta_0)(\psi_2^\dagger \psi_1 - \delta_0 + \delta_0) = -\delta_0^2 + \psi^\dagger(\begin{bmatrix} 0 & \delta_0^* \\ \delta_0 & 0 \end{bmatrix})\psi
\]

\[
= -\delta_0^2 + \psi^\dagger\begin{bmatrix} \delta_0 \cos \theta \sigma_x + \delta_0 \sin \theta \sigma_y \end{bmatrix}\psi
\]

\[
= -\frac{m_x^2 + m_y^2}{4} + \psi^\dagger\begin{bmatrix} m_x \sigma_x + m_y \sigma_y \end{bmatrix}\psi. \tag{E.11}
\]

where we have defined

\[
m_x = \delta_0 + \delta^*_0 = \langle \psi_1^\dagger \psi_2 + \psi_2^\dagger \psi_1 \rangle = \langle \Psi^\dagger_1\sigma_0 \Psi \rangle
\]

\[
m_y = \delta_0 - \delta^*_0 = \langle -i\psi_1^\dagger \psi_2 + i\psi_2^\dagger \psi_1 \rangle = \langle \Psi^\dagger_0\sigma_1 \Psi \rangle. \tag{E.12}
\]

In the previous section, we have shown that only mass terms that couples to either \( \sigma_y \) or \( \sigma_x \) will gap the system. So we will consider the contributions from order \( m_z \) and \( m_y \). In the mean field approximation, we write down the Hamiltonian and Fourier transform it into momentum space

\[
H = \sum_k \Psi^\dagger(k)H_{\text{eff}}(k)\Psi(k) + U\sum_r \{, - \}
\]

\[
= \sum_k \Psi^\dagger(k)H_{\text{eff}}(k)\Psi(k)
\]

\[
- \frac{U}{2}\sum_r \Psi^\dagger(r)(m_y \sigma_y + m_z \sigma_z)\Psi(r)
\]

\[
+ \frac{L^2U}{4}(m_y^2 + m_z^2)
\]

\[
= \sum_k \Psi^\dagger(k)H_{\text{eff}}(k)
\]

\[
- \frac{U}{2}(m_y \sigma_y + m_z \sigma_z)\Psi(k) + \frac{L^2U}{4}(m_y^2 + m_z^2)
\]

\[
= \sum_k \Psi^\dagger(k)H_{\text{eff}}(k)
\]

\[
- \frac{U}{2}(m_y \sigma_y + m_z \sigma_z)\Psi(k) + \frac{L^2U}{4}(m_y^2 + m_z^2). \tag{E.13}
\]

The above mass terms obviously both break time reversal symmetry, and thus gap the surface states. A difference between the two mass terms is that \( m_y \sigma_x \) breaks three fold rotation symmetry \( C_3 \), while \( m_z \sigma_z \) preserves \( C_3 \) symmetry. Therefore, \( m_y \sigma_x \) leads to a nematic phase. If the system is in this nematic regime, such nematic ordering could be detected via scanning tunneling microscopy (STM) measurement. In general, both terms here will gap the surface states and result in the spontaneous half quantum hall state on the surface.
Appendix F. Critical interaction solution for self-consistent equations

The zero temperature free energy $F$ is

$$F = \frac{L^2}{4}(m_y^2 + m_z^2) - \sum \sqrt{d_x(k)^2 + (d_y(k) - \Delta_y)^2 + (d_z(k) - \Delta_z)^2}$$

$$= \frac{L^2}{U}(\Delta_y^2 + \Delta_z^2) - \sum \sqrt{d_x(k)^2 + (d_y(k) - \Delta_y)^2 + (d_z(k) - \Delta_z)^2}, \quad (F.1)$$

where $\Delta_{y,z} = \frac{U}{2}m_{y,z}$. Self-consistent equations for order parameter $\Delta_i (i = y, z)$ can be obtained by minimizing $F$:

$$0 = \frac{\partial F}{\partial \Delta_i} = \frac{2L^2}{U} \Delta_i - \sum \frac{d_i(k)}{\sqrt{d_x(k)^2 + (d_y(k) - \Delta_y)^2 + (d_z(k) - \Delta_z)^2}}.$$

$$= \lim_{\Delta_{y,z} \to 0} \frac{\sum d_i(k)}{\Delta_i} \lim_{\Delta_{y,z} \to 0} \sum \sqrt{d_x(k)^2 + (d_y(k) - \Delta_y)^2 + (d_z(k) - \Delta_z)^2}$$

$$= \lim_{\Delta_{y,z} \to 0} \sum \frac{d_i(k)}{\Delta_i} \lim_{\Delta_{y,z} \to 0} \sum \sqrt{d_x(k)^2 + (d_y(k) - \Delta_y)^2 + (d_z(k) - \Delta_z)^2}.$$

$$= \lim_{\Delta_{y,z} \to 0} \sum \frac{d_i(k)}{\Delta_i} \lim_{\Delta_{y,z} \to 0} \sum \sqrt{d_x(k)^2 + (d_y(k) - \Delta_y)^2 + (d_z(k) - \Delta_z)^2}.$$

In two-dimensions, the DOS for a linear dispersion at the Dirac point is vanishing, but DOS for a cubic dispersion is diverging. Thus, in the presence of interactions, systems with a linear dispersion require finite interaction strength to develop ordering, while systems with a cubic dispersion will develop instabilities. In our system, there is an interesting competition between linear terms and cubic terms. When linear term is vanishingly small, we expect the system will spontaneously develop TR breaking ordering ($m_y$ and $m_z$). As we increase the linear term, the critical interaction required to develop the above order parameters also increases. To verify our expectation, we aim at finding a relation between critical interaction $U_c$ and linear coefficient $A_y$, while leaving the cubic coefficient $A_z$ fixed. Recall that the critical interaction happens when the limit $\Delta_i \to 0$ takes place.

$$\frac{1}{U_c} = \frac{1}{2L^2} \lim_{\Delta_{y,z} \to 0} \sum \frac{d_i(k)}{\sqrt{d_x(k)^2 + (d_y(k) - \Delta_y)^2 + (d_z(k) - \Delta_z)^2}}$$

$$= \frac{1}{2L^2} \lim_{\Delta_{y,z} \to 0} \sum \frac{d_i(k)}{\Delta_i} \lim_{\Delta_{y,z} \to 0} \sum \sqrt{d_x(k)^2 + (d_y(k) - \Delta_y)^2 + (d_z(k) - \Delta_z)^2}.$$

In the above equation, the first limit is easy to evaluate. For the second limit, the denominator $\Delta_i \to 0$, and meanwhile, the numerator also approaches zero under this limit, because $\sqrt{d_x(k)^2 + (d_y(k) - \Delta_y)^2 + (d_z(k) - \Delta_z)^2}$ is an odd function in $k$. Therefore, this limit is a $\frac{0}{0}$ type limit, which can be evaluated with the help of the L’Hospital’s rule:

References

[1] Zhang Y, Tan Y-W, Stormer H L and Kim P 2005 Nature 438 201
[2] Novoselov K, Geim A K, Morozov S, Jiang D, Katsnelson M, Grigorieva I, Dubonos S and Firsov A 2005 Nature 438 197
[3] Neto A C, Guinea F, Peres N, Novoselov K S and Geim A K 2009 Rev. Mod. Phys. 81 109
[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] Moore J E 2010 Nature 464 194
[7] Wallace P R 1947 Phys. Rev. 71 622
[8] Avouris P 2010 Nano Lett. 10 4285
[9] Das Sarma S, Adam S, Hwang E H and Rossi E 2011 Rev. Mod. Phys. 83 407
[10] Novoselov K S et al 2012 Nature 490 192
[11] Jackiw R 1984 Phys. Rev. D 29 2375
[12] Haldane F D M 1988 Phys. Rev. Lett. 61 2015
[13] Kane C L and Mele E J 2005 Phys. Rev. Lett. 95 226801
[14] Raghu S, Qi X-L, Honerkamp C and Zhang S-C 2008 Phys. Rev. Lett. 100 156401
[15] Kou J, Yan B, Hu F, Wu S-C, Wehling T O, Felser C, Chen C and Frauenheim T 2013 Nano Lett. 13 6251
[16] Jin K H and Jhi S-H 2013 Phys. Rev. B 87 075442
[17] Wyckoff R W G and Wyckoff R W 1960 Crystal Structures vol 2 (New York: Wiley)
[18] Hou C-Y, Chamon C and Mudry C 2007 Phys. Rev. Lett. 98 186809
[19] Yao Y, Ye F, Qi X-L, Zhang S-C and Fang Z 2007 Phys. Rev. B 75 041401
[20] Zhou S, Gweon G-H, Graf J, Fedorov A, Spataru C, Diehl R, Kopelevich Y, Lee D-H, Louie S G and Lanzara A 2006 Nat. Phys. 2 595
[21] Chen Y et al 2009 Science 325 178
[22] Xia Y et al 2009 Nat. Phys. 5 398
[23] Roushan P, Seo J, Parker C V, Hor Y, Hsieh D, Qian D, Richardella A, Hasan M Z, Cava R and Yazdani A 2009 Nature 460 1106
[24] Zhang T et al 2009 Phys. Rev. Lett. 103 266803
[25] Qi X-L, Hughes T L and Zhang S-C 2008 Phys. Rev. B 78 195424
[26] Esin A M, Moore J E and Vanderbilt D 2009 Phys. Rev. Lett. 102 146805
[27] Zhang F, Jung J, Fiete G A, Niu Q and MacDonald A H 2011 Phys. Rev. Lett. 106 156801
[28] Zhang F 2015 Synth. Met. 210 9
[29] Blochl P E 1994 Phys. Rev. B 50 17953
[30] Kresse G and Joubert D 1999 Phys. Rev. B 59 1758
[31] Kresse G and Furthmüller J 1996 Phys. Rev. B 54 11169
[32] Grimme S 2006 J. Comput. Chem. 27 1787
[33] Winkler R 2003 Spin–orbit Coupling Effects in Two-dimensional Electron and Hole Systems (Berlin: Springer)