Effects of Interactions and Magnetic Fields on Topological Surface States

by

Drew Shepherd

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Physics

University of Alberta

Abstract

We study the effects of electron-electron interactions and magnetic fields on surface states of three-dimensional topological insulators (3D TIs). In this work we use an effective Hamiltonian to describe a slab of the 3D TI Bi_2Se_3. In the non-interacting limit, and for slab thicknesses greater than approximately 40 Å, we observe a Dirac cone in the bulk bandgap whose states are highly localized to the surface. Similar behaviour is seen in the spectral function: near the surface it has peaks near the Dirac cone states, and in the bulk, the Dirac cone disappears. In addition, the Dirac cone gap closes at approximately 60 Å, which agrees well with experimental measurements, and the density of states near the Dirac point is found to be linear. Next, we incorporate short-range electron-electron interactions through the calculation of the second order self-energy, and its anti-Hermitian part, the broadening function. Examining the broadening function allows us to study the qualitative behaviour of the quasiparticle lifetime near the Fermi level. We obtain an infinite lifetime at the Fermi level, and a finite lifetime as we move away from this energy. Returning to the non-interacting regime, but this time adding a finite magnetic field, we observe Landau level (LL) peaks in the density of states. As the thickness of the slabs increases the total number of states increases, but the number of LLs localized to the surface remains constant. The observation of a zeroth LL and the LLs being linear in $\sqrt{|n|B}$, indicates that the surface states support a relativistic LL dispersion. Finally, electron-electron interactions were added to our topological insulator model for a finite magnetic field. Here, we derived an expression for the second order self-energy and its accompanying broadening function.

Acknowledgements

Table of Contents

List of Tables

List of Figures

List of Symbols

a	lattice constant
$A^{(0)}$	non-interacting spectral function
B	magnetic field
β	inverse temperature
$c, c^\dagger$	annihilation / creation operator
E_F	Fermi energy
G	real time Green's function
$\mathcal{G}$	Matsubara Green's function
Γ	broadening function
H_0	non-interacting Hamiltonian
ik_n, ip_n, iq_n	Matsubara frequencies
k, p, q	momentum
M	form factor
n	Landau level index
n_F, n_B	Fermi-Dirac distribution function, Bose-Einstein distribution function
ω, ϵ	energy
ψ	wavefunction
r	position / lattice site
Σ	self-energy
τ	imaginary time
V	interaction term

List of Acronyms

2D   two-dimensional

3D   three-dimensional

ARPES   angle-resolved photoemission spectroscopy

DOS   density of states

DC   Dirac cone

DP   Dirac point

ES   edge state

GF   Green's function

LL   Landau level

LHS   left-hand side

QSH   quantum spin Hall

QL   quintuple layer

STM   scanning tunnelling microscopy

STS   scanning tunnelling spectroscopy

SOC   spin-orbit coupling

SQUID   superconducting quantum interference device

SS   surface state

TR   time-reversal

TRIM   time-reversal invariant momentum

TI   topological insulator

Chapter 1

Introduction

Topological insulators (TIs) are a recently discovered topological phase of matter [1]. As such, they are insulating in the bulk, but unlike ordinary insulators they support topologically protected conducting edge or surface states [2–4]. Unlike previously discovered phases of matter, TIs are not classified by the symmetry they break and thus are not classified by a conventional order parameter. (For the non-topological ferromagnetic phase, rotational symmetry is broken and its order parameter is the magnetization of the material.) TIs do not break any symmetries and in order to differentiate a topological phase from a trivial phase, a quantized topological invariant is needed. Thus, for TIs, the order parameter is a topological invariant. The integer quantum Hall state is described by a non-zero Chern number, while the topological invariant describing the quantum spin Hall (QSH) state is the $\mathbb{Z}_2$ invariant. Topological invariants do not depend on local properties of the material. In order to transition between a trivial and a non-trivial topological state (both of which have a finite bulk bandgap), the bandgap must close. The closing of the gap is a second order quantum phase transition.

The potential practical applications of TIs include their use in low-power and spintronic devices, and in a topological quantum computer that uses Majorana fermions as qubits [2–4]. On a more fundamental level, TIs are being studied in the context of a broader search for novel exotic quantum phases of matter. For many of these applications and theoretical realizations to be manifested, a greater understanding of the electronic properties of TIs is required, in particular, the effects of electron-electron interactions and magnetic fields on surface states. In this introductory chapter, we review the basic physics of TIs and motivate

our theoretical study of electron-electron interactions and magnetic fields on TI surface states.

1.1 2D Topological Insulators

Two-dimensional (2D) TIs come in two known varieties: the Chern insulator, which breaks time-reversal (TR) symmetry, and the QSH insulator, which does not.

1.1.1 Chern Insulators

The first 2D TI was theorized in 1988 when Haldane constructed a model of the quantum anomalous Hall effect in 2D graphite, i.e., graphene [5]. (As will become clearer later, the quantum anomalous Hall effect signifies a quantum Hall effect in the absence of a net applied magnetic field). In this tight-binding model of a honeycomb lattice, TR symmetry is broken through the application of a periodic magnetic field that has an average magnetic flux of zero. The magnetic field opens a gap in the bandstructure at each of the two inequivalent Dirac points (DPs) of the graphene bandstructure, labelled K and K'. Mathematically, this gap is manifested in the addition of mass terms, m_K and $m_{K'}$, in the Dirac dispersions: $\epsilon_{K,K'} = \pm \hbar v |\boldsymbol{k}| \to \epsilon_{K,K'} = \pm\sqrt{(\hbar v \boldsymbol{k})^2 + (m_{K,K'} v^2)^2}$, where $\hbar$ is the reduced Planck constant, v is the Dirac fermion velocity, and $\boldsymbol{k} = (k_x, k_y)$ is the 2D wavevector measured with respect to K and K', respectively.

(a) Chiral edge states of the Chern insulator   (b) Helical edge states of the QSH insulator
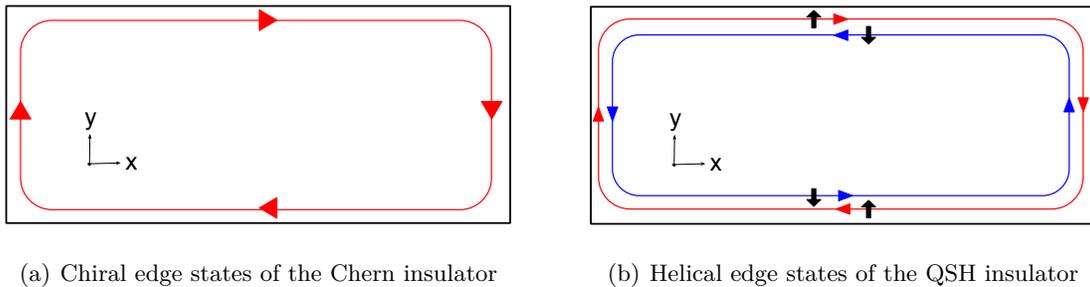
Figure 1.1: An overhead view of chiral (a) and helical (b) edge propagation for a 2D sample. The black box denotes the sample, the coloured arrows denote the direction of electron propagation on the sample edges, and in (b) the black arrows denote the spin of the electrons. (a) Electrons propagate in only one direction; electrons on opposite edges propagate in opposite directions. (b) Spin-up electrons (red) propagate in the opposite direction of spin-down electrons (blue). The helical edges are like two copies of chiral edges with opposite spin for opposite chiralities.

Depending on the relative signs of the mass terms at each of the two DPs, different phases of matter are realized [5]. When $\text{sgn}(m_K) = \text{sgn}(m_{K'})$, a conventional band insulator with a Hall conductance of $\sigma^{xy} = 0$ is predicted. On the other hand, if $\text{sgn}(m_K) \neq \text{sgn}(m_{K'})$, an integer quantum Hall state [6] with massless chiral edge states (ESs) and a Hall conductance of $\sigma^{xy} = \nu\frac{e^2}{h}$, where $\nu = \pm 1$ is predicted. Here, e is the charge on an electron and h is the Planck constant. In a chiral ES, electrons propagate in one direction (forward or backward) only (see Fig. 1.1a); such ESs are thus completely immune to both elastic and inelastic backscattering, provided the bulk bandgap remains open. In order to transition between these two states, the system must undergo a quantum phase transition. At the quantum critical point, one of the mass terms vanishes and the bulk bandgap closes. The phase where the masses have opposite signs is the Chern insulator state and its associated topological invariant is the Chern number, $\nu = \frac{1}{2}(\text{sgn}\,(m_K) - \text{sgn}(m_{K'}))$. It is essentially an integer quantum Hall state without Landau levels. By tuning the signs of the mass terms, the various quantum phases are accessed.

Chern insulators have been experimentally realized in magnetically doped 3D TIs [8–10] and in ultra-cold systems of fermionic atoms [11]. The bandstructure and probability densities of a representative Chern insulator model (distinct from the Haldane model, but exhibiting essentially the same physics) are shown in Fig. 1.2. The probability density of the ES decays exponentially away from the edge; this behaviour is well captured by an analytical solution of the 2D Dirac equation near a mass domain wall (Fig. 1.3). This is expected as one can think of the vacuum as a trivial insulator.

(a) Bandstructure

(b) Probability densities
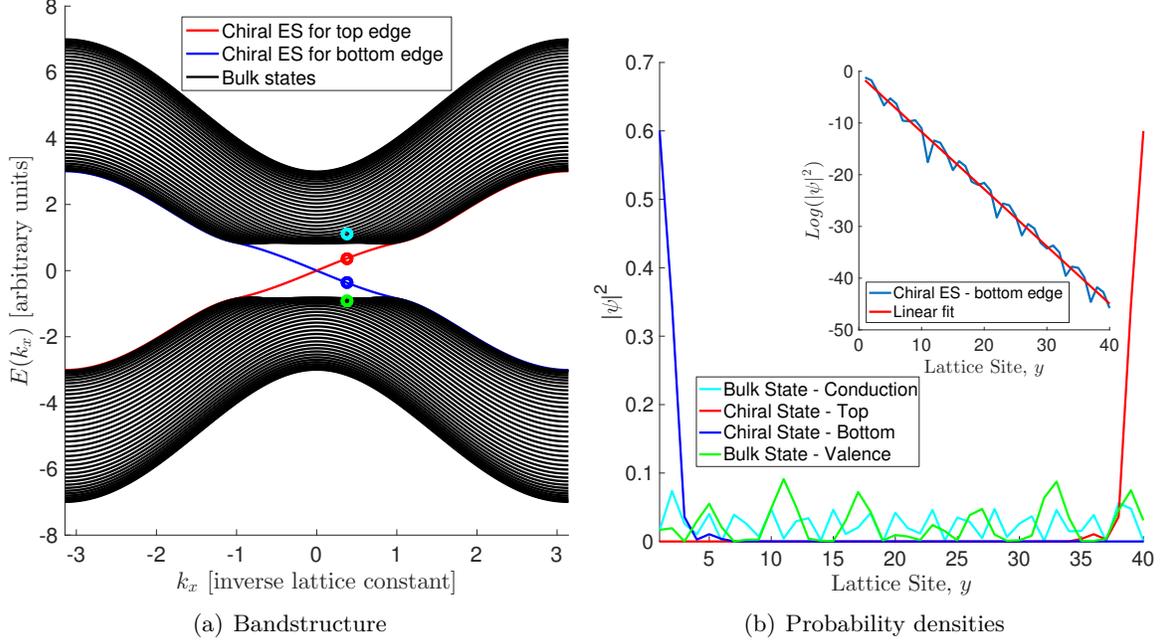
Figure 1.2: Numerical results for a model of a Chern insulator on a 2D square lattice with a periodic boundary condition in the x direction and an open boundary condition in the y direction [7]. (a) Bandstructure for a width of 40 sites. The chiral ESs disperse linearly near $k_x = 0$. The chirality of the ESs is manifested in the opposite velocities (slopes) for each edge. (b) Probability densities for the circled states in (a). The ESs are localized on opposite edges, while the valence and conduction band states have a roughly equal probability of being anywhere in the bulk. [Inset]: Logarithm of the probability density of the bottom chiral ES along with a linear fit to the result, indicating that the probability density decreases exponentially away from the edge.

$$\psi(y, k_x = 0) \propto$$
$$\exp\left(-2\frac{|m(y)vy|}{\hbar}\right)$$

$m(y) > 0 \qquad m(y) < 0$
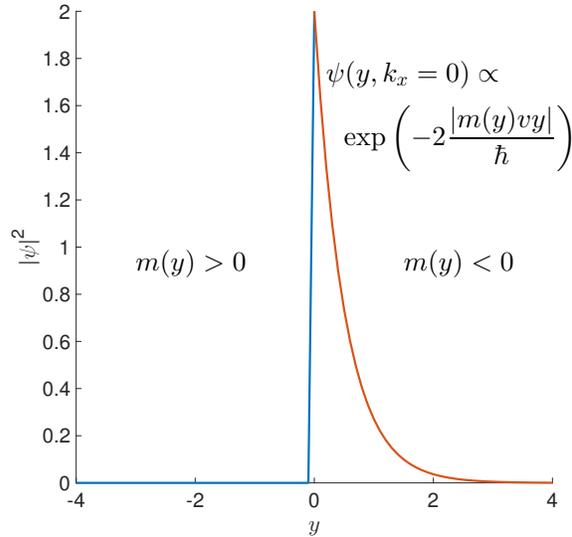
Figure 1.3: ES probability density derived from the 2D Dirac equation with a mass domain wall. The result is a plane wave in the x direction and a decreasing exponential in the y direction. If we let the masses become $m(y < 0) \to \infty$ and $m(y > 0)$ some finite negative value, then we can think of the region $y < 0$ as the vacuum and $y > 0$ as a Chern insulator with an edge at $y = 0$. As in Fig. 1.2b, the wavefunction is localized near the edge. As the mass changes sign at $y = 0$, the energy gap closes there. Thus, we can think of the mass domain wall as causing a phase transition spatially.

4

1.1.2 Quantum Spin Hall Effect

The QSH state, which was independently predicted by Kane and Mele in 2005 [12] and Bernevig and Zhang in 2006 [13], is another 2D TI. It can be thought of as two copies of the Haldane model described in Sec. 1.1.1, but where electrons of opposite spin have ESs of opposite chirality (see Fig. 1.1b). For this reason, these ESs are called helical. Unlike the Chern insulator, the QSH state preserves TR symmetry; however, like a Chern insulator, it is topologically distinct from a regular band insulator. While the integer quantum Hall state is characterized by a non-zero Chern number [14], a non-trivial $\mathbb{Z}_2$ topological invariant characterizes the QSH state [15, 16]. A $\mathbb{Z}_2$ invariant has only two possible values: trivial (0) or non-trivial (1). The $\mathbb{Z}_2$ topological invariant is only defined for TR invariant Hamiltonians. From now on, we call materials that exhibit the quantum anomalous Hall effect "Chern insulators", and materials that exhibit the QSH effect "2D TIs".

As TIs are band insulators, they have a bulk bandgap; however, they possess gapless ESs which, like the chiral ESs of the Chern insulator, exhibit topological properties [12]. The topological ESs exist because spin-orbit coupling (SOC) causes an inversion of the valence and conduction bands near the Fermi level, E_F [17]. Just as for the Chern insulator, this band inversion is equivalent to a change of sign of the mass in the effective Dirac description of the TI. The ESs are topologically protected because they are not affected by TR invariant local perturbations, as long as the bulk bandgap remains open [18]. If the bulk bandgap closes, electrons can elastically scatter from surface to bulk states, thereby destroying the topological protection of the ESs. An understanding of Kramers degeneracy is needed to understand why this is so.

Kramers' theorem [19] states that in a spin-half TR invariant system, every eigenstate must be at least doubly degenerate (e.g., $|\boldsymbol{k},\uparrow\rangle$ and $|-\boldsymbol{k},\downarrow\rangle$ are Kramers partners, where $\boldsymbol{k}$ denotes momentum). As a result, all TR invariant Hamiltonians have bandstructures that are symmetric about $\boldsymbol{k} = 0$ in the Brillouin zone. In addition, in the presence of a TR invariant perturbation, scattering between a state and its Kramers partner is not allowed. A qualitative explanation of this is given in Fig. 1.4. In a material such as graphene [20], where spin-up and spin-down Dirac cones (DCs) are degenerate, backscattering from $|\boldsymbol{k},\uparrow(\downarrow)\rangle$ to

$|-\boldsymbol{k}, \uparrow (\downarrow)\rangle$ is allowed. On the edge of the QSH state, backscattering is not allowed since there is no $|-\boldsymbol{k}, \uparrow (\downarrow)\rangle$ state into which the $|\boldsymbol{k}, \uparrow (\downarrow)\rangle$ state can scatter [3]. To scatter from $\boldsymbol{k}$ to $-\boldsymbol{k}$ requires a spin flip, which would violate TR symmetry [18]. In the limit of non-interacting electrons, this protection against elastic scattering leads to an infinite ES conductivity, i.e., ballistic transport on the edge.
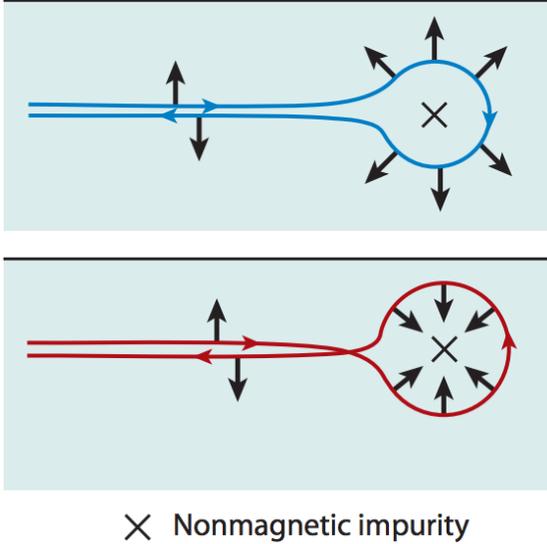
Figure 1.4: Intuitive picture for the topological protection of the QSH ES. When an electron scatters backwards off a non-magnetic impurity, we can imagine that its spin undergoes either a π or a $-\pi$ angle rotation (depending on the direction the electron "goes around" the impurity). A full 2π rotation of the electron's spin occurs when the relative phase difference between these two paths is considered. The spin-half nature of electrons ensures they pick up a minus sign while experiencing this 2π rotation. Therefore, the two backscattering paths interfere destructively and are not allowed. Figure taken from Ref. [18].

Another consequence of Kramers' theorem is the requirement that the crossing of the ESs in the edge Brillouin zone occur at TR invariant momentum (TRIM) points [2]. For the 1D ESs, these are $k = 0, \pm\frac{\pi}{a}$, where a is the lattice constant. The Fermi level can cross the bands connecting the TRIM points an odd or an even number of times. If there are an even number of crossings, a trivial insulator results as the ESs can be moved such that they are not crossed by the Fermi level [4]. If there are an odd number of crossings, the ESs are topologically protected as the Fermi level must cross the gapless ES. This is shown in Fig. 1.5. The degenerate crossing at TRIM points prevents the gapless ES from being gapped out without violating TR symmetry [3]. For this reason, the ESs of TIs are robust against TR symmetry-preserving disorder, as long as the bulk bandgap does not close [12]. If electron-electron interactions are accounted for however, then inelastic backscattering can occur, which leads to a finite conductivity in the ESs. Inelastic backscattering allows an electron to scatter a forward-moving state, $|\boldsymbol{k}, \uparrow\rangle$, to a backward-moving state, $|-\boldsymbol{k}', \downarrow\rangle$, at a different energy, which does not violate TR symmetry.
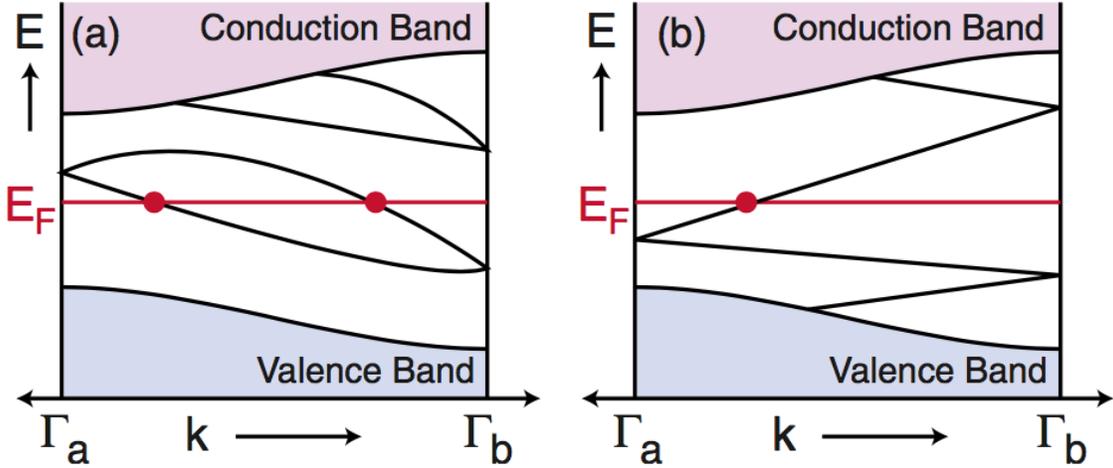
Figure 1.5: Bandstructure of a $\mathbb{Z}_2$ trivial insulator (a) and a $\mathbb{Z}_2$ non-trivial insulator (b). In each of the two bandstructures shown above, Kramers' theorem ensures that there is a double degeneracy at the TRIM points, Γ_a and Γ_b. (a) E_F crosses the ESs an even number of times and thus it is possible to move the ESs above or below the Fermi level, producing a trivial insulator. (b) There is an odd number of crossings, thereby ensuring the ESs cannot be gapped out. Figure taken from Ref. [2].

Due to the preservation of TR symmetry in 2D TIs, the ESs are helical, not chiral [18]. Because the electron's direction of propagation is linked to its spin, this phenomenon is called spin-momentum locking [2]. Thus, the ESs can transport both spin and charge currents [12]. Whereas the integer quantum Hall effect (e.g., in the Chern insulator) can be observed in macroscopic samples [21], the QSH effect is only robust when the phase coherence length, l_ϕ (the distance in the sample below which inelastic collisions can be neglected), is greater than the sample length, L [22]. Each ES, one for spin up electrons and one for spin down electrons, is predicted to contribute a conductance of $\frac{e^2}{h}$, for a total longitudinal conductance of $2\frac{e^2}{h}$ [13]. In the $L > l_\phi$ regime, inelastic backscattering can occur and the contribution from the ESs to the device's conductance is significantly reduced from its ballistic value.

1.1.3 Experimental Observation of 2D Topological Insulators

The original theoretical predictions of the QSH effect in graphene [15] and strained semiconductors [13] have not yet been experimentally realized. In 2006, Bernevig *et al.* predicted that HgTe/CdTe quantum wells can support the QSH state [17]. They proposed that by

varying the thickness, d, of the HgTe well from $d < d_c$ to $d > d_c$, where $d_c \approx 6.4$ nm is a critical thickness, the trivial insulator undergoes a phase transition to the QSH state. Changing the thickness of the well is akin to tuning the mass, M, of an effective Dirac Hamiltonian, as in the Chern insulator (see Sec. 1.1.1) [1, 17]. A summary of their predictions is shown in Table. 1.1.

Property	Trivial Insulator	Quantum Critical Point	QSH State
Well thickness	$d < d_c$	$d = d_c$	$d > d_c$
Dirac mass	$M < 0$	$M = 0$	$M > 0$
Bandstructure	Normal gap	Bands cross	Inverted gap
$\mathbb{Z}_2$	0	n/d	1
Helical ESs	No	n/d	Yes
Conductance	0	n/d	$2\frac{e^2}{h}$

Table 1.1: Trivial vs QSH insulator in HgTe/CdTe quantum wells. n/d means "not defined". Source [17].

Bernevig *et al.* proposed that the ES could be detected by a conductance measurement [17]. They predicted that in the trivially insulating phase, the longitudinal conductance is zero, but in the topologically insulating regime it is $2\frac{e^2}{h}$. This residual conductance arises from the ESs, and should therefore be independent of the sample width. In 2007, König *et al.* experimentally realized the QSH state in HgTe/CdTe quantum wells for $d > d_c \approx 6.3$ nm [1]. As shown in Fig. 1.6a, there is a much larger resistance in the $d < d_c$ regime, than in the $d > d_c$ regime. The smaller resistance in the $d > d_c$ regime is due to the residual conductance of the ESs. This conductance is independent of the sample width and is measured to be approximately $2\frac{e^2}{h}$ for sample lengths $L < l_\phi$, as predicted [17].

While the measurements by König *et al.* were convincing, to definitively prove that conduction in HgTe quantum wells was due to ESs, non-local transport experiments were performed [23]. The idea was to rule out conductance quantization from topologically trivial quasi-1D bulk channels in the ballistic regime. In 2009, Roth *et al.* conducted multiterminal resistance measurements on HgTe quantum wells in various non-local configurations. The non-local transport measurements are only explainable by current flow along the edges of the sample. Further evidence of the topological nature of HgTe quantum wells occurred in

2012 when the spin polarization of the helical ESs was measured via the inverse spin Hall effect [24].
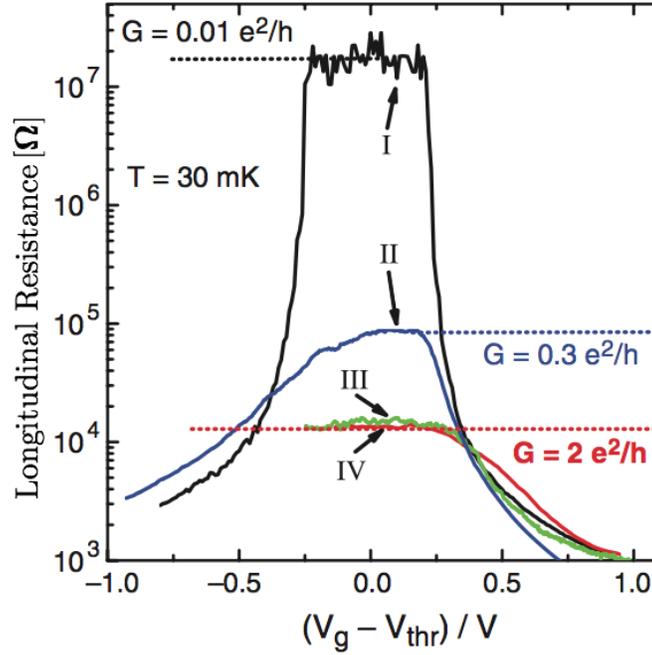
(a) Resistance of HgTe quantum wells (adapted from König *et al.* [1])

Line Colour	Thickness [μm]	Length $\times$ Width [μm]
Black	5.5	20.0 $\times$ 13.3
Blue	7.3	20.0 $\times$ 13.3
Green	7.3	1.0 $\times$ 1.0
Red	7.3	1.0 $\times$ 0.5

(b) Labels for the curves

Figure 1.6: QSH state in HgTe/CdTe quantum wells. (a) Longitudinal resistance measured using leads at the end points of a sample (think of leads at the left and right edges of Fig. 1.1b). (b) The labels and dimensions for each of the four curves shown in (a). The black curve is a trivially insulating phase, while the other three curves suggest the existence of ESs. The reason for the larger resistance in the blue curve, compared to the red and green curves, was the larger sample length, which is larger than the phase coherence length, $l_\phi \approx 1\,\mu$m. This allowed for inelastic scattering to occur, thereby increasing the resistance. The same measured residual conductance for the green and red curves indicates the conductance was independent of the sample width, as expected for ES transport [17]. Figure taken from Ref. [1].

In 2013, Nowack *et al.* used a superconducting quantum interference device (SQUID) to detect the ESs of HgTe quantum wells [25]. A SQUID directly images the magnetic fields produced by conducting ESs and thus produces images of the current density in a sample.

They measured simultaneous bulk and edge conduction and observed that even when the HgTe edges are much longer than the phase coherence length, the ESs are still a significant source of transport.

In 2012, Knez *et al.* conducted transport measurements that indicated the presence of helical ESs in InAs/GaSb quantum wells [22]. Like HgTe quantum wells, InAs/GaSb quantum wells exhibit a topologically distinct inverted bandstructure. Unlike in HgTe quantum wells however, InAs/GaSb quantum wells exhibit significant bulk conductivity, even larger than the ES conductance. Similar to the experiments conducted by Nowack *et al.*, Spanton *et al.* [26] used a SQUID to measure the current density of InAs/GaSb quantum wells. The current density images they produced showed predominantly edge conduction when the Fermi level was in the gap and uniform bulk conduction when it was in the conduction band. They used devices that were much larger than the phase coherence length; as such, they detected backscattering processes which they deduced occur along the sample edges. Du *et al.* doped InAs/GaSb quantum wells with Si to remove the bulk conductivity, and measured a $2\frac{e^2}{h}$ conductance of the helical edges when $L < l_\phi$ [27]. In the $L > l_\phi$ regime, the conductance was reduced and the resistance of the edges in the InAs/GaSb samples scaled with the sample length, owing to inelastic backscattering effects.

Taken together these results show that the QSH state has been realized in HgTe and InAs/GaSb quantum wells. Thus, we can call these materials 2D topological insulators.

1.2   3D Topological Insulators

Topological insulators have also been realized in 3D materials [2–4]. Similar to 2D TIs, they come in two types: strong and weak. We focus on strong TIs here.

1.2.1   From Two to Three Dimensions

Shortly after the prediction of 2D TIs, theoretical work began on generalizing the QSH effect to 3D [28–30]. Whereas 2D TIs are described by a single $\mathbb{Z}_2$ invariant, 3D TIs are described by four independent $\mathbb{Z}_2$ invariants. Of these four invariants, one of them discriminates between strong and weak TIs. Strong TIs, like the QSH state, are robust to disorder. Conversely, disorder destroys the topological nature of weak TIs and thus they easily revert

back to being trivial band insulators. From now on, when we refer to 3D TIs, it is implied that we mean strong 3D TIs.

Just as the QSH state has topologically protected ESs, 3D TIs have topologically protected surface states (SSs) [31]. The SS of a 3D TI is protected by the non-trivial π Berry's phase that an electron acquires when its path circles a Fermi surface enclosing an odd number of DPs [28]. In simple terms, this means states on opposite points of the Fermi surface (i.e., those related by backscattering) have opposite spin (see Fig. 1.7). A trivial Berry's phase is zero and results from a Fermi surface enclosing an even number of DPs. The π Berry's phase causes anti-localization of the electrons in the presence of disorder. The SS is a "2D topological metal" that can only exist on the boundary of a 3D system. For the easiest experimental observation of the SS, both the Fermi level and the DP must lie in the bulk bandgap [32].
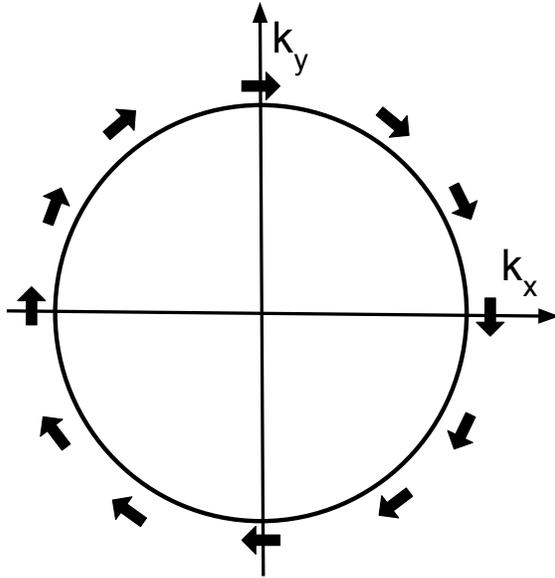
Figure 1.7: Constant energy contour of a DC in the 2D surface Brillouin zone of a 3D TI. The spin direction is given by the arrows. This shows the spin-momentum locking, or helical nature of the SS (the π Berry's phase). This is the 2D analog of the helical property of QSH ESs mentioned earlier.

1.2.2 Experimental Observation of 3D Topological Insulators

In 2007, Fu and Kane predicted that the $Bi_{1-x}Sb_x$ alloy is a 3D TI for $0.07 < x < 0.22$ [31]. The strong SOC in $Bi_{1-x}Sb_x$ inverts its bandstructure for $x > 0.07$ and the TI phase may be realized [33, 34]. In 2008, Hsieh *et al.* used angle-resolved photoemission spectroscopy (ARPES) to map the surface bandstructure of $Bi_{0.9}Sb_{0.1}$. They found five crossings of the

Fermi level with the gapless SS bands [33, 35]. As mentioned above, the odd number of crossings show that these states are topologically non-trivial.

Due to the complexity of $Bi_{1-x}Sb_x$ – five band crossings on the surface and its non-crystalline nature – 3D TIs with simpler SSs, simpler theoretical descriptions, and a stoichiometric crystal structure were, and still are, desired [35, 36]. In 2009, Zhang *et al.* predicted that the Bi_2Se_3 class of materials (Bi_2Se_3, Bi_2Te_3, Sb_2Te_3) are TIs with a single massless DC at the Γ point. This class of materials enjoys the same rhombohedral crystal structure arranged in quintuple layers (QLs) and exhibits a band inversion at the Γ point. In 2009, Xia *et al.* used ARPES measurements to map the bandstructure of the SSs of Bi_2Se_3. They found a single non-degenerate DC on the surface that exhibited a linear dispersion. An ARPES image of a Bi_2Se_3 DC is shown in Fig. 1.8a. Moreover, their results match the theoretical predictions as the DP lies at the Γ point in the bulk bandgap and a bandgap of 0.3 eV was measured [35]. The large bulk bandgap allows for topological properties to be realized at room temperature since 0.3 eV is much greater than the thermal energy at room temperature. Similar results were obtained in Bi_2Te_3 and Sb_2Te_3 where a non-degenerate massless DC on the surface was discovered for each material [37, 38]. An ARPES image of Bi_2Te_3 is shown in Fig. 1.8b.

To definitively prove that these materials are 3D TIs, the helical nature of the SSs, i.e., the π Berry's phase, needed to be measured [40]. Experimental confirmation of the helical nature of SSs in Ca-doped Bi_2Se_3 and Sn-doped Bi_2Te_3 occurred in 2009 [40]. Using spin-resolved ARPES, the spin texture of the SSs was found to exhibit a left-handed rotation around the Fermi surface. The helical nature of the SSs was subsequently determined using the same technique. In addition, using scanning tunnelling microscopy (STM), scanning tunnelling spectroscopy (STS), and ARPES, scattering between $|\boldsymbol{k}, \uparrow\rangle \rightarrow |-\boldsymbol{k}, \downarrow\rangle$ in the presence of non-magnetic disorder was not seen in $Bi_{1-x}Sb_x$ [41] or Bi_2Te_3 [42, 43]. This verifies the absence of disorder-caused elastic backscattering and further supports the topological nature of the SSs.

Perfectly crystalline Bi_2Se_3, Bi_2Te_3, and Sb_2Te_3 are predicted to be bulk insulators with the Fermi level in the bulk bandgap [3, 36]; however, vacancies and antisite defects in their

crystalline structures produce a residual bulk conductivity [37, 40]. The relatively large bulk-to-surface conductivity ratio makes it difficult to observe and exploit the desirable SS properties [32]. The Fermi level of imperfect Bi_2Se_3 and Bi_2Te_3 crystals lies in the bulk conduction band [35], and for Sb_2Te_3, in the bulk valence band. By hole doping Bi_2Se_3 with Ca and Bi_2Te_3 with Sn, their Fermi levels are tuneable into the bulk bandgap, which reduces the bulk conductivity. Nevertheless, even with doping, the bulk conductivity is still relatively large and new TI compounds with more desirable electronic properties have been investigated [44, 45].

(a) Bi_2Se_3 (b) Bi_2Te_3
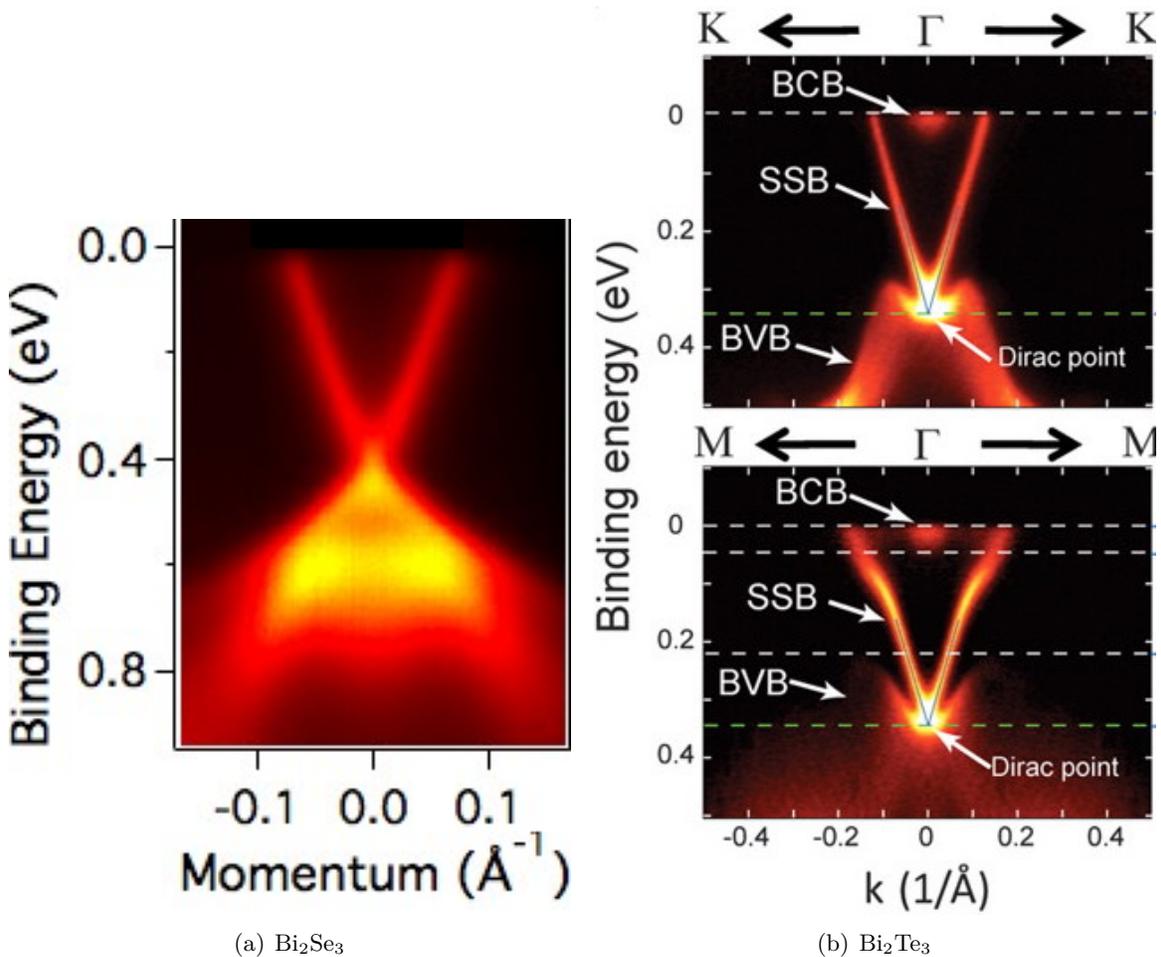
Figure 1.8: ARPES intensity plots showing the bandstructures of Bi_2Se_3 (a) and Bi_2Te_3 (b). The linear dispersion of the DCs is clearly seen. In (b), BCB is the bulk conduction band, BVB is the bulk valence band, and SSB is the surface state band. The brighter regions indicate more intense photoemission. Figures taken from Ref. [37, 39].

One such compound is Bi_2Te_2Se [44, 45]. It has a large bulk resistivity which yields a much greater surface than bulk carrier mobility. An issue with this material is that the DP lies below the bulk valence band which makes it impossible to isolate the SSs [32]. Another compound with a large bulk resistivity is Sn-doped $Bi_{1.1}Sb_{0.9}Te_2S$. It has a larger bandgap than the Bi_2Se_3 family of materials. Furthermore, its DP lies firmly in the bulk bandgap and its conduction is dominated by the SSs. Aside from the materials mentioned above, there are many other 3D TIs that have been both theoretically predicted and experimentally realized [3, 46].

3D TIs have been grown not only as bulk crystals but also as thin films [47]. If a 3D TI film is too thin, the SS wavefunctions on opposite surfaces can overlap and electrons on one surface can tunnel to the other. This opens a gap at the DP, which is seen in Bi_2Se_3 films thinner than six QLs (Fig. 1.9). At a critical thickness of six QLs, Bi_2Se_3 contains a gapless DC. Above a six QL thickness, Bi_2Se_3 is essentially a bulk insulating TI with a well-defined DC on the surface [48]. For Bi_2Se_3, a single QL has a thickness of approximately 1 nm [50]. In Sb_2Te_3, a DC is seen down to a thickness of five QL [49].
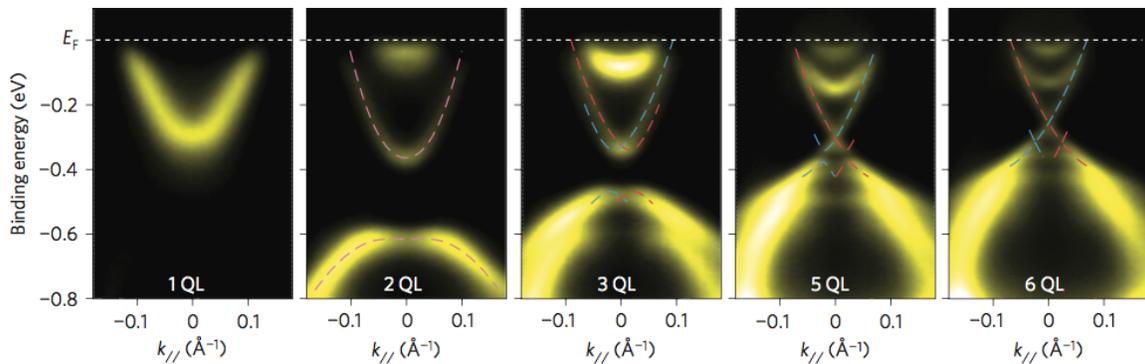
Figure 1.9: ARPES images of the Bi_2Se_3 bandstructure in the $\Gamma - K$ direction for various QL thicknesses. When the critical thickness of six QLs is reached, the gap disappears and a massless DC is realized. Figure taken from Ref. [47].

1.2.3 Electron-Electron Interactions

As TIs are a relatively new discovery in condensed matter physics, there are many phenomena that need to be studied in more depth. One such area is electron-electron interactions. Indeed, the above-cited theoretical descriptions have neglected the consequences of electron-

electron interactions, while in practice electrons interact via the Coulomb repulsion. However, a number of recent experimental observations cannot be explained by non-interacting theories. Some of these are described below.

In thin films of Bi_2Se_3, Wang *et al.*'s temperature and magnetic field dependent transport measurements do not agree with 2D transport theory for non-interacting electrons; however, when electron-electron interactions are considered, theoretical predictions match their data very well [51]. Another example of the necessity to incorporate electron-electron interactions into TI models is seen in the 2D limit of TIs [52, 53]. These experiments showed that at low temperature the resistivity of Bi_2Se_3 samples less than six QLs thick increases logarithmically as temperature decreases. This result stands in contrast with the expected behaviour of metallic substances, for which we expect the resistivity to decrease with decreasing temperature. Both experimenters conclude that strong electron-electron interactions are the only explanation for this deviation from the expected behaviour. More specifically, one of them states that the strong interactions occur from decreased Coulombic screening. Moreover, Park *et al.* used ARPES measurements on Bi_2Se_3 to examine quasiparticle scattering mechanisms and to estimate the SS self-energy [54]. They determined that in addition to electron-phonon and electron-impurity interactions, electron-electron interactions cause scattering from surface to bulk states.

Song *et al.* also observed electron-electron interaction-induced surface-to-bulk scattering [55]. They used STM and STS to measure the spatial decay of standing waves off steps in thin films of Bi_2Se_3 which allowed them to determine the inelastic quasiparticle lifetime. They found that the lifetime increases with the TI thin film thickness, which they attribute to decreased Coulombic screening, and goes as $(E-E_F)^{-2}$ (see Fig. 1.10). This enhancement of the quasiparticle lifetime near the Fermi level as caused by electron-electron interactions is also observed in the Landau quantization experiments described in the following section.
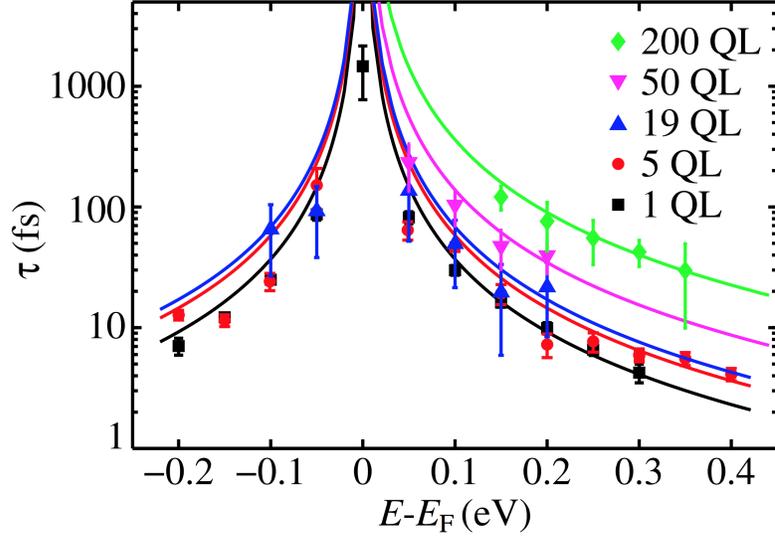
Figure 1.10: The electron quasiparticle lifetime, τ, near the Fermi level. The lifetime is greatest at the Fermi level and decreases away from this energy. Figure taken from Ref. [55].

1.2.4 Landau Levels

Due to the presence of the metallic SSs on 3D TIs, Landau level quantization in the surface density of states (DOS) can be observed. The detailed theory of LLs is reviewed in Sec. 4.1. In this section we simply state that a perpendicular magnetic field, B, applied to a metal causes charged particles to undergo quantized orbital motion. For a 2D metallic surface, these quantized orbits manifest as discrete levels in the dispersion relation, and are called LLs. In the limit of non-interacting electrons, the LLs appear as Dirac delta function peaks in the DOS and thus have an infinite lifetime.

The differential conductance (dI/dV) in scanning tunnelling experiments, which is a measure of the local DOS, is shown in Fig. 1.11 for a TI under zero magnetic field and in Fig. 1.12 in the presence of an applied magnetic field. For finite fields, LLs corresponding to peaks in the dI/dV trace are observed. A magnetic field-independent LL located at the DP is also observed. The observation of a field-independent LL is a strong indication that the measured LLs are due to the Dirac SSs, and not the bulk states [56]. This is because the bulk LLs would not be relativistic as they have parabolic bands.

Clearly, these LLs are not Dirac delta function peaks as they have a finite width. The broadened LL peaks in Fig. 1.12 indicate that the quasiparticles have a finite lifetime. The experiments show that the quasiparticle lifetime is enhanced near E_F and E_{DP}, and

decreases as one moves away from these energies [56]. Jiang *et al.* conclude, as does Pauly *et al.*, that electron-electron interactions are the only viable explanation for the observed peak width distribution [58, 59]. Moreover, Jiang *et al.* attribute the electron-electron interactions to scattering amongst SSs, not amongst surface to bulk states, due to the large bulk bandgap.
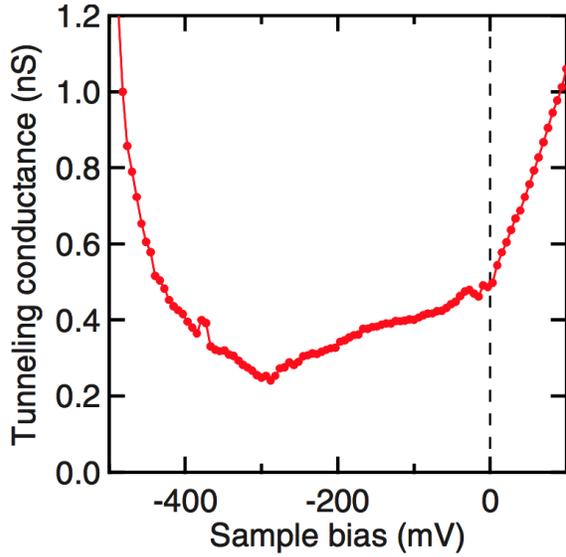
Figure 1.11: Differential conductance for zero magnetic field in Bi_2Se_3. The DP is located at -290 mV, which is the minimum of the curve. The sharp increase in conductance below -400 mV and above 0 mV is attributed to the bulk states. As is discussed in Sec. 3.2, the V-shaped curve and minimum at the DP are indicative of a surface 2D DC. Figure taken from Ref. [56].

(a) Bi$_2$Se$_3$

(b) Bi$_2$Se$_3$
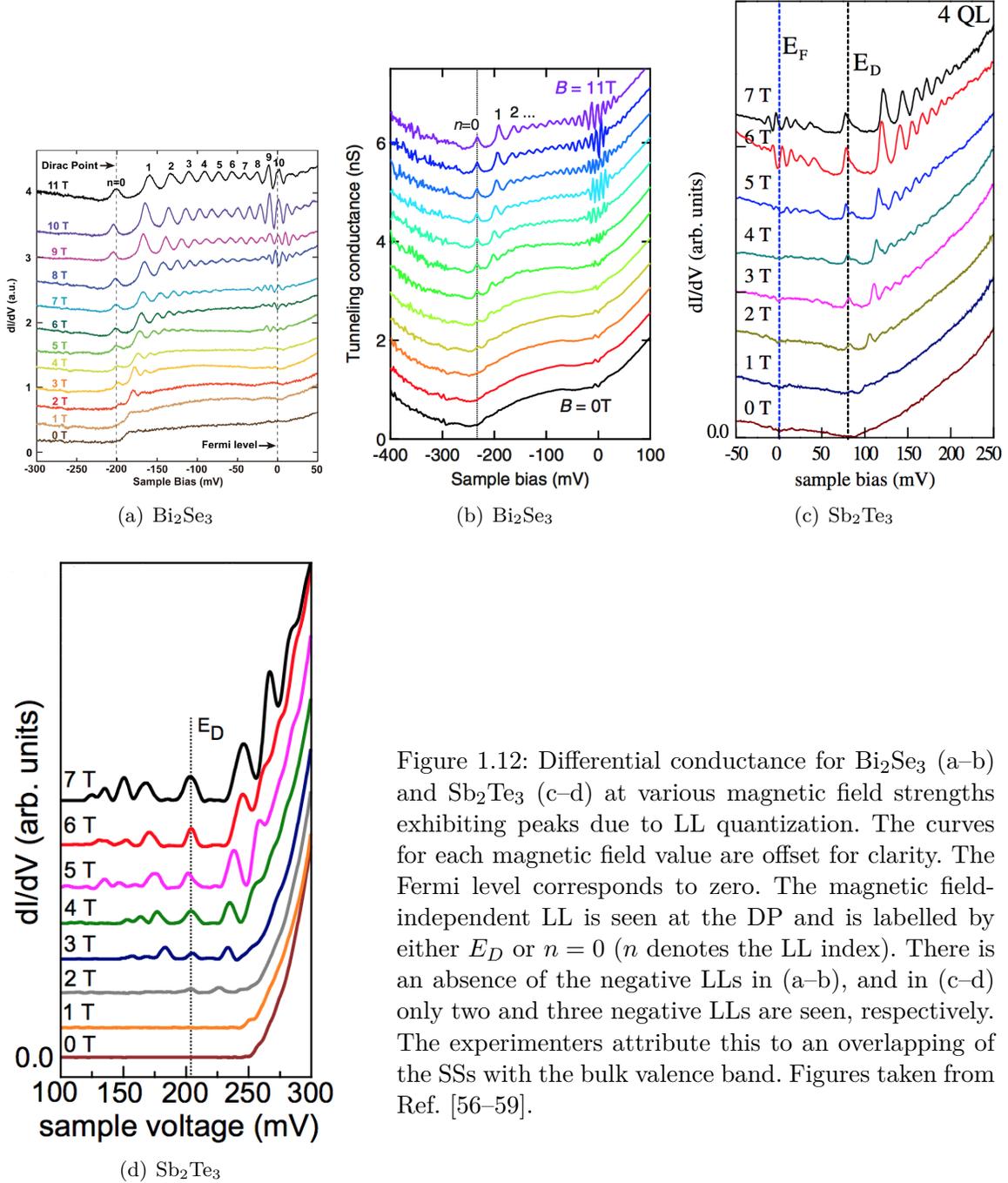
(c) Sb$_2$Te$_3$

(d) Sb$_2$Te$_3$

Figure 1.12: Differential conductance for Bi$_2$Se$_3$ (a–b) and Sb$_2$Te$_3$ (c–d) at various magnetic field strengths exhibiting peaks due to LL quantization. The curves for each magnetic field value are offset for clarity. The Fermi level corresponds to zero. The magnetic field-independent LL is seen at the DP and is labelled by either E_D or $n = 0$ (n denotes the LL index). There is an absence of the negative LLs in (a–b), and in (c–d) only two and three negative LLs are seen, respectively. The experimenters attribute this to an overlapping of the SSs with the bulk valence band. Figures taken from Ref. [56–59].

As shown in Sec. 4.1, the energies of Dirac SS LL peaks should be linear in $\sqrt{|n|B}$. The experimentally determined energies of the LL peaks are plotted in Fig. 1.13 against either $\sqrt{|n|B}$ or the LL momentum, k_n, where $k_n \equiv \sqrt{2e|n|B/\hbar}$. All of the dispersion relations in Fig. 1.13 deviate from linearity. Jiang *et al.* attribute the non-linearity to a tip-gating effect and were able to show that it is not an intrinsic effect [58].

Cheng *et al.* deposited Ag atoms on the surface of some samples to ensure that the measured LLs were due to the SSs [57]. The higher the density of Ag atoms, the more suppressed the LL spectrum became; thus, establishing that the results in Fig. 1.12a are due to SSs.

In addition to the above mentioned experimental investigations of LLs on 3D TI surfaces, there have been various theoretical studies. Shen used a 2D Dirac equation to study the integer quantum Hall effect in TI SSs [60], while Yang and Han used a four-band Hamiltonian (similar to the one described in Sec. 3.1) to examine the surface LLs [61]. Neither of these approaches incorporated any type of interaction in their models, nor did they model explicitly the effects of sample thickness. An electrostatic potential term was added to the 2D Dirac Hamiltonian by Schwab and Dzierzawa [62] to explain how the STM tip suppresses the negative LLs, as has been experimentally observed [56, 57]. Nevertheless, they did not incorporate electron-electron interactions into their approach.

(a) Bi_2Se_3

(b) Bi_2Se_3

(c) Sb_2Te_3

(d) Sb_2Te_3
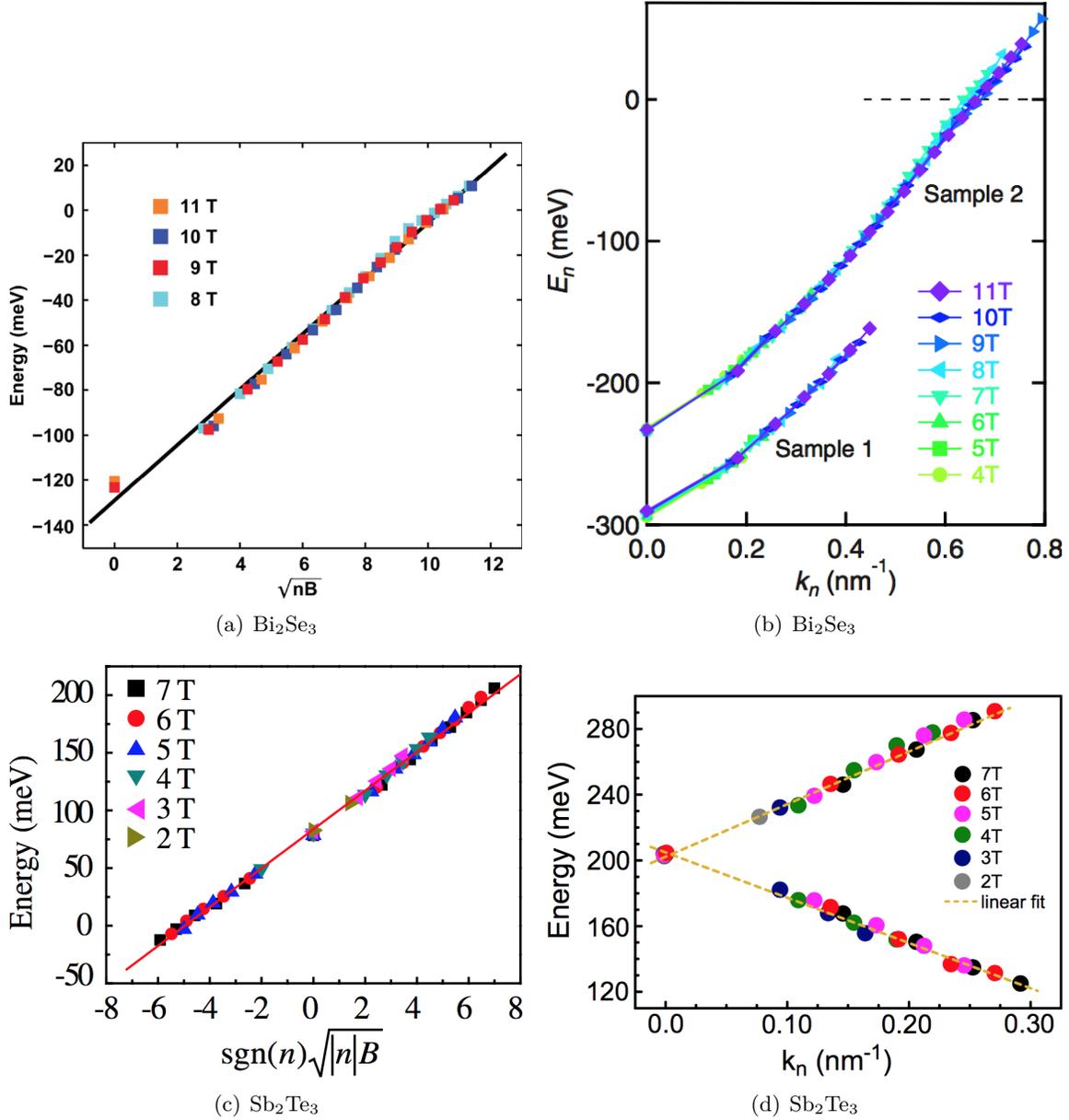
Figure 1.13: The LL energies and their fits to a $\sqrt{|n|B}$ or k_n dependence. The near linear nature of these dispersion relations indicates that these SS electrons are massless Dirac fermions (see Sec. 4.1). The curve fits in (b) appear less linear than those in the other three plots due to its much larger k_n domain, which goes much farther from the Γ point. For ease of comparison, $\sqrt{nB} = 10$ is approximately $k_n = 0.155$ nm^{-1}. Figures taken from Ref. [56–59]

1.3   Motivation and Overview

To understand the experimental results discussed in Sec. 1.2.3 and 1.2.4 it is necessary to incorporate electron-electron interactions and a finite magnetic field into a theoretical description of TIs. Thus far, there have been no theoretical TI models that have accomplished this while simultaneously taking into account the full TI bandstructure and the thickness dependence of the samples.

In this thesis three main quantities are examined. First, the effects of electron-electron interactions are examined through the calculation of the self-energy for zero magnetic field. This allows us to qualitatively understand the behaviour of the quasiparticle lifetime near the Fermi level. Second, the behaviour of electrons in the presence of a perpendicular magnetic field is analyzed for a non-interacting TI. From this we obtain the LL spectrum of the SSs, and how it is affected by the thickness of the TI film. Third, a many-body formalism is developed that allows for the calculation of the self-energy in the presence of a magnetic field. Although explicit numerical calculations were not performed here due to lack of time, this formalism allows one to calculate the DOS in the presence of interactions and should reproduce the broadened LL peak structure seen in the experiments mentioned in the previous section. The remainder of this thesis describes how these quantities were calculated and presents and discusses the corresponding results.

Ch. 2 describes the theoretical foundation on which the model developed for this thesis is constructed. Specifically, it describes the second quantization formalism, tight-binding and Hubbard models, many-body Green's functions and how to calculate them, and many-body perturbation theory (Feynman diagrams). Next, in Ch. 3, the Hamiltonian used in this model is described, as are results in the non-interacting limit, and results with interactions in zero magnetic field. In the final chapter, Ch. 4, we begin by reviewing the physics of LLs in both conventional and Dirac systems. Then, a magnetic field is added to our model, first in the non-interacting limit, and then in the presence of interactions. Where possible, the results obtained in Ch. 3 and 4 are compared to experiment and/or expected theoretical behaviour.

Chapter 2

Theoretical Methods

This chapter provides a general overview of the theoretical methods used in this thesis. First, the second quantization formalism, which is ideally suited to treat interacting many-particle systems, is introduced. Next, tight-binding and Hubbard models in a periodic crystal are reviewed. Then, the real and imaginary time many-body Green's functions and many-body perturbation theory (Feynman diagrams) are introduced. Finally, an algorithm that allows for the numerical calculation of the retarded Green's function is summarized. In general, the theory described in this section applies to both bosons and fermions; however, as the problem we wish to solve involves electrons, we focus on fermions. Here, and in the rest of this work, we work in natural units, $\hbar = k_B = 1$, unless otherwise specified.

2.1 Second Quantization

In first quantization, a generic N-particle fermionic wavefunction can be written as a linear combination of Slater determinants of the single particle eigenstates [63]. For large N, this can become cumbersome and alternative theoretical approaches are desirable. Second quantization accomplishes this as it is an inherently many-body approach. It has several advantages over the first quantization formalism. Some of these benefits include: particle statistics is intrinsic to the theory; perturbation theory is easily incorporated through Feynman diagrams and Dyson's equation; and the ability to create and destroy particles and excitations in a system is built in [64]. This last point allows for the study of quasiparticle excitations and their lifetimes, which is one of the goals of this thesis.

In second quantization, the basis states for a general many-body system are labelled by the occupation numbers for all single particle states,

$$N\text{-particle basis states: } |n_{\nu_1}, n_{\nu_2}, \ldots, n_{\nu_N}\rangle \tag{2.1}$$

The corresponding Hilbert space is known as Fock space. The number operator, $\hat{n}_{v_j}$, is defined such that

$$\hat{n}_{v_j} |n_{v_j}\rangle = n_{v_j} |n_{v_j}\rangle \tag{2.2}$$

where the corresponding eigenvalue, n_{v_j}, is the number of particles in state ν_j. This is why an alternative name for second quantization is the occupation number formalism. In general, we only write a hat (ˆ) on operators when we are in the interaction picture. An exception was made here to differentiate the number operator from the number eigenvalue.

Depending on the statistics to which the particles belong, n_{ν_j} can take on different values:

$$n_{\nu_j} = \begin{cases} 0, 1 & \text{fermions} \\ 0, 1, 2, \ldots & \text{bosons} \end{cases} \tag{2.3}$$

The number operator is defined as $\hat{n}_{v_j} \equiv c_{\nu_j}^\dagger c_{\nu_j}$, where $c_{\nu_j}^\dagger$ is the creation operator and c_{ν_j} is the annihilation operator. The creation operator creates a particle in state ν_j, while the annihilation operator destroys a particle in state ν_j. In second-quantization no explicit (anti-)symmetrization of the wavefunctions is required as particle statistics is automatically taken into account via the values n_{ν_j} can take and in the (anti-)commutation relations of the creation and annihilation operators [63]. As we are concerned with electrons, our operators obey Fermi statistics. The fermionic creation/annihilation operators obey the following anti-commutation rules:

$$\{c_{\nu_j}^\dagger, c_{\nu_k}^\dagger\} = 0\,, \qquad \{c_{\nu_j}, c_{\nu_k}\} = 0\,, \qquad \{c_{\nu_j}, c_{\nu_k}^\dagger\} = \delta_{\nu_j,\nu_k} \tag{2.4}$$

where δ_{ν_j,ν_k} is the Kronecker delta function.

2.2 Model Hamiltonians

In this thesis we use a tight-binding Hamiltonian to describe the propagation of non-interacting electrons, and a generalized Hubbard model to incorporate electron-electron interactions for electrons propagating through a 3D TI.

2.2.1 Tight-Binding Model

Consider a crystal lattice of regularly spaced atoms. If this spacing is comparable to the size of the orbitals of the atomic wavefunctions, then the orbitals of neighbouring atoms can overlap. If this occurs, valence electrons can tunnel between nearby atoms. A tight-binding model describes systems where electrons are "tightly bound" to the atom, but can tunnel or hop to nearby atoms due to the orbital overlap [65]. It is a single electron model that allows for the calculation of bandstructures.

In second quantization, the simplest type of tight-binding Hamiltonian, where an electron can tunnel only between adjacent lattice sites, is

$$\mathcal{H}_{TB} = -t \sum_{\langle \boldsymbol{rr'} \rangle, \sigma} c_{\boldsymbol{r}\sigma}^{\dagger} \, c_{\boldsymbol{r'}\sigma} - \mu \sum_{\boldsymbol{r}\sigma} c_{\boldsymbol{r}\sigma}^{\dagger} \, c_{\boldsymbol{r}\sigma} \tag{2.5}$$

where t is the hopping strength, $\langle \boldsymbol{rr'} \rangle$ denotes a sum over nearest-neighbour lattice sites, σ is the spin index, and μ is the chemical potential. Eq. 2.5 signifies that a particle is destroyed on the $\boldsymbol{r'}$ lattice site with spin σ and a new particle is created on one of its neighbouring sites, $\boldsymbol{r}$, also with spin σ. As there are no interaction terms (i.e., terms containing products of four or more creation/annihilation operators) in Eq. 2.5, the tight-binding model is a purely kinetic model. In this simple case, we assume there is only one orbital per site and hopping is spin-independent. For the work in this thesis we must consider more than one orbital per site, as well as the effect of SOC (see Sec. 3.1), and thus need a more general tight-binding Hamiltonian of the form

$$\mathcal{H}_{TB} = \sum_{\boldsymbol{rr'}} c_{\boldsymbol{r}\alpha}^{\dagger} \, h_{\boldsymbol{rr'}}^{\alpha\beta} \, c_{\boldsymbol{r'}\beta} \tag{2.6}$$

24

where α, β stand for combined orbital and spin degrees of freedom and h is a Hermitian matrix. Here, and thereafter, we use the Einstein summation convention for the orbital/spin indices, according to which indices repeated twice are summed over. In addition, the sum over $\boldsymbol{rr'}$ need not be restricted to nearest-neighbour lattice sites, and we can even have $\boldsymbol{r} = \boldsymbol{r'}$, which denotes an on-site energy. In particular, we include the chemical potential term in the definition of h.

2.2.2 Hubbard Model

The Hubbard model, which was first proposed in 1963 [66], improves upon the tight-binding model through the addition of an interaction between electrons on the same lattice site. It has been used to study ferromagnetism, high-temperature superconductors, Mott insulators, ultra-cold atom trapping, and many other physical systems [67, 68]. This on-site interaction arises from the Coulomb repulsion between nearby electrons. The interaction between electrons on different lattice sites is ignored due to the Coulomb interaction being screened by other electrons in the crystal [65]. The screened Coulomb interaction is modelled as a contact interaction. That is, the electrons must be on the same lattice site – in contact with one another – for the interaction to occur. Obviously, this is an oversimplification; however, it is a simple way of implementing at least some amount of interactions in a system described by a tight-binding model.

The original Hubbard model only concerns one atomic orbital [66]. For a spin-half system, such as the one described by Eq. 2.7, there are thus only four possible configurations on the $\boldsymbol{r}^{th}$ lattice site: no electron, one spin-up electron, one spin-down electron, or one spin-up and one spin-down electron. Of course, the interaction only occurs in the last case. The interaction cannot occur between electrons of the same spin state due to the Pauli exclusion principle. For spin-half fermions, the original Hubbard Hamiltonian is

$$\mathcal{H}_{\text{Hubbard}} = -t \sum_{\langle \boldsymbol{rr'} \rangle, \sigma} c_{\boldsymbol{r}\sigma}^{\dagger} c_{\boldsymbol{r'}\sigma} + U_{\uparrow\downarrow} \sum_{\boldsymbol{r}} n_{\boldsymbol{r}\uparrow} n_{\boldsymbol{r}\downarrow} - \mu \sum_{\boldsymbol{r}\sigma} c_{\boldsymbol{r}\sigma}^{\dagger} c_{\boldsymbol{r}\sigma} \tag{2.7}$$

where the hopping term is the same as in Eq. 2.5 and $U_{\uparrow\downarrow}$ is the strength of the on-site interaction between the spin-up ($\uparrow$) and spin-down ($\downarrow$) electrons. Any tight-binding model

augmented by an on-site interaction is termed a Hubbard model (or generalized Hubbard model). For the general tight-binding Hamiltonian in Eq. 2.6 we consider a Hubbard Hamiltonian of the form

$$\mathcal{H}_{\text{Hubbard}} = \sum_{\boldsymbol{rr'}} c_{\boldsymbol{r}\alpha}^{\dagger} \, h_{\boldsymbol{rr'}}^{\alpha\beta} \, c_{\boldsymbol{r'}\beta} + U_{\alpha\beta} \sum_{\boldsymbol{r}} n_{\boldsymbol{r}\alpha} \, n_{\boldsymbol{r}\beta} \tag{2.8}$$

where $U_{\alpha\beta}$ is a symmetric matrix with 0 along the diagonal. This accounts for the fact that interactions between electrons in different orbitals are also possible. A generalized Hubbard Hamiltonian of the form of Eq. 2.8 is the model used in this thesis.

2.3 Many-Body Green's Functions

The many-body Green's function (GF), $G(\nu, \nu', t, t')$, is the probability amplitude for a particle in one single-particle state, ν' at time t', to be found in another state, ν at a later time t [63, 69]. The GF can be thought of as a matrix, where the single-particle indices ν and ν' define the final and initial state, respectively. Although it does not contain as much information as the full many-body wavefunction, it does allow for the calculation of the ground state energy, some of the excited states, and the ground state expectation value of a single-particle operator [64]. The GF is calculated using the full many-body Hamiltonian, $H = H_0 + V$, where H_0 denotes the non-interacting part (for us, the tight-binding Hamiltonian) and V contains the interactions (for us, the Hubbard interaction). Thus, even though it only describes the propagation of one particle, this single particle propagation is affected by all other particles in the system. In addition, the many-body GF formalism enables us to calculate the self-energy of this propagating particle, from which we can calculate the broadening of single-particle levels and deduce the existence of a finite quasiparticle lifetime.

2.3.1 Real Time Green's Functions

If we choose the single-particle states ν and ν' to be localized orbital/spin states α and β at sites $\boldsymbol{r}$ and $\boldsymbol{r'}$, respectively, as in Eq. 2.6, the real time-ordered GF at zero-temperature

is defined as

$$G_{rr'}^{\alpha\beta}(t,t') = -i \left\langle \Psi_0 \left| T \left[c_{r\alpha}(t) \, c_{r'\beta}^{\dagger}(t') \right] \right| \Psi_0 \right\rangle \tag{2.9}$$

where $c_{r\alpha}(t)$, $c_{r'\beta}^{\dagger}(t')$ are annihilation/creation operators in the Heisenberg picture, T is the time-ordering operator, and $|\Psi_0\rangle$ is the ground state of the interacting many-body Hamiltonian. As we are concerned with causal processes, we require $t > t'$ [63, 64]. This condition produces the retarded GF,

$$G_{rr'}^{R\,\alpha\beta}(t,t') = -i\Theta(t-t') \left\langle \Psi_0 \left| \left\{ c_{r\alpha}(t), c_{r'\beta}^{\dagger}(t') \right\} \right| \Psi_0 \right\rangle \tag{2.10}$$

where the Heaviside step function, $\Theta(t-t')$, ensures $t > t'$ and $\{\dots\}$ is the anti-commutator.

For a time-independent Hamiltonian, the GF only depends on the time difference, $t-t'$. If a system is translationally invariant (i.e., has periodic boundary conditions), then the GF only depends on the spatial difference $r-r'$. We can then Fourier transform from $r-r'$, $t-t'$ to k, ω, where k is a momentum and ω is a frequency.

2.3.2 Imaginary Time Green's Functions

As defined above, the real time GF only gives results at zero temperature. Since experiments occur at finite temperature, and temperature effects are important considerations in any physical system, we want to be able to incorporate such effects into our model. This is possible using the Matsubara, or imaginary time, formalism where we let $t \rightarrow -i\tau$, where τ is real and is called the imaginary time [63, 64].

In the imaginary time Heisenberg picture, the annihilation and creation operators can be written as

$$c_r(\tau) = e^{H\tau} \, c_r \, e^{-H\tau} , \qquad c_{r'}^{\dagger}(\tau') = e^{H\tau'} \, c_{r'}^{\dagger} \, e^{-H\tau'} \tag{2.11}$$

where H is the full many-body Hamiltonian.

The position space Matsubara GF is

$$\mathcal{G}^{\alpha\beta}_{\boldsymbol{rr'}}(\tau, \tau') = -\left\langle T_\tau \left[c_{\boldsymbol{r}\alpha}(\tau) \, c^\dagger_{\boldsymbol{r'}\beta}(\tau') \right] \right\rangle \tag{2.12}$$

where T_τ is the imaginary time-ordering operator. The average in Eq. 2.12 is with respect to the equilibrium density matrix, $\rho = e^{-\beta H}/Z$, of the full interacting Hamiltonian, where $Z = \mathrm{Tr}\left[e^{-\beta H} \right]$ is the partition function and β is the inverse temperature. This is in contrast to Eq. 2.9 and 2.10, where the average is respect to the ground state. Writing Eq. 2.12 as

$$\mathcal{G}^{\alpha\beta}_{\boldsymbol{rr'}}(\tau, \tau') = -\frac{1}{Z} \mathrm{Tr}\left[e^{-\beta H} T_\tau \left[c_{\boldsymbol{r}\alpha}(\tau) \, c^\dagger_{\boldsymbol{r'}\beta}(\tau') \right] \right] \tag{2.13}$$

explicitly shows the average is with respect to the density matrix.

The main issue to be addressed is that one is not in general able to directly calculate averages with respect to the interacting density matrix. The main goal of many-body perturbation theory, which we now derive, is to express the interacting average, Eq. 2.13, in terms of averages with respect to the non-interacting density matrix, $\rho_0 = e^{-\beta H_0}/Z_0$, where $Z_0 = \mathrm{Tr}\left[e^{-\beta H_0} \right]$ is the non-interacting partition function. This derivation is best carried out with the help of the imaginary-time evolution operator, $\hat{U}(\tau, \tau')$, which evolves a state from τ' to τ [63]:

$$\hat{U}(\tau, \tau') = e^{H_0\tau} \, e^{-H(\tau - \tau')} \, e^{-H_0\tau'} \tag{2.14}$$

An integral equation for $\hat{U}(\tau, \tau')$ can be obtained by iteratively solving the following differential equation, and substituting in Eq. 2.14:

$$\frac{\partial \hat{U}(\tau, \tau')}{\partial \tau} = -\hat{V}(\tau) \, \hat{U}(\tau, \tau') \tag{2.15}$$

where $\hat{V}(\tau)$ denotes the interaction term in the interaction picture, i.e., $\hat{V} = e^{H_0\tau} \, V \, e^{-H_0\tau}$. In the remainder of this thesis the hat denotes operators in the interaction picture. Solving

order by order in $\hat{V}$, we obtain

$$\hat{U}(\tau, \tau') = \sum_{n=0}^{\infty} \frac{1}{n!} (-1)^n \int_{\tau'}^{\tau} d\tau_1 \cdots \int_{\tau'}^{\tau} d\tau_n \left[\hat{V}(\tau_1) \ldots \hat{V}(\tau_n) \right] = T_\tau \left[e^{-\int_{\tau'}^{\tau} d\tau_1 \hat{V}(\tau_1)} \right] \quad (2.16)$$

Using Eq. 2.14 and 2.16, we can write

$$e^{-\beta H} = e^{-\beta H_0} \hat{U}(\beta, 0) = e^{-\beta H_0} T_\tau \left[e^{-\int_0^\beta d\tau_1 \hat{V}(\tau_1)} \right] \quad (2.17)$$

where we have let $\tau' \to 0$ and $\tau \to \beta$. Substituting Eq. 2.17 into Eq. 2.13 yields

$$\mathcal{G}_{\boldsymbol{rr'}}^{\alpha\beta}(\tau, \tau') = -\frac{1}{\text{Tr}\left[e^{-\beta H_0} \hat{U}(\beta, 0)\right]} \text{Tr}\left[e^{-\beta H_0} T_\tau \left[\hat{U}(\beta, 0) \, \hat{c}_{\boldsymbol{r}\alpha}(\tau) \, \hat{c}_{\boldsymbol{r'}\beta}^\dagger(\tau')\right]\right] \quad (2.18)$$

Dividing the numerator and denominator by the non-interacting partition function, we obtain

$$\mathcal{G}_{\boldsymbol{rr'}}^{\alpha\beta}(\tau, \tau') = -\frac{\left\langle T_\tau \left[\hat{U}(\beta, 0) \, \hat{c}_{\boldsymbol{r}\alpha}(\tau) \, \hat{c}_{\boldsymbol{r'}\beta}^\dagger(\tau')\right]\right\rangle_0}{\left\langle \hat{U}(\beta, 0) \right\rangle_0} \quad (2.19)$$

The subscript "0" indicates that the expectation value is computed with respect to the non-interacting density matrix. Using Eq. 2.16, Eq. 2.19 becomes

$$\mathcal{G}_{\boldsymbol{rr'}}^{\alpha\beta}(\tau, \tau')$$
$$= -\frac{\sum_{n=0}^{\infty} \frac{1}{n!} (-1)^n \int_0^\beta d\tau_1 \cdots \int_0^\beta d\tau_n \left\langle T_\tau \left[\hat{V}(\tau_1) \ldots \hat{V}(\tau_n) \, \hat{c}_{\boldsymbol{r}\alpha}(\tau) \, \hat{c}_{\boldsymbol{r'}\beta}^\dagger(\tau')\right]\right\rangle_0}{\sum_{n=0}^{\infty} \frac{1}{n!} (-1)^n \int_0^\beta d\tau_1 \cdots \int_0^\beta d\tau_n \left\langle T_\tau \left[\hat{V}(\tau_1) \ldots \hat{V}(\tau_n)\right]\right\rangle_0} \quad (2.20)$$

All averages are now expressed in terms of the non-interacting density matrix, and can thus be computed in principle. The average calculated with respect to the non-interacting density matrix, or perhaps more explicitly the appearance of β, in Eq. 2.20, shows that we have incorporated temperature effects into the GF.

2.3.3 Many-body Perturbation Theory: Feynman Diagrams

For perturbation theory to be valid, the interaction term, $\hat{V}(\tau)$, must be small. Assuming this is true, the goal now is to evaluate the time-ordered averages in Eq. 2.20. As the interaction term can be expressed in terms of products of creation and annihilation operators, the main task is to evaluate averages of time-ordered products of creation and annihilation operators. Directly evaluating the time-ordered operations is complicated as all of the anti-commutators must be evaluated. Any time an anti-commutator does not yield zero, an extra term arises. Wick's theorem, which is not proven here, provides a systematic procedure for evaluating such averages for an arbitrary number of creation/annihilation operators [63, 64]. In short, the theorem states that the average of a product of creation/annihilation operators with respect to a non-interacting density matrix is given by the product of all possible contractions, i.e., averages of all possible pairs of creation/annihilation operators.

One way to understand Wick's theorem is to look at an analogous problem: calculating the average of $\langle x^n \rangle = \langle xx \ldots x \rangle$ for a Gaussian random variable, x. The results for various powers of x are shown in Table 2.1 in terms of the variance, σ^2.

n	$\langle x^n \rangle$
$n = \text{odd}$	zero
$n = 2$	$\sigma^2 = \langle xx \rangle$
$n = 4$	$3\,\sigma^4 = 3\,\langle xx \rangle \langle xx \rangle$
$n = 6$	$15\,\sigma^6 = 15\,\langle xx \rangle \langle xx \rangle \langle xx \rangle$
$\vdots$	$\vdots$

Table 2.1: Gaussian averages of various powers of the random variable x. All averages for even powers of x can be written in terms of $\sigma^2 = \langle xx \rangle$.

The result for any even power of x is

$$\langle x^{2n} \rangle = \frac{(2n)!}{2^n n!}\,\sigma^{2n}\,, \qquad n = 1, 2, 3, \ldots \tag{2.21}$$

The Gaussian average for $2n$ random variables is a product of the Gaussian average of two of those random variables. The coefficient in Eq. 2.21 is the number of ways that the $2n$ x's can be paired up or contracted. Wick's theorem works the same way when the creation

and annihilation operators are contracted in Eq. 2.20: it decomposes all of the possible contractions into pairs of operators that have a non-zero contribution to the GF [64].

The contributing terms can be drawn as Feynman diagrams [63]. Feynman diagrams are a visual representation of the perturbative expansion of the GFs and can be drawn in a systematic manner. There are two classes of Feynman diagrams: connected (Fig. 2.1a) and disconnected (Fig. 2.1b). Fortunately, the disconnected diagrams need not be calculated as the denominator in Eq. 2.20 cancels their contribution with that in the numerator. Thus, we can re-write Eq. 2.20 as

$$\mathcal{G}_{rr'}^{\alpha\beta}(\tau, \tau') =$$
$$-\sum_{n=0}^{\infty} \frac{1}{n!} (-1)^n \int_0^{\beta} d\tau_1 \cdots \int_0^{\beta} d\tau_n \left\langle T_\tau \left[\hat{V}(\tau_1) \ldots \hat{V}(\tau_n) \, \hat{c}_{r\alpha}(\tau) \, \hat{c}_{r'\beta}^\dagger(\tau') \right] \right\rangle_0^{\text{connected}} \quad (2.22)$$

(a) Connected
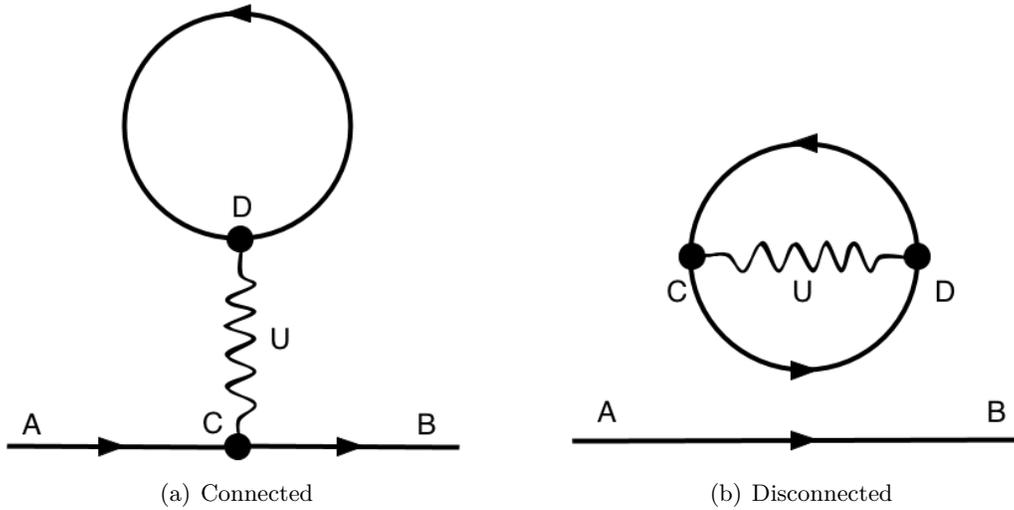
(b) Disconnected

Figure 2.1: First order connected (a) and disconnected (b) Feynman diagrams for a particle propagating from A to B. U denotes the interaction between a particle at C and a particle at D.

When drawing Feynman diagrams, a free particle GF is drawn as a solid line with an arrow, while an interaction is denoted by a squiggly line. For Eq. 2.22 expanded to n^{th} order, there are n interactions and $2n + 1$ GFs. Fig. 2.1a contains the Hartree diagram, which is one of the two connected first order diagrams. Examining Fig. 2.1a, we can think of the particle at A propagating freely to C, interacting with another particle through U, and

then continuing to propagate freely to B. Similar to Feynman diagrams in particle physics, higher order diagrams can be drawn.

2.3.4 Returning to Real Time Green's Functions

Although perturbation theory at finite temperature is best achieved in terms of imaginary time GFs, as we have seen, we are ultimately interested in physical properties, which are extracted from real time GFs. The retarded GF of Eq. 2.10 (or, rather, its Fourier transform in the frequency domain) can in fact be obtained from the Matsubara GF, as we briefly explain here. Two of the properties of Matsubara GFs are time-translational invariance,

$$\mathcal{G}(\tau, \tau') = \mathcal{G}(\tau - \tau') \tag{2.23}$$

stemming from the time independence of the Hamiltonian, and (anti-)periodicity in imaginary time,

$$\mathcal{G}(\tau) = \pm \mathcal{G}(\tau + \beta) \tag{2.24}$$

where fermions pick up the minus sign [63, 64]. As a consequence of this, the Matsubara GF can be expanded as a Fourier series using

$$\mathcal{G}(ik_n) = \int_0^\beta d\tau \, e^{ik_n \tau} \, \mathcal{G}(\tau) \tag{2.25}$$

where $k_n = \frac{(2n+1)\pi}{\beta}$ is the Matsubara frequency for fermions and n is an integer. ik_n can be thought of as the Matsubara formalism's analog of the frequency, ω.

The retarded GF of Eq. 2.10 is obtained from the analytic continuation of the Matsubara GF, i.e., by letting $ik_n \to \omega + i\eta$, where η is a positive infinitesimal:

$$G_{\boldsymbol{rr'}}^{R\,\alpha\beta}(\omega) = \lim_{ik_n \to \omega + i\eta} \mathcal{G}_{\boldsymbol{rr'}}^{\alpha\beta}(ik_n) \tag{2.26}$$

We have now returned to real time/frequency. The retarded GF in turn yields many physical properties, such as the spectral function and DOS (see Sec. 2.3.6).

2.3.5 Self-Energy

An electron propagating through a crystal interacts with other particles, such as other electrons and phonons [65, 69]. The propagating electron affects the other particles, and the other particles affect it. These interactions modify the electron's properties from those of an electron propagating alone. The self-energy, Σ, is a quantity that accounts for these modified properties of interacting electrons [63, 64]. To differentiate a particle propagating alone from a particle that is "dressed" by interactions with other particles, we call the "dressed" particle a quasiparticle.

The real part of the self-energy shifts or renormalizes the energy spectrum of the quasiparticle from that of a bare particle, while the imaginary part concerns the quasiparticle lifetime [63, 64, 69]. The quasiparticle lifetime is the time over which the quasiparticle can be considered to propagate coherently with its shifted energy spectrum. While non-interacting particles have an infinite lifetime, quasiparticles acquire a finite lifetime as a result of interactions. However, a quasiparticle on the Fermi surface has an infinite lifetime. Just as with a bare particle, a quasiparticle has a momentum and spin [70].

The retarded self-energy, Σ^R, enters Dyson's equation

$$G^R = G^{R(0)} + G^{R(0)} \, \Sigma^R \, G^R \tag{2.27}$$

where $G^{R(0)}$ is the GF in the absence of interactions. This equation can be understood as the re-summation of all possible Feynman diagrams in the perturbative expansion (Eq. 2.22) of the GF. In practice, the self-energy is calculated up to a certain finite order in perturbation theory. Once the self-energy is obtained, the interacting GF, which describes the propagation of the quasiparticle, can be calculated by solving Eq. 2.27 for G^R:

$$G^R = \left(\left(G^{R(0)} \right)^{-1} - \Sigma^R \right)^{-1} = G^{R(0)} + G^{R(0)} \, \Sigma^R \, G^{R(0)} + G^{R(0)} \, \Sigma^R \, G^{R(0)} \, \Sigma^R \, G^{R(0)} + \dots$$

$$\tag{2.28}$$

Thus, even when the self-energy is calculated only to finite order, the GF obtained from Dyson's equation contains processes to infinite order in perturbation theory.

In the simplest situation of a single band of electrons, the non-interacting retarded GF in the momentum representation is

$$G^{R(0)}(\boldsymbol{k}, \omega) = \frac{1}{\omega + i\eta - \xi_{\boldsymbol{k}}} \tag{2.29}$$

where the poles of this function give the energy, $\xi_{\boldsymbol{k}} = \epsilon_{\boldsymbol{k}} - \mu$, of the bare particle measured with respect to the chemical potential [63]. If electron-electron interactions are present, Eq. 2.29 becomes

$$G^{R}(\boldsymbol{k}, \omega) = \frac{1}{\omega + i\eta - \xi_{\boldsymbol{k}} - \Sigma^{R}(\boldsymbol{k}, \omega)} \tag{2.30}$$

It is clear that $\text{Re}\left[\Sigma^{R}(\boldsymbol{k}, \omega)\right]$ changes the dispersion relation from that of a bare particle because it shifts the position of the poles of the GF.

2.3.6 Spectral Function and Density of States

Roughly speaking, the spectral function is the probability that a particle with a given energy will be found in a given single-particle state [63]. Considering for instance the one-band problem just mentioned, the spectral function, $A(\boldsymbol{k}, \omega)$, is the probability that an electron with energy ω will be found in the single-particle state with momentum $\boldsymbol{k}$. It can be measured by ARPES. For non-interacting particles, the spectral function is a Dirac delta function,

$$A^{(0)}(\boldsymbol{k}, \omega) = 2\pi\, \delta(\omega - \xi_{\boldsymbol{k}}) \tag{2.31}$$

This reflects the fact that in the absence of interactions the energy of each individual electron is conserved and specified by the single-particle energy, $\xi_{\boldsymbol{k}}$. Interactions broaden the delta functions. The width of the broadened peaks yields the inverse lifetime of the quasiparticle in the given quantum state. In the presence of interactions, we obtain a Lorentzian-like

form,

$$A(\boldsymbol{k},\omega) = \frac{-2\,\mathrm{Im}\left[\Sigma^R(\boldsymbol{k},\omega)\right]}{\left(\omega - \xi_{\boldsymbol{k}} - \mathrm{Re}\left[\Sigma^R(\boldsymbol{k},\omega)\right]\right)^2 + \left(\mathrm{Im}\left[\Sigma^R(\boldsymbol{k},\omega)\right]\right)^2} \tag{2.32}$$

As was seen in Eq. 2.30, $\mathrm{Re}\left[\Sigma^R(\boldsymbol{k},\omega)\right]$ shifts the energy peak of the quasiparticle, and thus changes its dispersion relation. Conversely, $-2\,\mathrm{Im}\left[\Sigma^R(\boldsymbol{k},\omega)\right]$, which can be shown to be positive, is proportional to the width of the spectral function and thus to the inverse quasiparticle lifetime. Although these specific forms of the GF and spectral function are not used in this thesis, Eq. 2.32 illustrates the general feature that the width of the spectral function peaks, and thus the quasiparticle lifetime of interest in this thesis, are related to the imaginary part of the self-energy. If the self-energy is a matrix (which is the case in this work), then the Hermitian part of the self-energy concerns the renormalized energies and the anti-Hermitian part concerns the lifetime.

For the general interacting Hamiltonian, Eq. 2.8, the interacting spectral function is a matrix given by

$$A_{\boldsymbol{r}\boldsymbol{r}'}^{\alpha\beta}(\omega) = i\,\left(G^R(\omega) - \left(G^R(\omega)\right)^\dagger\right)_{\boldsymbol{r}\boldsymbol{r}'}^{\alpha\beta} \tag{2.33}$$

The DOS, which is what STM experiments effectively measure (recall Fig. 1.11 and 1.12), can be calculated from the spectral function,

$$\rho(\omega) = \frac{1}{2\pi}\,\mathrm{Tr}\left[A(\omega)\right] \tag{2.34}$$

where the trace is taken over spatial and orbital/spin indices.

2.4 Numerical Calculation of the Non-interacting Green's Function

In order to calculate the self-energy, which is given by Feynman diagrams involving free particle propagators, we must first calculate the non-interacting GF. For the problem considered in this thesis, the one-band GF of Eq. 2.29 is not sufficient, as we have multiple bands. Furthermore, because of SOC the GF is not diagonal in its orbital/spin indices.

Finally, we study systems with a surface in order to see the topological SSs, thus we do not have translational invariance in all directions and cannot completely Fourier transform to momentum space. Therefore, numerical methods are required to determine the GF, even at the non-interacting level. The non-interacting retarded GF at a given energy, ω, is given by the matrix inverse

$$G^R(\omega) = (z\mathbb{1} - h)^{-1} \tag{2.35}$$

where $z = \omega + i\eta$, $\mathbb{1}$ is the identity matrix, and h is the tight-binding Hamiltonian matrix in Eq. 2.6 [71]. For an $n \times n$ matrix, the computational cost of matrix inversion scales as $\mathcal{O}(n^3)$ when Gaussian elimination is used. Thus, for large matrices, such as the ones used in this thesis, this is a very computationally intensive calculation. Fortunately, the Hamiltonian used in this thesis is sparse. The sparseness of this matrix allows us to use certain algorithms to more quickly calculate the GF.

To calculate the surface and bulk GFs we use a matrix inversion algorithm developed by Wu *et al.* [71] specifically designed for problems with finite range hopping. We call this algorithm "Finite-N". The first step in computing the GF is to convert the Hamiltonian to block tri-diagonal form:

$$h = \begin{pmatrix} H_{00} & H_{01} & 0 & \dots \\ H_{10} & H_{00} & H_{01} & \\ 0 & H_{10} & H_{00} & \\ \vdots & & & \ddots \end{pmatrix} \tag{2.36}$$

Any matrix with a finite range hopping can be put in this form.

The advantage of this algorithm is that it operates on the much smaller H_{00}, H_{01}, and H_{10} blocks of the h matrix. The size of these blocks does not scale with the Hamiltonian matrix size, and thus we avoid the $\mathcal{O}(n^3)$ scaling of Gaussian elimination. In addition, this algorithm works on blocks of various sizes.

For a square matrix with n_{pl} blocks of the size of H_{00}, the GF can be calculated as follows:

1. Calculate the so-called backward GFs: $\Delta_1^-, \ldots, \Delta_{n_{pl}}^-$

   (a) $\Delta_1^- = (z\mathbb{1} - H_{00})^{-1}$

   (b) $\Delta_{i+1}^- = ((z\mathbb{1} - H_{00}) - H_{10}\,\Delta_i^-\,H_{01})^{-1}, \quad i = 1, 2, \ldots, n_{pl} - 1$

2. Calculate the so-called forward GFs: $\Delta_{n_{pl}}^+, \ldots, \Delta_1^+$

   (a) $\Delta_{n_{pl}}^+ = (z\mathbb{1} - H_{00})^{-1}$

   (b) $\Delta_{i-1}^+ = ((z\mathbb{1} - H_{00}) - H_{01}\,\Delta_i^+\,H_{10})^{-1}, \quad i = n_{pl}, n_{pl} - 1, \ldots, 2$

3. Calculate the diagonal blocks of the GF

   (a) $G_{1,1} = \Delta_1^+$

   (b) $G_{n,n} = \Delta_{n_{pl}}^-$

   (c) $G_{i,i} = \left((z\mathbb{1} - H_{00}) - H_{01}\,\Delta_{i+1}^+\,H_{10} - H_{10}\,\Delta_{i-1}^-\,H_{01}\right)^{-1}$

4. Calculate the off-diagonal blocks of the GF

   (a) $G_{i,i'} = G_{i,i}\,H_{01}\,\Delta_{i+1}^+\ldots H_{01}\,\Delta_{i'}^+, \quad i' > i$

   (b) $G_{i,i'} = G_{i,i}\,H_{10}\,\Delta_{i-1}^-\ldots H_{10}\,\Delta_{i'}^-, \quad i > i'$

5. Construct the full retarded GF matrix from the $G_{i,i'}$ blocks

Chapter 3

Interacting Surface States for Zero Magnetic Field

In this chapter, we introduce the effective Hamiltonian used to model the 3D TI Bi$_2$Se$_3$. From this Hamiltonian, we obtained the bandstructure of Bi$_2$Se$_3$ near its DP and found that states lying in the bulk bandgap are localized to the surface of the TI. The non-interacting spectral function and DOS were also calculated and yield the expected result for Dirac fermions. The first and second order self-energies were calculated from which we were able to infer the behaviour of the quasiparticle lifetime.

3.1 Non-Interacting Hamiltonian

The full Hamiltonian used in this thesis has the general form

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{3.1}$$

where $\hat{H}_0$ is the non-interacting part and $\hat{V}$ is the interacting part. More specifically, we use a generalized Hubbard model for $\hat{H}$, where $\hat{H}_0$ has the form of the first term in Eq. 2.8 and $\hat{V}$ has the form of the second term. The non-interacting part of the Hamiltonian used in this thesis is taken from ab initio calculations by Liu *et al.* [72]. This four band Hamiltonian describes isotropic bulk Bi$_2$Se$_3$, Bi$_2$Te$_3$, and Sb$_2$Te$_3$; however, we only show results for Bi$_2$Se$_3$. Only the valence orbitals, which are p orbitals for these TIs, are considered. The Hamiltonian is only accurate near the Γ point, $|\boldsymbol{k}| < 0.04$ Å^{-1}. This works fine for our purposes as we are interested in the SS physics near the DP, which is located at the Γ point.

The non-interacting Hamiltonian was obtained by considering various phenomena [72]. First, the hybridization between the outer p orbitals of Bi and Se was considered. Next, bonding and anti-bonding states were accounted for, as was crystal field splitting. Finally, SOC was added which inverts the energy bands about the Fermi level. The two orbitals closest to the Fermi level were retained and form the basis of the effective Hamiltonian. There are four orbital/spin basis states: spin-up and spin-down in the valence band, and spin-up and spin-down in the conduction band. Due to SOC, the electrons propagate as a linear combination of the orbital/spin states.

We ignore certain higher order effects, such as hexagonal warping of the DC [73], and thus only keep terms up to quadratic order in momentum from Liu *et al.*'s Hamiltonian. These terms are small as long as we are sufficiently close to the DP. The effective Hamiltonian used in this thesis is

$$h_{\text{eff}}(\boldsymbol{k}) = \epsilon_{\boldsymbol{k}} \, \mathbb{1} + \mathcal{M}(\boldsymbol{k}) \, \Gamma_5 + B_0 \, k_z \, \Gamma_4 + A_0 \, (k_y \, \Gamma_1 - k_x \, \Gamma_2) \tag{3.2}$$

where

$$\epsilon_{\boldsymbol{k}} = C_0 + C_1 \, k_z^2 + C_2 \, \boldsymbol{k}_{\parallel}{}^2 \tag{3.3}$$

$$\mathcal{M}(\boldsymbol{k}) = M_0 + M_1 \, k_z^2 + M_2 \, \boldsymbol{k}_{\parallel}{}^2 \tag{3.4}$$

$\boldsymbol{k}_{\parallel} = (k_x, k_y)$, a is the lattice constant, and $C_{0,1,2}$, $M_{0,1,2}$, A_0, and B_0 are material parameters whose values were determined by Liu *et al.*. The $\Gamma_{1,2,4,5}$ are 4×4 Dirac matrices, which are defined as

$$\Gamma_1 = \sigma_1 \otimes \sigma_1 \tag{3.5}$$

$$\Gamma_2 = \sigma_2 \otimes \sigma_1 \tag{3.6}$$

$$\Gamma_4 = \mathbb{1} \otimes \sigma_2 \tag{3.7}$$

$$\Gamma_5 = \mathbb{1} \otimes \sigma_3 \tag{3.8}$$

where $\sigma_{1,2,3}$ are the 2×2 Pauli matrices.

3.1.1   Periodizing the Effective Hamiltonian

This Hamiltonian, Eq. 3.2, describes electrons propagating in the momentum space continuum of a 3D TI. Since we want to implement a Hubbard model Hamiltonian, we require a discretized Hamiltonian that can describe an electron propagating through a lattice. This requires the Hamiltonian to be in real space in one direction so that a finite size TI slab[1] can be created. This is also essential for studying SSs. Therefore, Eq. 3.2 must be periodized, and we choose to do so in the z direction. The slab is therefore a 1D lattice in the z direction with infinite momentum planes in the x and y directions. The periodized Hamiltonian has the same periodicity as the lattice. When this periodized Hamiltonian is expanded to quadratic order in $\boldsymbol{k}$, Eq. 3.2 is recovered.

To periodize Eq. 3.2, we let

$$k_z \to \frac{1}{a}\,\sin(k_z a) \tag{3.9}$$

$$k_z^2 \to \frac{2}{a^2}\,(1 - \cos(k_z a)) \tag{3.10}$$

The periodized momentum space Hamiltonian is therefore

$$h'_{\text{eff}}(\boldsymbol{k}) = \epsilon'_{\boldsymbol{k}}\,\mathbb{1} + \mathcal{M}'(\boldsymbol{k})\,\Gamma_5 + \frac{1}{a}\,B_0\,\sin(k_z a)\,\Gamma_4 + A_0\,(k_y\,\Gamma_1 - k_x\,\Gamma_2) \tag{3.11}$$

where

$$\epsilon'_{\boldsymbol{k}} = C_0 + \frac{2}{a^2}C_1\,(1 - \cos(k_z a)) + C_2\,\boldsymbol{k}_{\|}^{\,2} \tag{3.12}$$

$$\mathcal{M}'(\boldsymbol{k}) = M_0 + \frac{2}{a^2}M_1\,(1 - \cos(k_z a)) + M_2\,\boldsymbol{k}_{\|}^{\,2} \tag{3.13}$$

and the prime is used to differentiate the periodized from the un-periodized Hamiltonian. In second quantization formalism, the periodized Hamiltonian in the interaction picture is

$$\hat{\mathcal{H}}' = \sum_{\boldsymbol{k}} \hat{c}^{\dagger}_{\boldsymbol{k}\alpha}\,h'^{\alpha\beta}_{\text{eff}}(\boldsymbol{k})\,\hat{c}_{\boldsymbol{k}\beta} \tag{3.14}$$

---

[1]"Slab" is used to refer to the 3D TI sample in the model.

where $h'^{\alpha\beta}_{\text{eff}}(\boldsymbol{k})$ is Eq. 3.11 and α, β are orbital/spin indices representing the four basis states.

We diagonalize Eq. 3.2 and 3.11 near the Γ point and plot the energy eigenvalues to compare the un-periodized and periodized bulk bandstructures, respectively (see Fig. 3.1). The bandstructures of both Eq. 3.2 and 3.11 are indistinguishable near the Γ point (in fact, in Fig. 3.1a they are identical), which validates our decision to periodize the Hamiltonian.
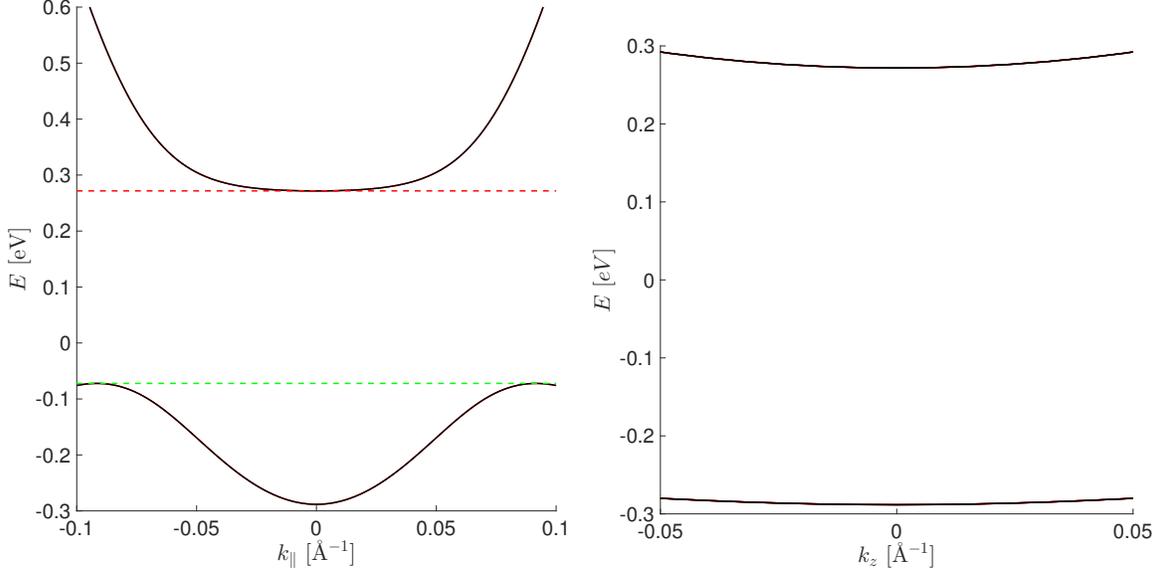
Figure 3.1: Bulk bandstructure of Bi_2Se_3 in the $\boldsymbol{k}_\parallel$ (left) and k_z (right) directions. The Fermi level is at $E = 0$ eV; thus, Eq. 3.2 and 3.11 describe a band insulator as is required for a TI. (Left) The red line denotes the bottom of the conduction band, and the green line denotes the top of the valence band. These figures reproduce results from Ref. [72].

The effective Hamiltonian, Eq. 3.11, describes a bulk slab of a TI that is infinite in all three directions; partially Fourier transforming Eq. 3.14 from k_z to z adds a boundary to the 3D slab and allows the SSs to be studied. This is done by substituting

$$\hat{c}^\dagger_{\boldsymbol{k}\alpha} = \frac{1}{\sqrt{L_z}} \sum_z e^{ik_z z}\, \hat{c}^\dagger_{\boldsymbol{k}_\parallel z\alpha}, \qquad \hat{c}_{\boldsymbol{k}\beta} = \frac{1}{\sqrt{L_z}} \sum_{z'} e^{-ik_z z'}\, \hat{c}_{\boldsymbol{k}_\parallel z'\beta} \tag{3.15}$$

into Eq. 3.14. $L_z = N_z a$ where a is the effective lattice constant and N_z is the number of z lattice sites. We generally refer to z as representing the number of layers of infinite k_x, k_y planes, as opposed to the number of lattice sites. By increasing the number of layers, we increase the thickness of the slab. For the remainder of this thesis $a = 1$ Å. a is not the physical lattice constant, but should be thought of as a discretization parameter that can

be chosen arbitrarily, as long as it is sufficiently small (if a is taken too large, the range of wavevectors for which the effective Hamiltonian is valid decreases).

The periodized effective Hamiltonian for finite z is

$$\hat{\mathcal{H}} = \sum_{\boldsymbol{k}_\parallel} \sum_{z=1}^{N_z} \left[\hat{c}^\dagger_{\boldsymbol{k}_\parallel z \alpha} \, h^0_{\alpha\beta}(\boldsymbol{k}_\parallel) \, \hat{c}_{\boldsymbol{k}_\parallel z \beta} - t^\dagger_{\alpha\beta} \, \hat{c}^\dagger_{\boldsymbol{k}_\parallel, z+1, \alpha} \, \hat{c}_{\boldsymbol{k}_\parallel z \beta} - t_{\alpha\beta} \, \hat{c}^\dagger_{\boldsymbol{k}_\parallel z \alpha} \, \hat{c}_{\boldsymbol{k}_\parallel, z+1, \beta} \right] \qquad (3.16)$$

where the on-site Hamiltonian is

$$h^0(\boldsymbol{k}_\parallel) = (C_0 + 2\,C_1 + C_2\,\boldsymbol{k}_\parallel^2)\,\mathbb{1} + (M_0 + 2\,M_1 + M_2\,\boldsymbol{k}_\parallel^2)\,\Gamma_5 + A_0\,(k_y\,\Gamma_1 - k_x\,\Gamma_2) \qquad (3.17)$$

and the nearest-neighbour hopping term in the z direction is

$$t^\dagger = C_1\,\mathbb{1} + M_1\,\Gamma_5 - \frac{i}{2} B_0\,\Gamma_4 \qquad (3.18)$$

The size of the Hamiltonian matrix depends on the number of layers, N_z. There are N_z $h^0(\boldsymbol{k}_\parallel)$ blocks and $N_z - 1$ $t, t^\dagger$ blocks. The full Hamiltonian matrix is a block tri-diagonal $4N_z \times 4N_z$ matrix. It is important to note that a single QL in a real TI does not have a one-to-one correspondence to a single layer in the Hamiltonian, as this is an effective Hamiltonian, not a microscopic one.

In order to calculate the non-interacting GF using the Finite-N algorithm, Eq. 3.16 must be put in the block tri-diagonal form of Eq. 2.36. This is easy to do as it simply requires letting $H_{00} = h^0(\boldsymbol{k}_\parallel)$, $H_{01} = -t$, and $H_{10} = -t^\dagger$. Diagonalizing Eq. 3.16 and plotting the eigenvalues yields the band structure near the Fermi level. Bandstructure and probability density plots of Bi_2Se_3 are shown in Fig. 3.2 and 3.3, respectively.

In Fig. 3.2a, the DC is gapped out as a result of the hybridization between the top and bottom SSs in a thin slab. This happens when the slab thickness is smaller than the SS penetration depth. Furthermore, the bulk valence and conduction bands are far from their infinite N_z values. Fig. 3.2b shows a more distinct DC, but a gap still exists. However, (though not plotted) the states on the DC are localized to the surface. N_z is large enough such that this can be considered a 3D TI even though the bulk valence and conduction bands

are still far from their infinite N_z values. We see in Fig. 3.2c that the DC gap has essentially closed. Recall that one QL has a thickness of approximately 1 nm [50]; considering our choice of effective lattice constant (1 Å), one QL is roughly equivalent to $N_z = 10$ layers. Thus, our results in this figure match well with experimental data that the DC gap closes at six QLs (see Fig. 1.9). As N_z increases, the bulk valence and conduction bands converge toward their values for an infinitely thick slab, Fig. 3.1. Finally, in Fig. 3.2d, the bulk valence and conduction bands have converged to their infinite N_z values. The bandstructure plots are 2D and thus only a slice of the DC is seen. One can imagine rotating this bandstructure around the k_y states to obtain the 3D DC. In Fig. 3.3, the penetration depth of the SSs is seen to be roughly 20 layers; this is consistent with the observation that the DC is gapped out for $N_z = 20$.

(a) $N_z = 20$ Layers

(b) $N_z = 40$ Layers

(c) $N_z = 60$ Layers

(d) $N_z = 200$ Layers
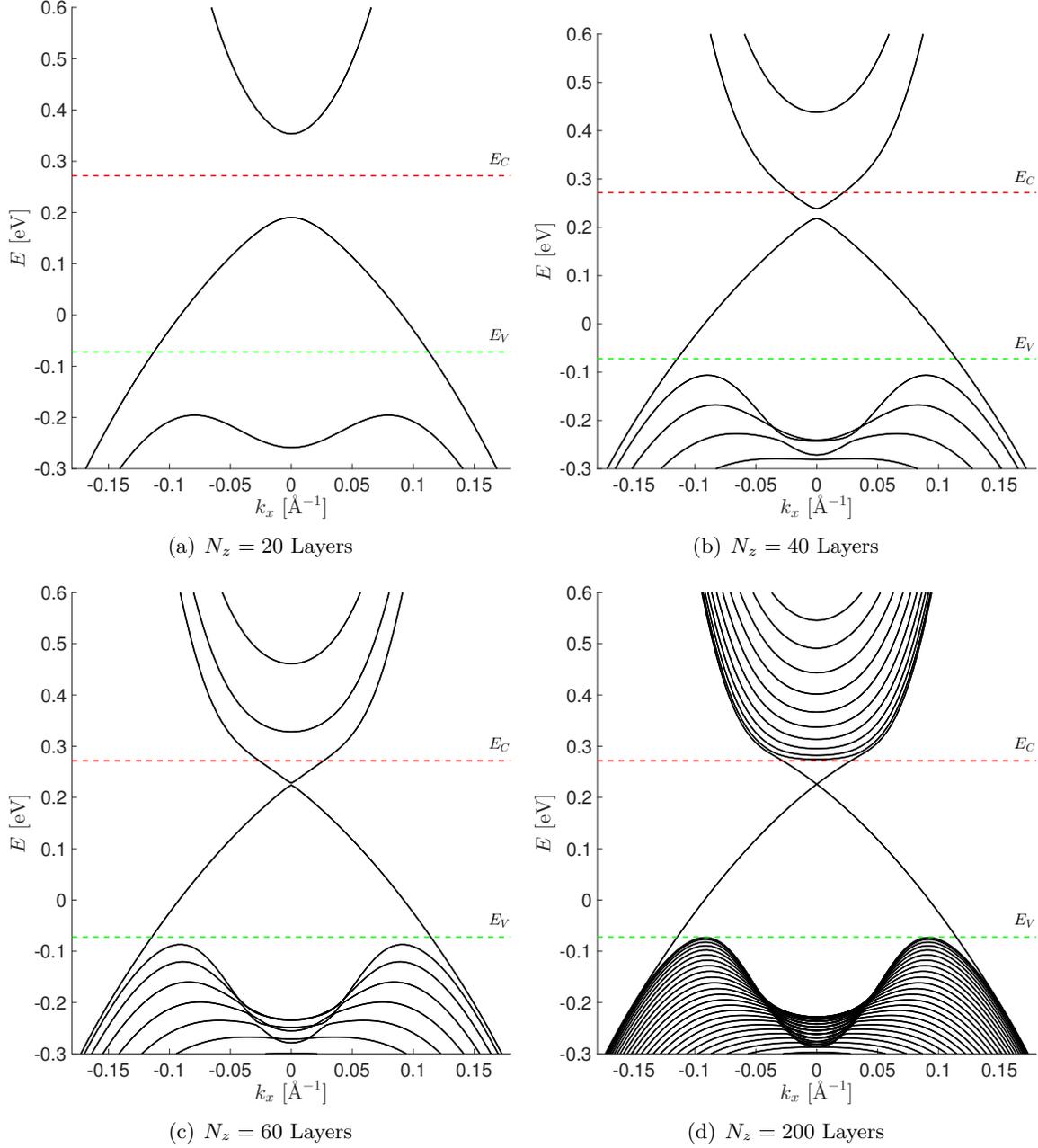
Figure 3.2: The bandstructure near the Fermi level ($E_F = 0$ eV) at $k_y = 0$ Å^{-1} for various thicknesses of Bi$_2$Se$_3$. The top of the valence band is denoted by the green line, E_V, while the bottom of the conduction band is denoted by the red line, E_C. The values for E_V and E_C were taken from their infinite N_z values in Fig. 3.1. Although the Hamiltonian is only quantitatively accurate for $|\boldsymbol{k}| < 0.04$ Å^{-1}, a larger domain is shown so as to see where the SSs connect to the bulk states.
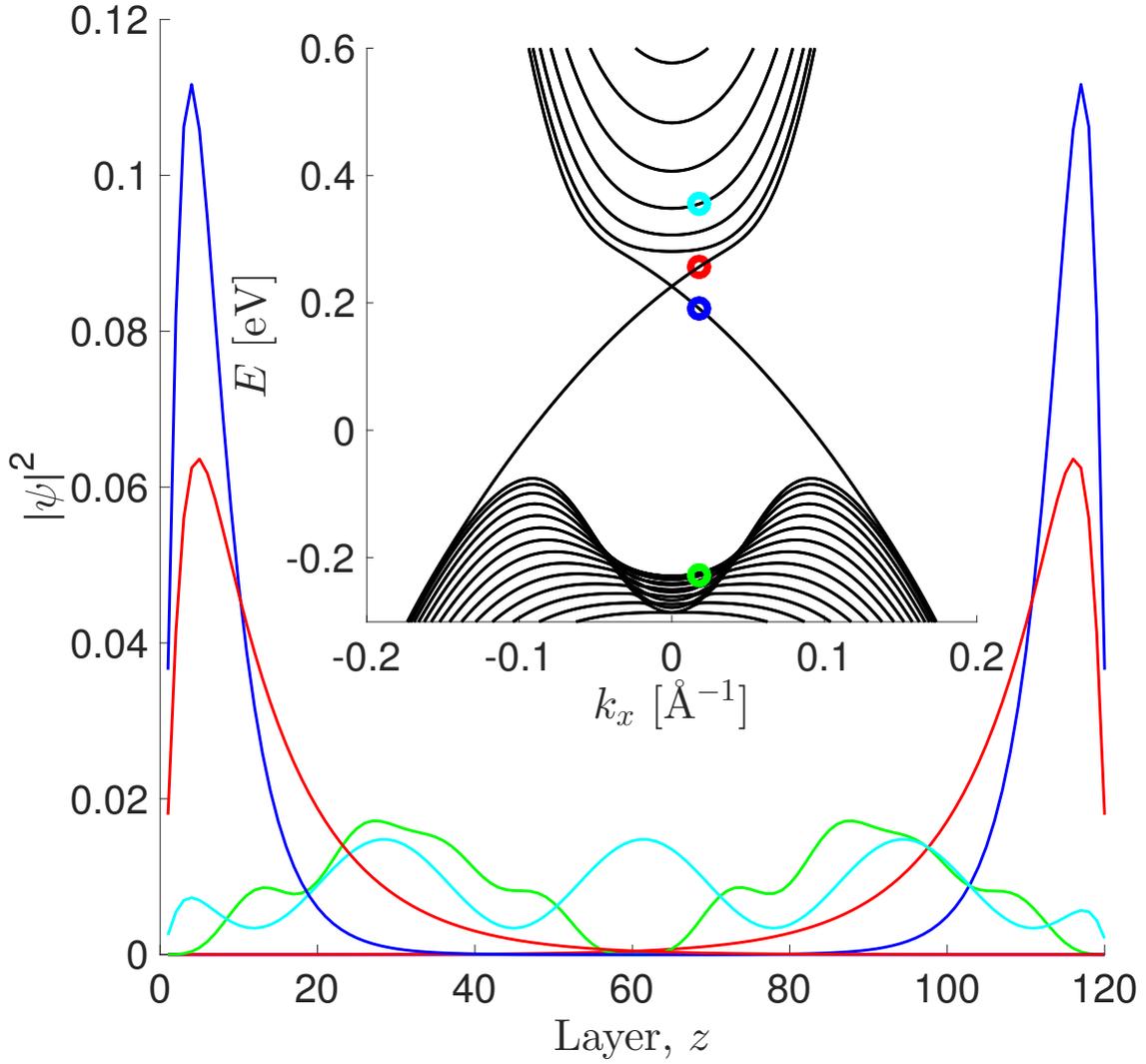
Figure 3.3: The probability densities for valence (green), conduction (light blue), and SSs (blue and red) are plotted for $N_z = 120$ layers of Bi_2Se_3 at $k_x = 0.0125$ Å^{-1}. [Inset]: The bandstructure for $N_z = 120$ layers. Each band is doubly degenerate: one band for the top surface and one band for the bottom surface. For the two circles on the DC, both the top and bottom states are plotted (for a total of four SSs). Conversely, only one of the states circled for the bulk conduction and valence bands are plotted. The plotted states are circled with the same colour as in the inset. The SSs are found on the DC; states in the bulk valence and conduction bands are found throughout the bulk slab.

3.2 Non-interacting Spectral Function and Density of States

As discussed in Ch. 1, two primary means of studying the SSs experimentally are ARPES, which probes the spectral function (see Fig. 1.8 and 1.9) and STM which essentially measures the surface local DOS (see Fig. 1.11 and 1.12). We first calculate the non-interacting surface and bulk GFs using the algorithm outlined in Sec. 2.4. From this, we obtain the non-interacting spectral function and DOS[2] using Eq. 2.33 and 3.19, respectively. Both ARPES and STM probe only the surfaces of the materials of interest, and since the experiments we reviewed used these probes, we focus on quantities calculated on the slab surface. We calculate the surface spectral function from the non-interacting GF by setting $z = z' = 1$ in Eq. 2.33, corresponding to the top layer, and Fourier transforming $x - x'$ and $y - y'$ to k_x and k_y, respectively. We could have just as easily set $z = z' = N_z$ since there is reflection symmetry in the z direction. In our model, the spectral function is a 4×4 matrix in orbital/spin space. Numerically calculating the spectral function and DOS resulted in run times scaling quadratically with N_z, and linearly with the number of k_x, k_y, and ω values.

In Fig. 3.4, we plot the trace of the spectral function, corresponding to the sum over all orbitals/spins, at various layers, z, for a slab of $N_z = 120$ layers. We see that for $z = 4$, the DC states are brightest, and the bulk states are comparatively dim. This agrees with the probability density plot in Fig. 3.3, where the greatest probability of finding an electron in a layer was for $z = 4$. As we examine layers away from $z = 4$, the DC states become dimmer. In addition, for $z = 10$, we see the bulk states becoming brighter. For $z = 30$, the DC is virtually invisible and no SSs exist. Note the similarity in shape between these plots and the bandstructure plots in Fig. 3.2. This is more easily seen in the side-by-side comparison of bandstructure (left panel) and spectral intensity (centre panel) plots in Fig. 3.5.

---

[2]The system is translationally invariant in the x and y directions; therefore, the DOS and local DOS are the same quantity.

(a) $z = 1$

(b) $z = 4$

(c) $z = 10$

(d) $z = 30$
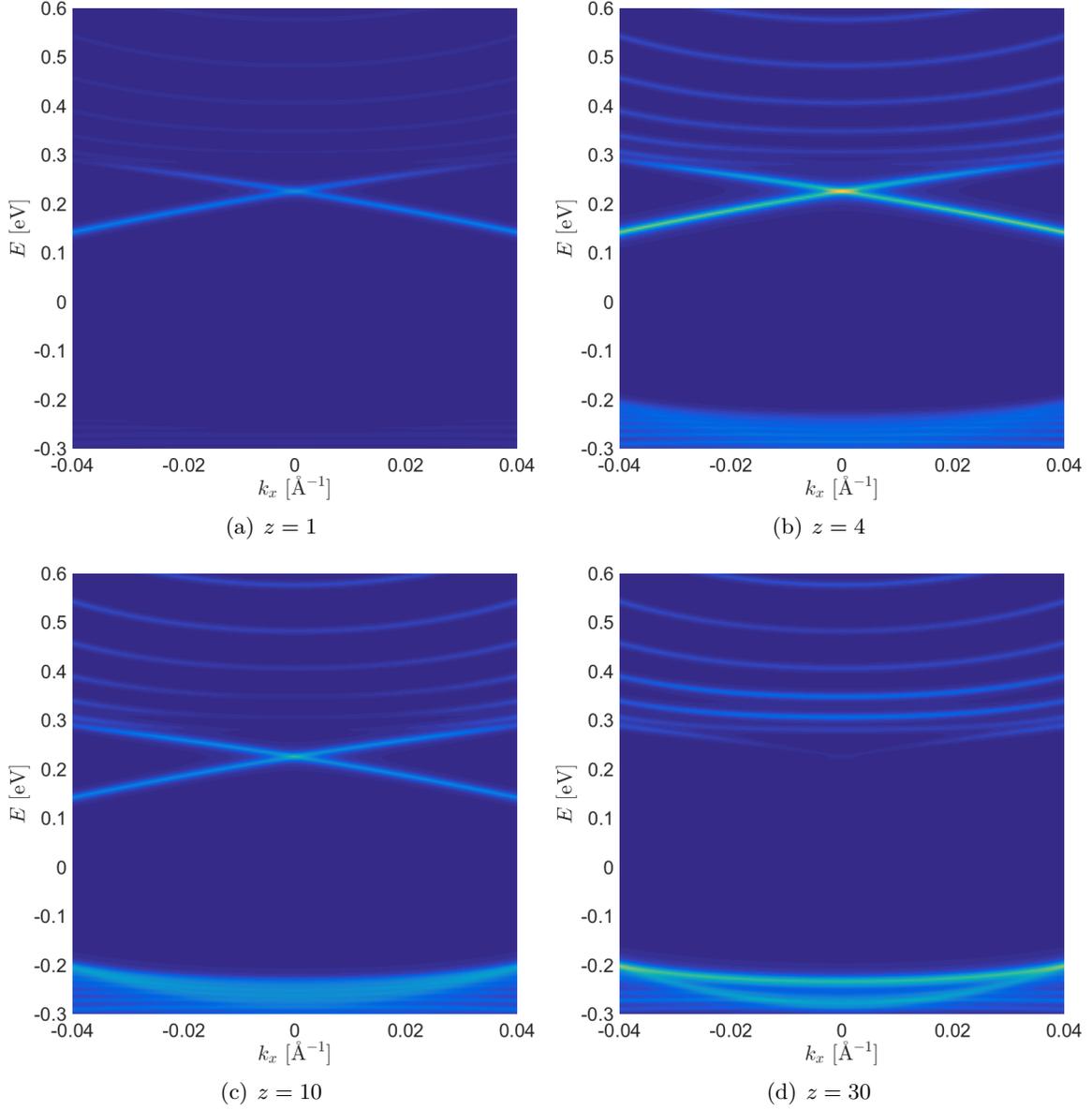
Figure 3.4: Intensity plots of the trace of the spectral function for various layers. The brighter the plot, the more distinguishable are the peaks in the spectral function. These plots are analogous to the ARPES plots in Fig. 1.8 and 1.9.

The surface DOS is

$$\rho_{\text{surface}}(\omega) = \frac{1}{2\pi} \frac{1}{L_x L_y} \sum_{\boldsymbol{k}_\parallel} \sum_{\alpha\beta} A^{\alpha\beta}_{z,z'=1}(\boldsymbol{k}_\parallel, \omega)\, \delta_{\alpha\beta} \tag{3.19}$$

where $L_{x,y} = N_{x,y}\, a$, where N_x, N_y are the number of values summed over in the k_x, k_y directions. This is plotted in the right panel of Fig. 3.5.

47

As mentioned previously, the linear Dirac dispersion on a 2D DC near the DP is

$$E_{\boldsymbol{k}} = \pm v_F |\boldsymbol{k}| \tag{3.20}$$

where v_F is the Fermi velocity. The DOS is

$$\rho(\omega) = \int \frac{d^2\boldsymbol{k}}{(2\pi)^2} \delta(\omega - E_{\boldsymbol{k}}) \tag{3.21}$$

and switching to polar coordinates gives

$$\rho(\omega) = \int_0^{2\pi} \frac{d\theta}{2\pi} \int_0^\infty \frac{dk}{2\pi} k \, \delta(\omega - E_{\boldsymbol{k}}) \tag{3.22}$$

The dispersion only depends on the magnitude of $\boldsymbol{k}$, and thus $\int \frac{d\theta}{2\pi} = 1$. Substituting in Eq. 3.20 yields

$$\rho(\omega) = \int_0^\infty \frac{dk}{2\pi} k \, \delta(\omega \mp v_F k) \tag{3.23}$$

The δ function yields $k = |\omega|/v_F$, and the DOS is therefore

$$\rho(\omega) = \frac{1}{(2\pi)^2} \frac{|\omega|}{v_F} \tag{3.24}$$

which is linear in ω. More specifically, a Dirac DOS resembles an absolute value function: a V-shape, with a minimum at $\omega = 0$. Recall that a near linear differential conductance is observed in Fig. 1.11 near the DP.

We see a roughly linear surface DOS in the right panel of Fig. 3.5 near the DP ($0.15 \leq \omega \leq 0.26$ eV). The DOS was calculated for $-0.18 < k_x < 0.18$ Å^{-1}, which is larger than the k_x domain plotted here, so as to include all states in the energy range $-0.3 < \omega < 0.6$ eV. The rapid increase in the DOS below the green line is due to contributions from the bulk valence states. This was also seen experimentally, as shown in Fig. 1.11. The contribution from the conduction band is not as large, which can be understood from the fact that the conduction band has a lower effective mass than the valence band (as can be seen in Fig. 3.1

and 3.2). This also explains why experimentally (Fig. 1.11) the rise in tunnelling differential conductance is faster below the DP than above it.
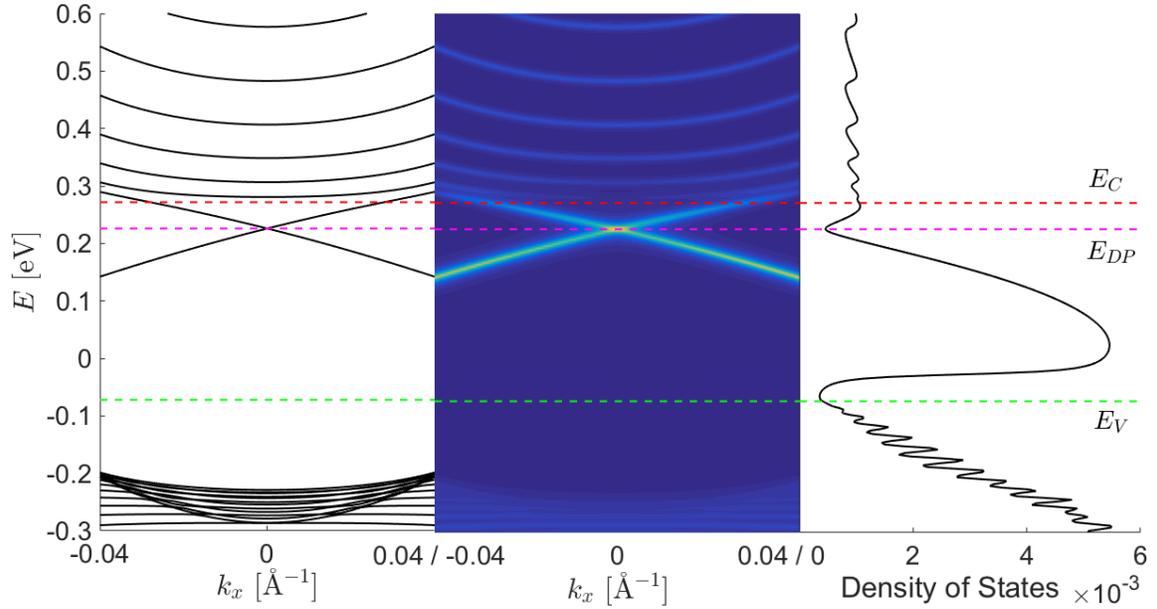
Figure 3.5: (Left) Bandstructure for $N_z = 120$ layers. (Centre) Intensity plot of the trace of the surface spectral function. (Right) Surface DOS. The green line denotes the top of the bulk valence band, the magenta line denotes the location of the DP, and the red line denotes the bottom of the bulk conduction band.

3.3 First Order Self-Energy

Sec. 3.1 and 3.2 only considered the non-interacting part of the Hamiltonian; we now focus on the interacting part, $\hat{V}$, so as to include electron-electron interactions in our model. This allows us to examine how interactions affect the quasiparticle lifetime. In this section (and for the rest of this chapter), we assume the interactions are small so that we can apply perturbation theory. As described in Sec. 2.2.2, the screened Coulomb interaction can be modelled as a delta function contact interaction. In the calculation of the self-energy we use the full GF (including both surface and bulk GFs) since both surface-to-surface and surface-to-bulk scattering processes are in principle possible, as discussed in Ch. 1. We begin by calculating the self-energy to first order in perturbation theory [63, 65].

Since the effective Hamiltonian, Eq. 3.16, is in mixed momentum, $\boldsymbol{k}_\parallel$, and position, z, space, our GFs must also have these indices. First, we expand Eq. 2.22 to first order, and

set $\tau' = 0$:

$$\mathcal{G}_{zz'}^{\alpha\beta}(\boldsymbol{k}_\parallel, \boldsymbol{k}'_\parallel, \tau) = -\left\langle T_\tau \left[\hat{c}_{\boldsymbol{k}_\parallel z\alpha}(\tau) \hat{c}_{\boldsymbol{k}'_\parallel z'\beta}^\dagger(0) \right] \right\rangle_0 + \int_0^\beta d\tau_1 \left\langle T_\tau \left[\hat{V}(\tau_1) \hat{c}_{\boldsymbol{k}_\parallel z\alpha}(\tau) \hat{c}_{\boldsymbol{k}'_\parallel z'\beta}^\dagger(0) \right] \right\rangle_0$$

$$(3.25)$$

where from Eq. 2.12 we see that the first term (the zeroth order term) is the free particle GF, $\mathcal{G}_{zz'}^{(0)\alpha\beta}(\boldsymbol{k}_\parallel, \tau)$. We ignore this term as we are concerned with obtaining the first order Feynman diagrams. In the second term,

$$\hat{V}(\tau_1) = U_{\gamma\delta} \sum_{\boldsymbol{r}} \hat{c}_{\boldsymbol{r}\gamma}^\dagger(\tau_1) \hat{c}_{\boldsymbol{r}\gamma}(\tau_1) \hat{c}_{\boldsymbol{r}\delta}^\dagger(\tau_1) \hat{c}_{\boldsymbol{r}\delta}(\tau_1) \tag{3.26}$$

where γ, δ are orbital/spin indices. Here, and below, there is an implied sum over γ and δ, even though these indices are repeated more than twice. We write $\sum_{\boldsymbol{r}} \to \int d^2\boldsymbol{r}_\parallel \sum_{z_1}$ since we are continuous in the x, y directions and discrete in z. Next, we Fourier transform Eq. 3.26 from $\boldsymbol{r}_\parallel \to \boldsymbol{p}_\parallel$, where $\boldsymbol{p}_\parallel$ is a momentum, to obtain the same basis as Eq. 3.25. Here, $\boldsymbol{p}_\parallel$ is a discrete quantity as it will be calculated numerically, but we sum over enough values such that it is a good approximation to an integral. In other words, we are concerned with a system that is finite in the planar directions, L_x and L_y. The Fourier transform is performed by substituting Eq. 3.27 into Eq. 3.26:

$$\hat{c}_{\boldsymbol{r}\gamma}^\dagger(\tau_1) = \frac{1}{\sqrt{L_x L_y}} \sum_{\boldsymbol{p}_\parallel} e^{-i\boldsymbol{p}_\parallel \cdot \boldsymbol{r}_\parallel} \hat{c}_{\boldsymbol{p}_\parallel z_1\gamma}^\dagger(\tau_1), \qquad \hat{c}_{\boldsymbol{r}\gamma}(\tau_1) = \frac{1}{\sqrt{L_x L_y}} \sum_{\boldsymbol{p}'_\parallel} e^{i\boldsymbol{p}'_\parallel \cdot \boldsymbol{r}_\parallel} \hat{c}_{\boldsymbol{p}'_\parallel z_1\gamma}(\tau_1)$$

$$(3.27)$$

We now write 3.26 in its Fourier transformed form as

$$\hat{V}(\tau_1) = \frac{U_{\gamma\delta}}{(L_x L_y)^2} \int d^2\boldsymbol{r}_\parallel \sum_{z_1} \sum_{\boldsymbol{p}_\parallel \boldsymbol{p}'_\parallel \boldsymbol{q}_\parallel \boldsymbol{q}'_\parallel}$$

$$\times \hat{c}_{\boldsymbol{p}'_\parallel z_1\gamma}^\dagger(\tau_1) \hat{c}_{\boldsymbol{p}_\parallel z_1\gamma}(\tau_1) \hat{c}_{\boldsymbol{q}'_\parallel z_1\delta}^\dagger(\tau_1) \hat{c}_{\boldsymbol{q}_\parallel z_1\delta}(\tau_1) e^{-i(\boldsymbol{p}'_\parallel - \boldsymbol{p}_\parallel + \boldsymbol{q}'_\parallel - \boldsymbol{q}_\parallel) \cdot \boldsymbol{r}_\parallel} \quad (3.28)$$

Evaluating $\int d^2\boldsymbol{r}_\parallel$ yields $L_x L_y \, \delta_{-\boldsymbol{p}'_\parallel + \boldsymbol{p}_\parallel - \boldsymbol{q}'_\parallel + \boldsymbol{q}_\parallel, 0}$. This Kronecker delta function ensures that conservation of momentum is obeyed and allows one of the momentum sums to be removed.

We also make a change of variable, which gives

$$\hat{V}(\tau_1) = \frac{U_{\gamma\delta}}{L_x L_y} \sum_{z_1} \sum_{pp'q} \hat{c}^\dagger_{p'+q,z_1,\gamma}(\tau_1)\, \hat{c}_{p',z_1,\gamma}(\tau_1)\, \hat{c}^\dagger_{p-q,z_1,\delta}(\tau_1)\, \hat{c}_{p,z_1,\delta}(\tau_1) \qquad (3.29)$$

where we have dropped the $\parallel$ subscript as it is understood that all momenta are only in the x, y directions.
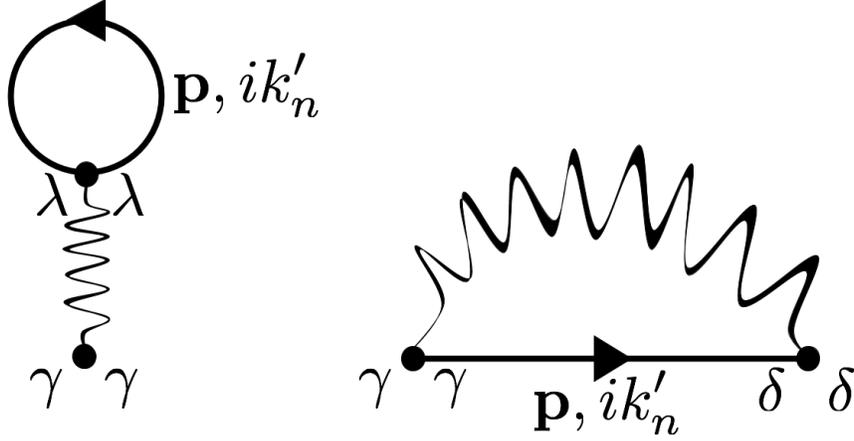
Figure 3.6: The Hartree (left) and Fock (right) Feynman diagrams with labelled orbital/spin indices, momenta, and Matsubara frequencies. These are the two connected first order diagrams.

We now apply Wick's theorem to the second term in 3.25, where $\hat{V}(\tau_1)$ takes the form of Eq. 3.29. There are two non-trivial connected diagrams: the Hartree diagram and the Fock diagram. These are shown in Fig. 3.6. Each of these diagrams appears twice after using Wick's theorem and thus carry a prefactor of ± 2. (Whether a diagram carries a plus or minus sign arises from how the fermionic operators are paired up in Wick's theorem. The factor of two can be obtained by swapping the orbital/spin indices, $\gamma \leftrightarrow \delta$. This does not change the topology of the diagrams.) We also now write the expectation value of the product of two creation/annihilation operators as Matsubara GFs (see Eq. 2.12). The first

order correction to the GF becomes

$$\mathcal{G}_{zz'}^{(1)\alpha\beta}(\boldsymbol{k},\boldsymbol{k}',\tau) = 2\frac{U_{\gamma\delta}}{L_x L_y}\int_0^\beta d\tau_1 \sum_{z_1}\sum_{\boldsymbol{p}'\boldsymbol{pq}}$$

$$\times\,[\mathcal{G}_{zz_1}^{(0)\alpha\gamma}(\boldsymbol{k},\boldsymbol{p}'+\boldsymbol{q},\tau-\tau_1)\,\mathcal{G}_{z_1z_1}^{(0)\delta\delta}(\boldsymbol{p},\boldsymbol{p}-\boldsymbol{q},0)\,\mathcal{G}_{z_1z'}^{(0)\gamma\beta}(\boldsymbol{p}',\boldsymbol{k}',\tau_1)$$

$$-\,\mathcal{G}_{zz_1}^{(0)\alpha\gamma}(\boldsymbol{k},\boldsymbol{p}'+\boldsymbol{q},\tau-\tau_1)\,\mathcal{G}_{z_1z_1}^{(0)\gamma\delta}(\boldsymbol{p}',\boldsymbol{p}-\boldsymbol{q},0)\,\mathcal{G}_{z_1z'}^{(0)\delta\beta}(\boldsymbol{p},\boldsymbol{k}',\tau_1)] \quad (3.30)$$

where the first term is the Hartree diagram and the second term is the Fock diagram. Taking advantage of the translational invariance in our system (recall that Eq. 3.16 is diagonal in $\boldsymbol{k}$), such that

$$\mathcal{G}_{zz'}^{\alpha\beta}(\boldsymbol{k},\boldsymbol{k}',\tau-\tau') = \mathcal{G}_{zz'}^{\alpha\beta}(\boldsymbol{k},\tau-\tau')\,\delta_{\boldsymbol{kk}'} \qquad (3.31)$$

and evaluating the Kronecker delta functions, Eq. 3.30 becomes

$$\mathcal{G}_{zz'}^{(1)\alpha\beta}(\boldsymbol{k},\tau) = 2\frac{U_{\gamma\delta}}{L_x L_y}\int_0^\beta d\tau_1 \sum_{z_1}\sum_{\boldsymbol{p}}$$

$$\times\,[\mathcal{G}_{zz_1}^{(0)\alpha\gamma}(\boldsymbol{k},\tau-\tau_1)\,\mathcal{G}_{z_1z_1}^{(0)\delta\delta}(\boldsymbol{p},0)\,\mathcal{G}_{z_1z'}^{(0)\gamma\beta}(\boldsymbol{k},\tau_1)$$

$$-\,\mathcal{G}_{zz_1}^{(0)\alpha\gamma}(\boldsymbol{k},\tau-\tau_1)\,\mathcal{G}_{z_1z_1}^{(0)\gamma\delta}(\boldsymbol{p},0)\,\mathcal{G}_{z_1z'}^{(0)\delta\beta}(\boldsymbol{k},\tau_1)] \quad (3.32)$$

Fourier transforming from τ to ik_n, we replace each of the GFs in Eq. 3.32 with

$$\mathcal{G}_{zz'}^{\alpha\beta}(\boldsymbol{k},\tau) = \frac{1}{\beta}\sum_{ik_n}e^{-ik_n\tau}\,\mathcal{G}_{zz'}^{\alpha\beta}(\boldsymbol{k},ik_n) \qquad (3.33)$$

which yields

$$\frac{1}{\beta}\sum_{ik_n}e^{-ik_n\tau}\,\mathcal{G}_{zz'}^{(1)\alpha\beta}(\boldsymbol{k},ik_n) = \frac{2}{\beta^3}\frac{U_{\gamma\delta}}{L_x L_y}\sum_{z_1}\sum_{\boldsymbol{p}}\sum_{ik_nik_n'ik_n''}e^{-ik_n\tau}\int_0^\beta d\tau_1\,e^{(ik_n-ik_n'')\tau_1}$$

$$\times\,[\mathcal{G}_{zz_1}^{(0)\alpha\gamma}(\boldsymbol{k},ik_n)\,\mathcal{G}_{z_1z_1}^{(0)\delta\delta}(\boldsymbol{p},ik_n')\,\mathcal{G}_{z_1z'}^{(0)\gamma\beta}(\boldsymbol{k},ik_n'')$$

$$-\,\mathcal{G}_{zz_1}^{(0)\alpha\gamma}(\boldsymbol{k},ik_n)\,\mathcal{G}_{z_1z_1}^{(0)\gamma\delta}(\boldsymbol{p},ik_n')\,\mathcal{G}_{z_1z'}^{(0)\delta\beta}(\boldsymbol{k},ik_n'')] \quad (3.34)$$

where ik_n, ik'_n, ik''_n are fermionic Matsubara frequencies. Evaluating $\int d\tau_1$, we obtain

$$\frac{1}{\beta} \sum_{ik_n} e^{-ik_n\tau} \mathcal{G}^{(1)\alpha\beta}_{zz'}(\boldsymbol{k}, ik_n) = \frac{2}{\beta^2} \frac{U_{\gamma\delta}}{L_x L_y} \sum_{z_1} \sum_{\boldsymbol{p}} \sum_{ik_n ik'_n} e^{-ik_n\tau}$$

$$\times \left[\mathcal{G}^{(0)\alpha\gamma}_{zz_1}(\boldsymbol{k}, ik_n)\, \mathcal{G}^{(0)\delta\delta}_{z_1 z_1}(\boldsymbol{p}, ik'_n)\, \mathcal{G}^{(0)\gamma\beta}_{z_1 z'}(\boldsymbol{k}, ik_n) \right.$$

$$\left. - \mathcal{G}^{(0)\alpha\gamma}_{zz_1}(\boldsymbol{k}, ik_n)\, \mathcal{G}^{(0)\gamma\delta}_{z_1 z_1}(\boldsymbol{p}, ik'_n)\, \mathcal{G}^{(0)\delta\beta}_{z_1 z'}(\boldsymbol{k}, ik_n) \right] \quad (3.35)$$

Since this expression must be true for all τ, we have

$$\mathcal{G}^{(1)\alpha\beta}_{zz'}(\boldsymbol{k}, ik_n) = \frac{2}{\beta} \frac{U_{\gamma\delta}}{L_x L_y} \sum_{z_1} \sum_{\boldsymbol{p}} \sum_{ik'_n}$$

$$\times \left[\mathcal{G}^{(0)\alpha\gamma}_{zz_1}(\boldsymbol{k}, ik_n)\, \mathcal{G}^{(0)\delta\delta}_{z_1 z_1}(\boldsymbol{p}, ik'_n)\, \mathcal{G}^{(0)\gamma\beta}_{z_1 z'}(\boldsymbol{k}, ik_n) \right.$$

$$\left. - \mathcal{G}^{(0)\alpha\gamma}_{zz_1}(\boldsymbol{k}, ik_n)\, \mathcal{G}^{(0)\gamma\delta}_{z_1 z_1}(\boldsymbol{p}, ik'_n)\, \mathcal{G}^{(0)\delta\beta}_{z_1 z'}(\boldsymbol{k}, ik_n) \right] \quad (3.36)$$

In order to extract the self-energy from this equation we need to compare it to Dyson's equation, which to first order is

$$\mathcal{G}^{(1)\alpha\beta}_{zz'}(\boldsymbol{k}, ik_n) = \mathcal{G}^{(0)\alpha\beta}_{zz'}(\boldsymbol{k}, ik_n) + \sum_{z_1} \mathcal{G}^{(0)\alpha\gamma}_{zz_1}(\boldsymbol{k}, ik_n)\, \Sigma^{\gamma\delta}_{z_1 z_1}(\boldsymbol{k}, ik_n)\, \mathcal{G}^{(0)\delta\beta}_{z_1 z'}(\boldsymbol{k}, ik_n) \quad (3.37)$$

Writing Eq. 3.36 as

$$\mathcal{G}^{(1)\alpha\beta}_{zz'}(\boldsymbol{k}, ik_n) = \sum_{z_1} \mathcal{G}^{(0)\alpha\gamma}_{zz_1}(\boldsymbol{k}, ik_n)$$

$$\times \frac{2}{\beta} \frac{1}{L_x L_y} \sum_{\boldsymbol{p}} \sum_{ik'_n} \left[\delta_{\gamma\delta}\, U_{\gamma\lambda}\, \mathcal{G}^{(0)\lambda\lambda}_{z_1 z_1}(\boldsymbol{p}, ik'_n) - U_{\gamma\delta}\, \mathcal{G}^{(0)\gamma\delta}_{z_1 z_1}(\boldsymbol{p}, ik'_n) \right]$$

$$\times \mathcal{G}^{(0)\delta\beta}_{z_1 z'}(\boldsymbol{k}, ik_n) \quad (3.38)$$

we see that the first order self-energy is the expression on the second line. We add a factor of $e^{ik'_n\eta}$, where η is a positive infinitesimal, to each term as required by Feynman rules since the first term forms a closed loop and the second term has a GF linked by the same

interaction line [63], and thus obtain

$$\Sigma_{z_1 z_1}^{(1)\gamma\delta} = \frac{2}{\beta} \frac{1}{L_x L_y} \sum_{\boldsymbol{p}} \sum_{ik_n'} \left[\delta_{\gamma\delta} \, U_{\gamma\lambda} \, \mathcal{G}_{z_1 z_1}^{(0)\lambda\lambda}(\boldsymbol{p}, ik_n') - U_{\gamma\delta} \, \mathcal{G}_{z_1 z_1}^{(0)\gamma\delta}(\boldsymbol{p}, ik_n') \right] e^{ik_n'\eta} \qquad (3.39)$$

which is independent of $\boldsymbol{k}$ and ik_n.

Writing the GFs in terms of the non-interacting spectral function, $A^{(0)}$,

$$\mathcal{G}_{z_1 z_1}^{(0)\alpha\beta}(\boldsymbol{k}, ik_n) = \int \frac{d\omega}{2\pi} \frac{A_{z_1 z_1}^{(0)\alpha\beta}(\boldsymbol{k}, \omega)}{ik_n - \omega} \qquad (3.40)$$

where ω is the energy of the quasiparticle described by the GF, and then evaluating the Matsubara sum, $\frac{1}{\beta} \sum_{ik_n'}$, yields

$$\Sigma_{z_1 z_1}^{(1)\gamma\delta} = \frac{2}{L_x L_y} \sum_{\boldsymbol{p}} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \left[\delta_{\gamma\delta} \, U_{\gamma\lambda} \, A_{z_1 z_1}^{(0)\lambda\lambda}(\boldsymbol{p}, \omega) - U_{\gamma\delta} \, A_{z_1 z_1}^{(0)\gamma\delta}(\boldsymbol{p}, \omega) \right] n_F(\omega) \qquad (3.41)$$

where we have taken $\lim_{\eta \to 0}$ and n_F is the Fermi function. Analytically continuing Eq. 3.41, as was shown in Eq. 2.26, yields the same result as Eq. 3.41. This is obvious due to the lack of an ik_n dependence.

We now need to determine if Eq. 3.41 is Hermitian. If it is Hermitian it yields an infinite quasiparticle lifetime. To calculate the Hermitian conjugate of Eq. 3.41, we take the complex conjugate and swap the orbital/spin indices:

$$\left(\Sigma_{z_1 z_1}^{(1)\dagger} \right)^{\gamma\delta} = \frac{2}{L_x L_y} \sum_{\boldsymbol{p}} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \left[\delta_{\delta\gamma} \, U_{\delta\lambda}^* \left(A_{z_1 z_1}^{(0)\lambda\lambda}(\boldsymbol{p}, \omega) \right)^* - U_{\delta\gamma}^* \left(A_{z_1 z_1}^{(0)\delta\gamma}(\boldsymbol{p}, \omega) \right)^* \right] n_F(\omega)$$

$$(3.42)$$

$U_{\gamma\delta} = U_{\delta\gamma}^*$ and $A_{z_1 z_1}^{(0)\gamma\delta} = \left(A_{z_1 z_1}^{(0)\delta\gamma} \right)^*$ as they are both Hermitian matrices (in fact, U is real and symmetric). Since the Kronecker delta is symmetric $\delta_{\delta\gamma} = \delta_{\gamma\delta}$, we thus obtain

$$\left(\Sigma_{z_1 z_1}^{(1)\dagger} \right)^{\gamma\delta} = \frac{2}{L_x L_y} \sum_{\boldsymbol{p}} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \left[\delta_{\gamma\delta} \, U_{\gamma\lambda} \, A_{z_1 z_1}^{(0)\lambda\lambda}(\boldsymbol{p}, \omega) - U_{\gamma\delta} \, A_{z_1 z_1}^{(0)\gamma\delta}(\boldsymbol{p}, \omega) \right] n_F(\omega) \qquad (3.43)$$

which is the same as Eq. 3.41. Thus, $\Sigma_{z_1 z_1}^{(1)} = \Sigma_{z_1 z_1}^{(1)\dagger}$ and the quasiparticle lifetime remains unchanged from its non-interacting result. The first order self-energy does give us the renor-

malized energy levels that arise from interactions, but we are not concerned with such re-
sults. Now, we must go to second order in perturbation theory to obtain a finite quasiparticle
lifetime.

3.4 Second Order Self-Energy

To obtain the second order diagrams, we follow the same approach as that used in Sec. 3.3.
Expanding Eq. 2.22 to second order, and setting $\tau' = 0$, the second order correction to the
GF is

$$\mathcal{G}_{zz'}^{(2)\alpha\beta}(\boldsymbol{k}_\|, \tau) = \frac{-1}{2!} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \left\langle T_\tau \left[\hat{V}(\tau_1)\hat{V}(\tau_2) \hat{c}_{\boldsymbol{k}_\| z\alpha}(\tau) \hat{c}_{\boldsymbol{k}_\| z'\beta}^\dagger(0) \right] \right\rangle_0 \qquad (3.44)$$

where

$$\hat{V}(\tau_1) = U_{\gamma\delta} \sum_{\boldsymbol{r}} \hat{c}_{\boldsymbol{r}\gamma}^\dagger(\tau_1) \hat{c}_{\boldsymbol{r}\gamma}(\tau_1) \hat{c}_{\boldsymbol{r}\delta}^\dagger(\tau_1) \hat{c}_{\boldsymbol{r}\delta}(\tau_1) \qquad (3.45)$$

$$\hat{V}(\tau_2) = U_{\mu\nu} \sum_{\boldsymbol{r}'} \hat{c}_{\boldsymbol{r}'\mu}^\dagger(\tau_2) \hat{c}_{\boldsymbol{r}'\mu}(\tau_2) \hat{c}_{\boldsymbol{r}'\nu}^\dagger(\tau_2) \hat{c}_{\boldsymbol{r}'\nu}(\tau_2) \qquad (3.46)$$

and μ, ν are orbital/spin indices. Writing $\sum_{\boldsymbol{r}} \to \int d^2\boldsymbol{r}_\| \sum_{z_1}$ and $\sum_{\boldsymbol{r}'} \to \int d^2\boldsymbol{r}_\|' \sum_{z_2}$, using
Eq. 3.27 to Fourier transform Eq. 3.45 and 3.46 from $\boldsymbol{r}_\| \to \boldsymbol{p}_\|$ and $\boldsymbol{r}_\|' \to \boldsymbol{p}_\|'$, respectively,
and substituting these results into Eq. 3.44, gives

$$\mathcal{G}_{zz'}^{(2)\alpha\beta}(\boldsymbol{k}_\|, \tau) = \frac{-1}{2!} \frac{1}{(L_x L_y)^4} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2$$

$$\times \left\langle T_\tau \left[U_{\gamma\delta} \int d^2\boldsymbol{r}_\| \sum_{z_1} \sum_{\boldsymbol{p}_\| \boldsymbol{p}_\|' \boldsymbol{q}_\| \boldsymbol{q}_\|'} \hat{c}_{\boldsymbol{p}_\|' z_1 \gamma}^\dagger(\tau_1) \hat{c}_{\boldsymbol{p}_\| z_1 \gamma}(\tau_1) \hat{c}_{\boldsymbol{q}_\|' z_1 \delta}^\dagger(\tau_1) \hat{c}_{\boldsymbol{q}_\| z_1 \delta}(\tau_1) \, e^{-i(\boldsymbol{p}_\|' - \boldsymbol{p}_\| + \boldsymbol{q}_\|' - \boldsymbol{q}_\|) \cdot \boldsymbol{r}_\|} \right.$$

$$\times U_{\mu\nu} \int d^2\boldsymbol{r}_\|' \sum_{z_2} \sum_{\boldsymbol{l}_\| \boldsymbol{l}_\|' \boldsymbol{m}_\| \boldsymbol{m}_\|'} \hat{c}_{\boldsymbol{l}_\|' z_2 \mu}^\dagger(\tau_2) \hat{c}_{\boldsymbol{l}_\| z_2 \mu}(\tau_2) \hat{c}_{\boldsymbol{m}_\|' z_2 \nu}^\dagger(\tau_2) \hat{c}_{\boldsymbol{m}_\| z_2 \nu}(\tau_2) \, e^{-i(\boldsymbol{l}_\|' - \boldsymbol{l}_\| + \boldsymbol{m}_\|' - \boldsymbol{m}_\|) \cdot \boldsymbol{r}_\|'}$$

$$\left. \times \hat{c}_{\boldsymbol{k}_\| z\alpha}(\tau) \hat{c}_{\boldsymbol{k}_\| z'\beta}^\dagger(0) \right] \right\rangle_0 \qquad (3.47)$$

Integrating $\int d^2 \boldsymbol{r}_\parallel \int d^2 \boldsymbol{r}'_\parallel$, evaluating the Kronecker delta functions that result, making a change of variable, and dropping the $\parallel$ subscripts as done before, we obtain

$$
\mathcal{G}_{zz'}^{(2)\alpha\beta}(\boldsymbol{k}, \tau) = \frac{-1}{2!} \frac{1}{(L_x L_y)^2} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2
$$
$$
\times \left\langle T_\tau \left[U_{\gamma\delta} \sum_{z_1} \sum_{\boldsymbol{pp'q}} \hat{c}_{\boldsymbol{p'+q},z_1,\gamma}^\dagger(\tau_1) \hat{c}_{\boldsymbol{p'},z_1,\gamma}(\tau_1) \hat{c}_{\boldsymbol{p-q},z_1,\delta}^\dagger(\tau_1) \hat{c}_{\boldsymbol{p},z_1,\delta}(\tau_1) \right.\right.
$$
$$
\times U_{\mu\nu} \sum_{z_2} \sum_{\boldsymbol{ll'm}} \hat{c}_{\boldsymbol{l'+m},z_2,\mu}^\dagger(\tau_2) \hat{c}_{\boldsymbol{l'},z_2,\mu}(\tau_2) \hat{c}_{\boldsymbol{l-m},z_2,\nu}^\dagger(\tau_2) \hat{c}_{\boldsymbol{l},z_2,\nu}(\tau_2)
$$
$$
\left.\left. \times \hat{c}_{\boldsymbol{k}z\alpha}(\tau) \hat{c}_{\boldsymbol{k}z'\beta}^\dagger(0) \right] \right\rangle_0 \quad (3.48)
$$

We apply Wick's theorem to Eq. 3.48 and find six connected diagrams. Each of these diagrams appears eight times (there are eight different ways to permute the four orbital/spin indices) and thus each of these diagrams has a prefactor of ± 8. Four of these diagrams are Hermitian and thus, like the first order diagrams, do not contribute to a finite quasiparticle lifetime. The two non-Hermitian diagrams are the pair-bubble diagram and what we call the double exchange diagram (see Fig. 3.7). We examine these in the following sections.

(a) Pair-Bubble Diagram   (b) Double Exchange Diagram
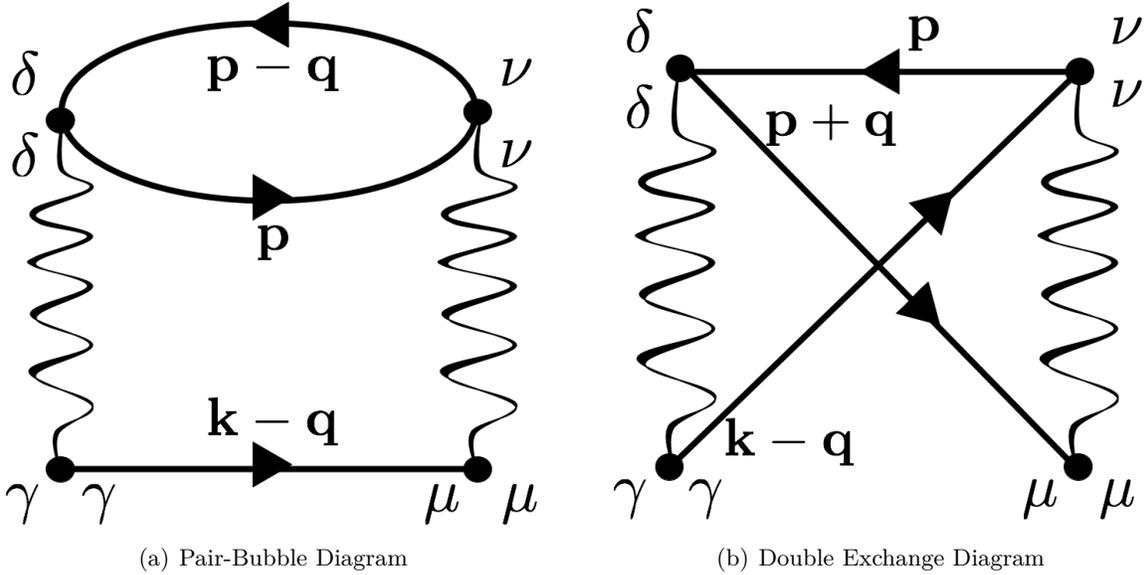
Figure 3.7: The pair-bubble (a) and double exchange (b) Feynman diagrams with labelled orbital/spin indices and momenta. These are the two non-Hermitian second order diagrams.

3.4.1 Pair-Bubble Diagram

We begin by examining the pair-bubble diagram, which is given by

$$
\mathcal{G}_{zz'}^{(2a)\alpha\beta}(\boldsymbol{k}, \tau) = \frac{-8}{2!} \frac{U_{\gamma\delta} U_{\mu\nu}}{(L_x L_y)^2} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \sum_{z_1 z_2} \sum_{\boldsymbol{pp'q}} \sum_{\boldsymbol{ll'm}}
$$

$$
\times \left[\mathcal{G}_{zz_1}^{(0)\alpha\gamma}(\boldsymbol{k}, \boldsymbol{p'} + \boldsymbol{q}, \tau - \tau_1) \, \mathcal{G}_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{p'}, \boldsymbol{l'} + \boldsymbol{m}, \tau_1 - \tau_2) \, \mathcal{G}_{z_2 z'}^{(0)\mu\beta}(\boldsymbol{l'}, \boldsymbol{k}, \tau_2) \right.
$$

$$
\left. \times \mathcal{G}_{z_1 z_2}^{(0)\delta\nu}(\boldsymbol{p}, \boldsymbol{l} - \boldsymbol{m}, \tau_1 - \tau_2) \, \mathcal{G}_{z_2 z_1}^{(0)\nu\delta}(\boldsymbol{l}, \boldsymbol{p} - \boldsymbol{q}, \tau_2 - \tau_1) \right] \quad (3.49)
$$

The "a" in the superscript on the left-hand side is used to distinguish this diagram from the double exchange second order diagram, which we label "b". For a translationally invariant system, we apply Eq. 3.31, and Fourier transform from τ to ik_n using the substitution in Eq. 3.33. Thus, we obtain

$$
\frac{1}{\beta} \sum_{ik_n} e^{-ik_n \tau} \, \mathcal{G}_{zz'}^{(2a)\alpha\beta}(\boldsymbol{k}, ik_n) = \frac{-4}{\beta^5} \frac{U_{\gamma\delta} U_{\mu\nu}}{(L_x L_y)^2} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \sum_{z_1 z_2} \sum_{\boldsymbol{pq}} \sum_{ik_n, ik_n', ip_n, ip_n', ip_n''}
$$

$$
\times \left[e^{-ik_n(\tau - \tau_1)} \, \mathcal{G}_{zz_1}^{(0)\alpha\gamma}(\boldsymbol{k}, ik_n) \, e^{-ip_n(\tau_1 - \tau_2)} \, \mathcal{G}_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{k} - \boldsymbol{q}, ip_n) \, e^{-ik_n'(\tau_2)} \, \mathcal{G}_{z_2 z'}^{(0)\mu\beta}(\boldsymbol{k}, ik_n') \right.
$$

$$
\left. \times e^{-ip_n'(\tau_1 - \tau_2)} \, \mathcal{G}_{z_1 z_2}^{(0)\delta\nu}(\boldsymbol{p}, ip_n') \, e^{-ip_n''(\tau_2 - \tau_1)} \, \mathcal{G}_{z_2 z_1}^{(0)\nu\delta}(\boldsymbol{p} - \boldsymbol{q}, ip_n'') \right] \quad (3.50)
$$

where $ik_n, ik_n', ip_n, ip_n', ip_n''$ are fermionic Matsubara frequencies. Evaluating $\int d\tau_1 \int d\tau_2$, and since this expression must be true for all τ, we obtain

$$
\mathcal{G}_{zz'}^{(2a)\alpha\beta}(\boldsymbol{k}, ik_n) = \frac{-4}{\beta^2} \frac{U_{\gamma\delta} U_{\mu\nu}}{(L_x L_y)^2} \sum_{z_1 z_2} \sum_{\boldsymbol{pq}} \sum_{ip_n', ip_n''}
$$

$$
\times \mathcal{G}_{zz_1}^{(0)\alpha\gamma}(\boldsymbol{k}, ik_n) \, \mathcal{G}_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{k} - \boldsymbol{q}, ik_n - ip_n' + ip_n'')
$$

$$
\times \mathcal{G}_{z_2 z'}^{(0)\mu\beta}(\boldsymbol{k}, ik_n) \, \mathcal{G}_{z_1 z_2}^{(0)\delta\nu}(\boldsymbol{p}, ip_n') \, \mathcal{G}_{z_2 z_1}^{(0)\nu\delta}(\boldsymbol{p} - \boldsymbol{q}, ip_n'') \quad (3.51)
$$

Making a change of variable and extracting the self-energy from this equation, using the same technique as in the previous section, results in a self-energy of the form

$$\Sigma_{z_1z_2}^{(2a)\gamma\mu}(\boldsymbol{k}, ik_n) = \frac{-4}{\beta^2}\frac{U_{\gamma\delta}\,U_{\mu\nu}}{(L_xL_y)^2}\sum_{\boldsymbol{pq}}\sum_{ip_niq_n}$$
$$\times\,\mathcal{G}_{z_1z_2}^{(0)\gamma\mu}(\boldsymbol{k}+\boldsymbol{q}, ik_n+iq_n)\,\mathcal{G}_{z_1z_2}^{(0)\delta\nu}(\boldsymbol{p}, ip_n)\,\mathcal{G}_{z_2z_1}^{(0)\nu\delta}(\boldsymbol{p}+\boldsymbol{q}, ip_n+iq_n) \quad (3.52)$$

where iq_n is a bosonic Matsubara frequency. By convention, we write this as

$$\Sigma_{z_1z_2}^{(2a)\gamma\mu}(\boldsymbol{k}, ik_n) = \frac{-4}{\beta}\frac{U_{\gamma\delta}\,U_{\mu\nu}}{(L_xL_y)^2}\sum_{\boldsymbol{q}}\sum_{iq_n}\mathcal{G}_{z_1z_2}^{(0)\gamma\mu}(\boldsymbol{k}+\boldsymbol{q}, ik_n+iq_n)\,\Pi_{z_1z_2}^{(0)\delta\nu}(\boldsymbol{q}, iq_n) \quad (3.53)$$

where

$$\Pi_{z_1z_2}^{(0)\delta\nu}(\boldsymbol{q}, iq_n) = \frac{1}{\beta}\sum_{\boldsymbol{p}}\sum_{ip_n}\mathcal{G}_{z_1z_2}^{(0)\delta\nu}(\boldsymbol{p}, ip_n)\,\mathcal{G}_{z_2z_1}^{(0)\nu\delta}(\boldsymbol{p}+\boldsymbol{q}, ip_n+iq_n) \quad (3.54)$$

is called the polarization function. There is no summation over the ν, δ in Eq. 3.54. The spectral representation of the polarization function is

$$\Pi_{z_1z_2}^{(0)\delta\nu}(\boldsymbol{q}, iq_n) = \frac{1}{\beta}\int\frac{d\epsilon}{2\pi}\frac{\Pi_{z_1z_2}''^{(0)\delta\nu}(\boldsymbol{q}, \epsilon)}{iq_n - \epsilon} \quad (3.55)$$

where we have used the form of Eq. 3.40. Substituting Eq. 3.55 into Eq. 3.53, and writing $\mathcal{G}_{z_1z_2}^{(0)\gamma\mu}(\boldsymbol{k}+\boldsymbol{q}, ik_n+iq_n)$ in its spectral representation using Eq. 3.40 as well, we obtain

$$\Sigma_{z_1z_2}^{(2a)\gamma\mu}(\boldsymbol{k}, ik_n) = \frac{-4}{\beta^2}\frac{U_{\gamma\delta}\,U_{\mu\nu}}{(L_xL_y)^2}\sum_{\boldsymbol{q}}\sum_{iq_n}\int\frac{d\epsilon_1}{2\pi}\frac{A_{z_1z_2}^{(0)\gamma\mu}(\boldsymbol{k}+\boldsymbol{q}, \epsilon_1)}{ik_n+iq_n-\epsilon_1}\int\frac{d\epsilon}{2\pi}\frac{\Pi_{z_1z_2}''^{(0)\delta\nu}(\boldsymbol{q}, \epsilon)}{iq_n - \epsilon} \quad (3.56)$$

Applying Eq. 3.40 to Eq. 3.54, and computing the fermionic Matsubara sum, $\frac{1}{\beta}\sum_{ip_n}$, gives

$$\frac{1}{\beta}\int\frac{d\epsilon}{2\pi}\frac{\Pi_{z_1z_2}''^{(0)\delta\nu}(\boldsymbol{q}, \epsilon)}{iq_n - \epsilon}$$
$$= \sum_{\boldsymbol{p}}\int\frac{d\epsilon_2}{2\pi}A_{z_1z_2}^{(0)\delta\nu}(\boldsymbol{p}, \epsilon_2)\int\frac{d\epsilon_3}{2\pi}A_{z_2z_1}^{(0)\nu\delta}(\boldsymbol{p}+\boldsymbol{q}, \epsilon_3)\left(\frac{n_F(\epsilon_2) - n_F(\epsilon_3 - iq_n)}{iq_n - (\epsilon_3 - \epsilon_2)}\right) \quad (3.57)$$

Recognizing that $\epsilon = \epsilon_3 - \epsilon_2$ and $n_F(\epsilon_3 - iq_n) = n_F(\epsilon_3)$. We can find $\Pi_{z_1 z_2}''^{(0)\delta\nu}(\boldsymbol{q}, \epsilon)$ by inspection:

$$\Pi_{z_1 z_2}''^{(0)\delta\nu}(\boldsymbol{q}, \epsilon) = \beta \sum_{\boldsymbol{p}} \int \frac{d\epsilon_2}{2\pi} A_{z_1 z_2}^{(0)\delta\nu}(\boldsymbol{p}, \epsilon_2) \int \frac{d\epsilon_3}{2\pi} A_{z_2 z_1}^{(0)\nu\delta}(\boldsymbol{p} + \boldsymbol{q}, \epsilon_3)$$

$$\times \left(n_F(\epsilon_2) - n_F(\epsilon_3) \right) \delta(\epsilon + \epsilon_2 - \epsilon_3) \quad (3.58)$$

Substituting this into Eq. 3.56 yields

$$\Sigma_{z_1 z_2}^{(2a)\gamma\mu}(\boldsymbol{k}, ik_n) = \frac{-4}{\beta} \frac{U_{\gamma\delta} U_{\mu\nu}}{(L_x L_y)^2} \sum_{\boldsymbol{p q}} \sum_{iq_n} \int \frac{d\epsilon_1}{2\pi} \frac{A_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{k} + \boldsymbol{q}, \epsilon_1)}{iq_n + ik_n - \epsilon_1} \int \frac{d\epsilon_2}{2\pi} A_{z_1 z_2}^{(0)\delta\nu}(\boldsymbol{p}, \epsilon_2)$$

$$\times \int \frac{d\epsilon_3}{2\pi} A_{z_2 z_1}^{(0)\nu\delta}(\boldsymbol{p} + \boldsymbol{q}, \epsilon_3) \left(n_F(\epsilon_2) - n_F(\epsilon_3) \right) \int \frac{d\epsilon}{2\pi} \delta(\epsilon + \epsilon_2 - \epsilon_3) \frac{1}{iq_n - \epsilon} \quad (3.59)$$

Computing the bosonic Matsubara sum, $\frac{1}{\beta} \sum_{iq_n}$, recognizing that $n_B(\epsilon_1 - ik_n) = -n_F(\epsilon_1)$, where n_B is the Bose function, and evaluating $\int d\epsilon$, gives

$$\Sigma_{z_1 z_2}^{(2a)\gamma\mu}(\boldsymbol{k}, ik_n) = -4 \frac{U_{\gamma\delta} U_{\mu\nu}}{(L_x L_y)^2} \sum_{\boldsymbol{p q}} \int \frac{d\epsilon_1}{2\pi} A_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{k} + \boldsymbol{q}, \epsilon_1) \int \frac{d\epsilon_2}{2\pi} A_{z_1 z_2}^{(0)\delta\nu}(\boldsymbol{p}, \epsilon_2)$$

$$\times \int \frac{d\epsilon_3}{2\pi} A_{z_2 z_1}^{(0)\nu\delta}(\boldsymbol{p} + \boldsymbol{q}, \epsilon_3) \left(n_F(\epsilon_2) - n_F(\epsilon_3) \right) \left(\frac{n_F(\epsilon_1) + n_B(\epsilon_3 - \epsilon_2)}{-ik_n + \epsilon_1 + \epsilon_2 - \epsilon_3} \right) \quad (3.60)$$

Analytically continuining this result, $\lim_{ik_n \to \omega + i\delta}$, and letting $q \to -q$, we arrive at the retarded second order self-energy for the pair-bubble diagram, which is

$$\Sigma_{z_1 z_2}^{R(2a)\gamma\mu}(\boldsymbol{k}, \omega) = 4 \frac{U_{\gamma\delta} U_{\mu\nu}}{(L_x L_y)^2} \sum_{\boldsymbol{p q}} \int \frac{d\epsilon_1}{2\pi} A_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{k} - \boldsymbol{q}, \epsilon_1) \int \frac{d\epsilon_2}{2\pi} A_{z_1 z_2}^{(0)\delta\nu}(\boldsymbol{p}, \epsilon_2)$$

$$\times \int \frac{d\epsilon_3}{2\pi} A_{z_2 z_1}^{(0)\nu\delta}(\boldsymbol{p} - \boldsymbol{q}, \epsilon_3) \left(\frac{(n_F(\epsilon_2) - n_F(\epsilon_3))(n_F(\epsilon_1) + n_B(\epsilon_3 - \epsilon_2))}{\omega + i\delta - \epsilon_1 - \epsilon_2 + \epsilon_3} \right) \quad (3.61)$$

This is non-Hermitian, and will thus allow us to examine the behaviour of the quasiparticle lifetime.

3.4.2 Double Exchange Diagram

We now examine the other non-Hermitian diagram: the double exchange diagram. Interestingly, this diagram vanishes identically in the conventional Hubbard model without SOC,

by the Pauli principle, because it would describe an electron of fixed spin interacting with itself. In the presence of SOC, the spin of an electron oscillates between up and down as it propagates, and this diagram is non-zero. It is derived following the same steps detailed in Sec. 3.4.1. As such, only the final result is included here. The second order retarded self-energy for the double exchange diagram is

$$\Sigma_{z_1 z_2}^{R(2b)\gamma\mu}(\boldsymbol{k}, \omega) = -4 \frac{U_{\gamma\delta} U_{\mu\nu}}{(L_x L_y)^2} \sum_{\boldsymbol{pq}} \int \frac{d\epsilon_1}{2\pi} A_{z_1 z_2}^{(0)\gamma\nu}(\boldsymbol{k} - \boldsymbol{q}, \epsilon_1) \int \frac{d\epsilon_2}{2\pi} A_{z_2 z_1}^{(0)\nu\delta}(\boldsymbol{p}, \epsilon_2)$$

$$\times \int \frac{d\epsilon_3}{2\pi} A_{z_1 z_2}^{(0)\delta\mu}(\boldsymbol{p} + \boldsymbol{q}, \epsilon_3) \left(\frac{(n_F(\epsilon_3) - n_F(\epsilon_2))(n_F(\epsilon_1) + n_B(\epsilon_2 - \epsilon_3))}{\omega + i\delta - \epsilon_1 + \epsilon_2 - \epsilon_3} \right) \quad (3.62)$$

The non-Hermitian second order self-energy is the sum of each of these diagrams:

$$\Sigma_{z_1 z_2}^{R(2)\gamma\mu}(\boldsymbol{k}, \omega) = \Sigma_{z_1 z_2}^{R(2a)\gamma\mu}(\boldsymbol{k}, \omega) + \Sigma_{z_1 z_2}^{R(2b)\gamma\mu}(\boldsymbol{k}, \omega) \quad (3.63)$$

3.4.3 Broadening Function

We can write the retarded self-energy as

$$\Sigma_{z_1 z_2}^{R(2)\gamma\mu}(\boldsymbol{k}, \omega) = \Lambda_{z_1 z_2}^{\gamma\mu}(\boldsymbol{k}, \omega) - \frac{i}{2} \Gamma_{z_1 z_2}^{\gamma\mu}(\boldsymbol{k}, \omega) \quad (3.64)$$

where

$$\Lambda_{z_1 z_2}^{\gamma\mu}(\boldsymbol{k}, \omega) = \frac{\Sigma_{z_1 z_2}^{R(2)\gamma\mu}(\boldsymbol{k}, \omega) + \left(\Sigma_{z_2 z_1}^{R(2)\mu\gamma}(\boldsymbol{k}, \omega) \right)^*}{2} \quad (3.65)$$

is the Hermitian part of the self-energy, and $\Gamma_{z_1 z_2}^{\gamma\mu}(\boldsymbol{k}, \omega)$ is the negative of twice its anti-Hermitian part, given by

$$\Gamma_{z_1 z_2}^{\gamma\mu}(\boldsymbol{k}, \omega) = i \left(\Sigma_{z_1 z_2}^{R(2)\gamma\mu}(\boldsymbol{k}, \omega) - \left(\Sigma_{z_2 z_1}^{R(2)\mu\gamma}(\boldsymbol{k}, \omega) \right)^* \right) \quad (3.66)$$

We call this the broadening function; it gives us the frequency dependence of the quasiparticle lifetime and is the quantity we focus on for the rest of this chapter. Substituting Eq.

3.63 into Eq. 3.66 yields

$$\Gamma^{\gamma\mu}_{z_1z_2}(\boldsymbol{k},\omega) = \Gamma^{(a)\gamma\mu}_{z_1z_2}(\boldsymbol{k},\omega) + \Gamma^{(b)\gamma\mu}_{z_1z_2}(\boldsymbol{k},\omega) \tag{3.67}$$

where

$$\Gamma^{(a)\gamma\mu}_{z_1z_2}(\boldsymbol{k},\omega) = 4\frac{U_{\gamma\delta}\,U_{\mu\nu}}{(L_xL_y)^2} \sum_{\boldsymbol{pq}} \int \frac{d\epsilon_1}{2\pi} A^{(0)\gamma\mu}_{z_1z_2}(\boldsymbol{k}-\boldsymbol{q},\epsilon_1) \int \frac{d\epsilon_2}{2\pi} A^{(0)\delta\nu}_{z_1z_2}(\boldsymbol{p},\epsilon_2) \int \frac{d\epsilon_3}{2\pi}$$

$$\times\, A^{(0)\nu\delta}_{z_2z_1}(\boldsymbol{p}-\boldsymbol{q},\epsilon_3)\,(n_F(\epsilon_2)-n_F(\epsilon_3))(n_F(\epsilon_1)+n_B(\epsilon_3-\epsilon_2))\,2\pi\,\delta(\omega-(\epsilon_1+\epsilon_2-\epsilon_3))$$

$$\tag{3.68}$$

and

$$\Gamma^{(b)\gamma\mu}_{z_1z_2}(\boldsymbol{k},\omega) = -4\frac{U_{\gamma\delta}\,U_{\mu\nu}}{(L_xL_y)^2} \sum_{\boldsymbol{pq}} \int \frac{d\epsilon_1}{2\pi} A^{(0)\gamma\nu}_{z_1z_2}(\boldsymbol{k}-\boldsymbol{q},\epsilon_1) \int \frac{d\epsilon_2}{2\pi} A^{(0)\nu\delta}_{z_2z_1}(\boldsymbol{p},\epsilon_2) \int \frac{d\epsilon_3}{2\pi}$$

$$\times\, A^{(0)\delta\mu}_{z_1z_2}(\boldsymbol{p}+\boldsymbol{q},\epsilon_3)\,(n_F(\epsilon_3)-n_F(\epsilon_2))(n_F(\epsilon_1)+n_B(\epsilon_2-\epsilon_3))\,2\pi\,\delta(\omega-(\epsilon_1-\epsilon_2+\epsilon_3))$$

$$\tag{3.69}$$

These two expressions were obtained using the identity $1/(x+i\delta) = \mathcal{P}(1/x) - i\pi\delta(x)$, where $\mathcal{P}$ denotes the principal value. Performing $\int d\epsilon_3$, we obtain $\epsilon_3 = \epsilon_1 + \epsilon_2 - \omega$ from Eq. 3.68 and $\epsilon_3 = \omega - \epsilon_1 + \epsilon_2$ from Eq. 3.69. Thus, Eq. 3.68 becomes

$$\Gamma^{(a)\gamma\mu}_{z_1z_2}(\boldsymbol{k},\omega) = 4\frac{U_{\gamma\delta}\,U_{\mu\nu}}{(L_xL_y)^2} \sum_{\boldsymbol{pq}} \int \frac{d\epsilon_1}{2\pi} A^{(0)\gamma\mu}_{z_1z_2}(\boldsymbol{k}-\boldsymbol{q},\epsilon_1) \int \frac{d\epsilon_2}{2\pi} A^{(0)\delta\nu}_{z_1z_2}(\boldsymbol{p},\epsilon_2)$$

$$\times\, A^{(0)\nu\delta}_{z_2z_1}(\boldsymbol{p}-\boldsymbol{q},\epsilon_1+\epsilon_2-\omega)(n_F(\epsilon_2)-n_F(\epsilon_1+\epsilon_2-\omega))(n_F(\epsilon_1)+n_B(\epsilon_1-\omega)) \tag{3.70}$$

and Eq. 3.69 becomes

$$\Gamma^{(b)\gamma\mu}_{z_1z_2}(\boldsymbol{k},\omega) = -4\frac{U_{\gamma\delta}\,U_{\mu\nu}}{(L_xL_y)^2} \sum_{\boldsymbol{pq}} \int \frac{d\epsilon_1}{2\pi} A^{(0)\gamma\nu}_{z_1z_2}(\boldsymbol{k}-\boldsymbol{q},\epsilon_1) \int \frac{d\epsilon_2}{2\pi} A^{(0)\nu\delta}_{z_2z_1}(\boldsymbol{p},\epsilon_2)$$

$$\times\, A^{(0)\delta\mu}_{z_1z_2}(\boldsymbol{p}+\boldsymbol{q},\omega-\epsilon_1+\epsilon_2)\,(n_F(\omega-\epsilon_1+\epsilon_2)-n_F(\epsilon_2))(n_F(\epsilon_1)+n_B(\epsilon_1-\omega)) \tag{3.71}$$

So far we have considered a general 4×4 contact interaction matrix, U, in our Hubbard model, Eq. 2.8. To simplify our model further, we assume that all orbitals/spins interact in the same fashion, and thus the interaction matrix is written as

$$U_{\gamma\delta} = \frac{U}{2}\left(1 - \delta_{\gamma\delta}\right) \tag{3.72}$$

where U is a constant that determines the strength of the interaction and we divide by two to eliminate double counting. (We, however, emphasize that our formalism can be applied to any interaction matrix.) As mentioned earlier, all repeated orbital/spin indices are summed over. Following this convention, Eq. 3.70 becomes

$$\Gamma_{z_1 z_2}^{(a)\gamma\mu}(\boldsymbol{k},\omega) = \frac{U^2}{(L_x L_y)^2}\sum_{\boldsymbol{pq}}\int\frac{d\epsilon_1}{2\pi}A_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{k}-\boldsymbol{q},\epsilon_1)\int\frac{d\epsilon_2}{2\pi}\left(n_F(\epsilon_2) - n_F(\epsilon_1 + \epsilon_2 - \omega)\right)$$

$$\times\left(n_F(\epsilon_1) + n_B(\epsilon_1 - \omega)\right)\left[\mathrm{Tr}\left[A_{z_1 z_2}^{(0)}(\boldsymbol{p},\epsilon_2)\,A_{z_2 z_1}^{(0)}(\boldsymbol{p}-\boldsymbol{q},\epsilon_1 + \epsilon_2 - \omega)\right]\right.$$

$$-\left(A_{z_1 z_2}^{(0)}(\boldsymbol{p},\epsilon_2)\,A_{z_2 z_1}^{(0)}(\boldsymbol{p}-\boldsymbol{q},\epsilon_1 + \epsilon_2 - \omega)\right)_{\gamma\gamma} -\left(A_{z_1 z_2}^{(0)}(\boldsymbol{p},\epsilon_2)\,A_{z_2 z_1}^{(0)}(\boldsymbol{p}-\boldsymbol{q},\epsilon_1 + \epsilon_2 - \omega)\right)_{\mu\mu}$$

$$\left. + A_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{p},\epsilon_2)\,A_{z_2 z_1}^{(0)\mu\gamma}(\boldsymbol{p}-\boldsymbol{q},\epsilon_1 + \epsilon_2 - \omega)\right] \tag{3.73}$$

and Eq. 3.71 becomes

$$\Gamma_{z_1 z_2}^{(b)\gamma\mu}(\boldsymbol{k},\omega) = -\frac{U^2}{(L_x L_y)^2}\sum_{\boldsymbol{pq}}\int\frac{d\epsilon_1}{2\pi}\int\frac{d\epsilon_2}{2\pi}$$

$$\times\left(n_F(\omega - \epsilon_1 + \epsilon_2) - n_F(\epsilon_2)\right)\left(n_F(\epsilon_1) + n_B(\epsilon_1 - \omega)\right)$$

$$\times\left[\left(A_{z_1 z_2}^{(0)}(\boldsymbol{k}-\boldsymbol{q},\epsilon_1)\,A_{z_2 z_1}^{(0)}(\boldsymbol{p},\epsilon_2)\,A_{z_1 z_2}^{(0)}(\boldsymbol{p}+\boldsymbol{q},\omega - \epsilon_1 + \epsilon_2)\right)_{\gamma\mu}\right.$$

$$-\left(A_{z_1 z_2}^{(0)}(\boldsymbol{k}-\boldsymbol{q},\epsilon_1)\,A_{z_2 z_1}^{(0)}(\boldsymbol{p},\epsilon_2)\right)_{\gamma\gamma}A_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{p}+\boldsymbol{q},\omega - \epsilon_1 + \epsilon_2)$$

$$-A_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{k}-\boldsymbol{q},\epsilon_1)\left(A_{z_2 z_1}^{(0)}(\boldsymbol{p},\epsilon_2)A_{z_1 z_2}^{(0)}(\boldsymbol{p}+\boldsymbol{q},\omega - \epsilon_1 + \epsilon_2)\right)_{\mu\mu}$$

$$\left. + A_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{k}-\boldsymbol{q},\epsilon_1)\,A_{z_2 z_1}^{(0)\mu\gamma}(\boldsymbol{p},\epsilon_2)\,A_{z_1 z_2}^{(0)\gamma\mu}(\boldsymbol{p}+\boldsymbol{q},\omega - \epsilon_1 + \epsilon_2)\right] \tag{3.74}$$

In the next section we discuss how to numerically evaluate Eq. 3.73 and 3.74.

3.4.4 Numerical Calculation of the Broadening Function

Due to time constraints and the extremely large computational requirement involved in evaluating the broadening function, it was only possible to evaluate Eq. 3.73 and 3.74 at a single k_x, k_y point, and for a narrow range of ω. We chose to examine the behaviour of the broadening function near the Fermi surface as it is in this range of energies that the quasiparticle picture is valid [63, 64, 69]. We chose $k_x = k_F$ and $k_y = 0$, and since the Fermi surface is rotationally invariant, the same ω dependence applies for $\boldsymbol{k}$ anywhere on the Fermi surface.

The bounds on the energy integrals, $\int d\epsilon_1 \int d\epsilon_2$, are determined by the Fermi and Bose functions. In this thesis, while the expressions derived so far can be used to determine the broadening function at any temperature, for computational simplicity, all numerical results are given at zero temperature, and therefore the Fermi functions are step functions, and the bounds on the integrals are finite. Eq. 3.73 and 3.74 each have the term $n_F(\epsilon_1) + n_B(\epsilon_1 - \omega)$. Evaluating this at zero temperature yields

$$n_F(\epsilon_1) + n_B(\epsilon_1 - \omega) = \begin{cases} -1, & \omega > \epsilon_1 > 0 \\ 1, & 0 > \epsilon_1 > \omega \\ 0, & \text{otherwise} \end{cases} \tag{3.75}$$

Examining $n_F(\epsilon_2) - n_F(\epsilon_1 + \epsilon_2 - \omega)$ in Eq. 3.73 gives the bounds on ϵ_2:

$$n_F(\epsilon_2) - n_F(\epsilon_1 + \epsilon_2 - \omega) = \begin{cases} -1, & \omega - \epsilon_1 > \epsilon_2 > 0 \\ 1, & 0 > \epsilon_2 > \omega - \epsilon_1 \\ 0, & \text{otherwise} \end{cases} \tag{3.76}$$

To find the ϵ_2 bounds in Eq. 3.74, we evaluate $n_F(\omega - \epsilon_1 + \epsilon_2) - n_F(\epsilon_2))$, which yields

$$n_F(\omega - \epsilon_1 + \epsilon_2) - n_F(\epsilon_2)) = \begin{cases} -1, & 0 > \epsilon_2 > \epsilon_1 - \omega \\ 1, & \epsilon_1 - \omega > \epsilon_2 > 0 \\ 0, & \text{otherwise} \end{cases} \qquad (3.77)$$

In determining the bounds on $\sum_{pq}$, we use the fact that for energies near the Fermi level, the non-interacting spectral function is non-zero only near the Fermi surface (recall Eq. 2.31 and 2.32 as well as the middle panel of Fig. 3.5). Thus, in summing over p and q, we need to only sum over a thin annulus of width 2ζ centred around the Fermi momentum (see Fig. 3.8). We set $\zeta = 0.1\,k_F$. The p and q values we sum over in Eq. 3.73 and 3.74 are thus

$$p_x = Q_p \cos(\theta_p)\,, \qquad p_y = Q_p \sin(\theta_p) \qquad (3.78)$$

$$q_x = Q_q \cos(\theta_q) + k_F\,, \qquad q_y = Q_q \sin(\theta_q) \qquad (3.79)$$

where $Q_{p,q}$ is the magnitude of p or q, and $\theta_{p,q}$ is the angle of p or q with respect to the x axis and goes from 0 to 2π. Therefore, the integrals over p and q are best evaluated in polar coordinates, and we make the following substitution:

$$\sum_{pq} \rightarrow \sum_{Q_p} Q_p \sum_{\theta_p} \sum_{Q_q} Q_q \sum_{\theta_q} \qquad (3.80)$$

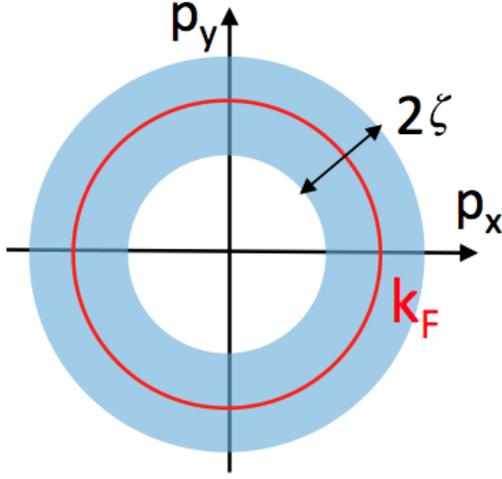where the additional factors of $Q_{p,q}$ arise from the Jacobian.

Figure 3.8: The thin annulus of integration for $\boldsymbol{p}$. The same area of integration applies to $\boldsymbol{q}$, but it is shifted by $-k_F\,\hat{x}$.

Writing the bounds on the integrals (we convert $\sum_{\boldsymbol{pq}}$ to an integral) for $\boldsymbol{p}, \boldsymbol{q}, \epsilon_1$, and ϵ_2, we arrive at the final expression for $\Gamma^{\gamma\mu}_{z_1 z_2}(\boldsymbol{k}, \omega)$:

$$\Gamma^{(a)\gamma\mu}_{z_1 z_2}(\boldsymbol{k}, \omega) = U^2 \int_{k_F - \zeta}^{k_F + \zeta} \frac{dQ_q}{2\pi}\, Q_q \int_0^{2\pi} \frac{d\theta_q}{2\pi} \int_{k_F - \zeta}^{k_F + \zeta} \frac{dQ_p}{2\pi}\, Q_p \int_0^{2\pi} \frac{d\theta_p}{2\pi} \int_0^{\omega} \frac{d\epsilon_1}{2\pi} \int_0^{\omega - \epsilon_1} \frac{d\epsilon_2}{2\pi}$$

$$\left[A^{(0)\gamma\mu}_{z_1 z_2}(\boldsymbol{k} - \boldsymbol{q}, \epsilon_1) \operatorname{Tr}\left[A^{(0)}_{z_1 z_2}(\boldsymbol{p}, \epsilon_2)\, A^{(0)}_{z_2 z_1}(\boldsymbol{p} - \boldsymbol{q}, \epsilon_1 + \epsilon_2 - \omega) \right] \right.$$

$$- \left(A^{(0)}_{z_1 z_2}(\boldsymbol{p}, \epsilon_2)\, A^{(0)}_{z_2 z_1}(\boldsymbol{p} - \boldsymbol{q}, \epsilon_1 + \epsilon_2 - \omega) \right)_{\gamma\gamma} - \left(A^{(0)}_{z_1 z_2}(\boldsymbol{p}, \epsilon_2)\, A^{(0)}_{z_2 z_1}(\boldsymbol{p} - \boldsymbol{q}, \epsilon_1 + \epsilon_2 - \omega) \right)_{\mu\mu}$$

$$\left. + A^{(0)\gamma\mu}_{z_1 z_2}(\boldsymbol{p}, \epsilon_2)\, A^{(0)\mu\gamma}_{z_2 z_1}(\boldsymbol{p} - \boldsymbol{q}, \epsilon_1 + \epsilon_2 - \omega) \right] \quad (3.81)$$

and

$$\Gamma^{(b)\gamma\mu}_{z_1 z_2}(\boldsymbol{k}, \omega) = U^2 \int_{k_F - \zeta}^{k_F + \zeta} \frac{dQ_q}{2\pi}\, Q_q \int_0^{2\pi} \frac{d\theta_q}{2\pi} \int_{k_F - \zeta}^{k_F + \zeta} \frac{dQ_p}{2\pi}\, Q_p \int_0^{2\pi} \frac{d\theta_p}{2\pi} \int_0^{\omega} \frac{d\epsilon_1}{2\pi} \int_0^{\epsilon_1 - \omega} \frac{d\epsilon_2}{2\pi}$$

$$\times \left[\left(A^{(0)}_{z_1 z_2}(\boldsymbol{k} - \boldsymbol{q}, \epsilon_1)\, A^{(0)}_{z_2 z_1}(\boldsymbol{p}, \epsilon_2)\, A^{(0)}_{z_1 z_2}(\boldsymbol{p} + \boldsymbol{q}, \omega - \epsilon_1 + \epsilon_2) \right)_{\gamma\mu} \right.$$

$$- \left(A^{(0)}_{z_1 z_2}(\boldsymbol{k} - \boldsymbol{q}, \epsilon_1)\, A^{(0)}_{z_2 z_1}(\boldsymbol{p}, \epsilon_2) \right)_{\gamma\gamma} A^{(0)\gamma\mu}_{z_1 z_2}(\boldsymbol{p} + \boldsymbol{q}, \omega - \epsilon_1 + \epsilon_2)$$

$$- A^{(0)\gamma\mu}_{z_1 z_2}(\boldsymbol{k} - \boldsymbol{q}, \epsilon_1) \left(A^{(0)}_{z_2 z_1}(\boldsymbol{p}, \epsilon_2) A^{(0)}_{z_1 z_2}(\boldsymbol{p} + \boldsymbol{q}, \omega - \epsilon_1 + \epsilon_2) \right)_{\mu\mu}$$

$$\left. + A^{(0)\gamma\mu}_{z_1 z_2}(\boldsymbol{k} - \boldsymbol{q}, \epsilon_1)\, A^{(0)\mu\gamma}_{z_2 z_1}(\boldsymbol{p}, \epsilon_2)\, A^{(0)\gamma\mu}_{z_1 z_2}(\boldsymbol{p} + \boldsymbol{q}, \omega - \epsilon_1 + \epsilon_2) \right] \quad (3.82)$$

In practice, we convert the integrals to sums, since we evaluate $\Gamma^{\gamma\mu}_{z_1z_2}(\boldsymbol{k},\omega)$ numerically. Thus, we make the following substitutions in the above two expressions:

$$\int \frac{d\epsilon_{1,2}}{2\pi} \to \sum_{\epsilon_{1,2}} \frac{\Delta\epsilon_{1,2}}{2\pi} \tag{3.83}$$

$$\int \frac{dQ_{p,q}}{2\pi} \int \frac{d\theta_{p,q}}{2\pi} Q_{p,q} \to \frac{1}{N_{Q_{p,q}} N_{\theta_{p,q}}} \sum_{Q_{p,q}} \sum_{\theta_{p,q}} Q_{p,q} \tag{3.84}$$

where $\Delta\epsilon_{1,2}$ is a small step size between adjacent $\epsilon_{1,2}$ values, and $N_{Q_{p,q}}$ and $N_{\theta_{p,q}}$ are the number of $Q_{p,q}$ and $\theta_{p,q}$ values that are summed over, respectively. In a TI, the Fermi level only intersects the SSs, not the bulk states. If ω is chosen to be sufficiently small compared to the bulk bandgap, we see from Eq. 3.81 and 3.82 that the non-interacting spectral functions are evaluated at small arguments, since ϵ_1 and ϵ_2 are of order ω. In other words, for small ω we only need to know the spectral function near the Fermi level.

Due to the large number of sums required to evaluate each Feynman diagram, the numerical evaluation of the GFs needed to be parallelized. As such, each value of ω, θ_q, was run on a separate CPU core. Each of these partial contributions to the Feynman diagrams was then collected and summed. Unlike for the non-interacting calculations, these second order self-energy calculation run times scaled greater than quadratically in N_z. However, they scaled approximately linearly with the number of Q_q, Q_p, θ_q, and θ_p, internal energy ($\epsilon_{1,2}$), and ω values.

We have so far ignored the Hermitian part, $\Lambda^{\gamma\mu}_{z_1z_2}(\boldsymbol{k},\omega)$, in Eq. 3.64. In principle, to calculate the interacting spectral function (which is needed to obtain the quasiparticle lifetime), one should include the Hermitian part as well. This is because a frequency-dependent contribution to the Hermitian part of the self-energy indirectly affects the width of the spectral function via the so-called quasiparticle residue, $Z_{\boldsymbol{k}}$ [63]. Unfortunately, the Hermitian part is given by a triple energy integral, compared to the double energy integrals in Eq. 3.73 and 3.74, which makes its numerical evaluation prohibitively difficult and beyond the scope of this thesis. However, the frequency dependence of the lifetime is controlled by the broadening function, Eq. 3.73 and 3.74, which therefore gives us a qualitative understanding of this frequency dependence.

Three plots of the trace of $\Gamma^{\gamma\mu}_{z_1 z_2}(k_x = k_F, k_y = 0, \omega)$ for ω near the Fermi level and a given layer index, $z_1 = z_2$, are shown in Fig. 3.9, 3.10, and 3.11. In these plots we have chosen $\mu = 0.21$ eV so as to ensure $k_x = |k_F| < 0.04$ Å^{-1}. This corresponds to a Fermi level slightly below the DP. Note that experimentally it is possible to tune the chemical potential in TI thin films by electrostatic gating [74]. We also set the interaction strength, U, to be 1 eV; U only affects the overall magnitude of the broadening function but not its frequency dependence. Unlike in Fig. 3.2, here we only show results for $N_z \geq 40$. This is because for $N_z < 40$, there are not well-defined SSs on which we can examine the behaviour of the broadening function.

In Fig. 3.9, we see how the trace of the broadening function varies with layer index, z, energy, ω, and slab thickness, N_z. In this figure, and in the two that follow, the $N_z \geq 60$ plots are virtually indistinguishable as their DCs and Fermi surfaces are the essentially the same. We find $\mathrm{Tr}[\Gamma(k_x = k_F, k_y = 0, \omega = 0)] = 0$ eV, which means there is an infinite lifetime at the Fermi level, as expected. Indeed, for quasiparticles exactly at the Fermi level, there is no phase space for electron-electron scattering due to the Pauli exclusion principle [63, 65]. As we move away from $\omega = 0$ eV, a finite lifetime results. For a fixed (small) frequency, ω, as a function of layer index, the broadening function follows essentially the probability density of the SSs (see Fig. 3.3) and the spectral intensity (see Fig. 3.4). This is expected from Eq. 3.81 and 3.82, which show that the broadening function for a given layer index only involves the non-interacting spectral function on that layer. This, in turn, is due to the on-site nature of the generalized Hubbard interaction in Eq. 2.8; we expect that a longer range interaction would result in a slower decay of the broadening function with layer index.

(a) $N_z = 40$ Layers

(b) $N_z = 60$ Layers

(c) $N_z = 80$ Layers

(d) $N_z = 200$ Layers
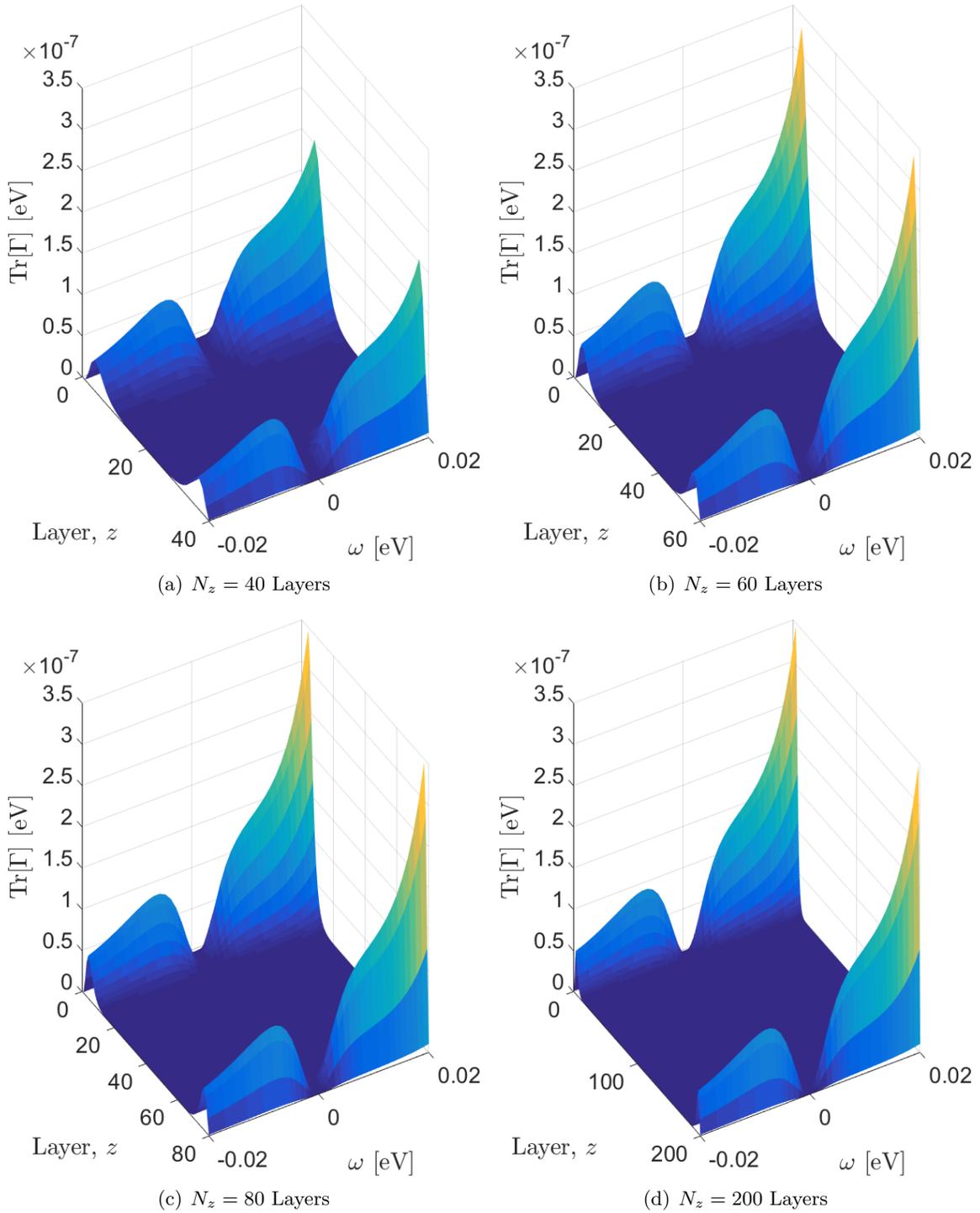
Figure 3.9: The trace of the broadening function. If we imagine a 2D cross-section for constant ω contours, we see a similar curve to that in Fig. 3.3, where $\text{Tr}[\Gamma]$ is highly peaked near the surfaces (they peak at $z = 4$) and as we move away from $z = 4$, $\text{Tr}[\Gamma]$ falls off rapidly.

Fig. 3.10 provides 2D slices through the 3D plots in Fig. 3.9 so as to examine the ω dependence of the broadening function more clearly. Little dependence on the slab thickness is found in this range of thicknesses. Although for $N_z = 40$ the DC is clearly still gapped due to intersurface hybridization, while for higher values of N_z this gap eventually closes (Fig. 3.2). For a chemical potential sufficiently far from the DP the Fermi surface does not vary much with slab thickness. For small ω the broadening only depends on Fermi surface properties, and is thus not affected much by N_z. The only visible change on going from $N_z = 40$ to $N_z = 60$ is the $N_z = 40$ curves are slightly flatter. This is most noticeable for the largest positive values of ω considered ($\omega > 15$ meV). These values probe the behaviour of the spectral function near the DP, which is gapped for $N_z = 40$ but gapless for $N_z \geq 60$. Indeed, the DP is located approximately 16 meV above the Fermi level for the value of μ considered. Since the gapless DP corresponds to an additional scattering channel for electrons, this explains why the broadening function near the DP increases from $N_z = 40$ to $N_z = 60$ but then stays roughly the same.

(a) $N_z = 40$ Layers

(b) $N_z = 60$ Layers

(c) $N_z = 80$ Layers

(d) $N_z = 200$ Layers

Figure 3.10: The trace of the broadening function for various slab thicknesses. These are constant z contours from Fig. 3.9. Due to the reflection symmetry between the top and bottom halves of the the slab, only half the layers need to be shown.

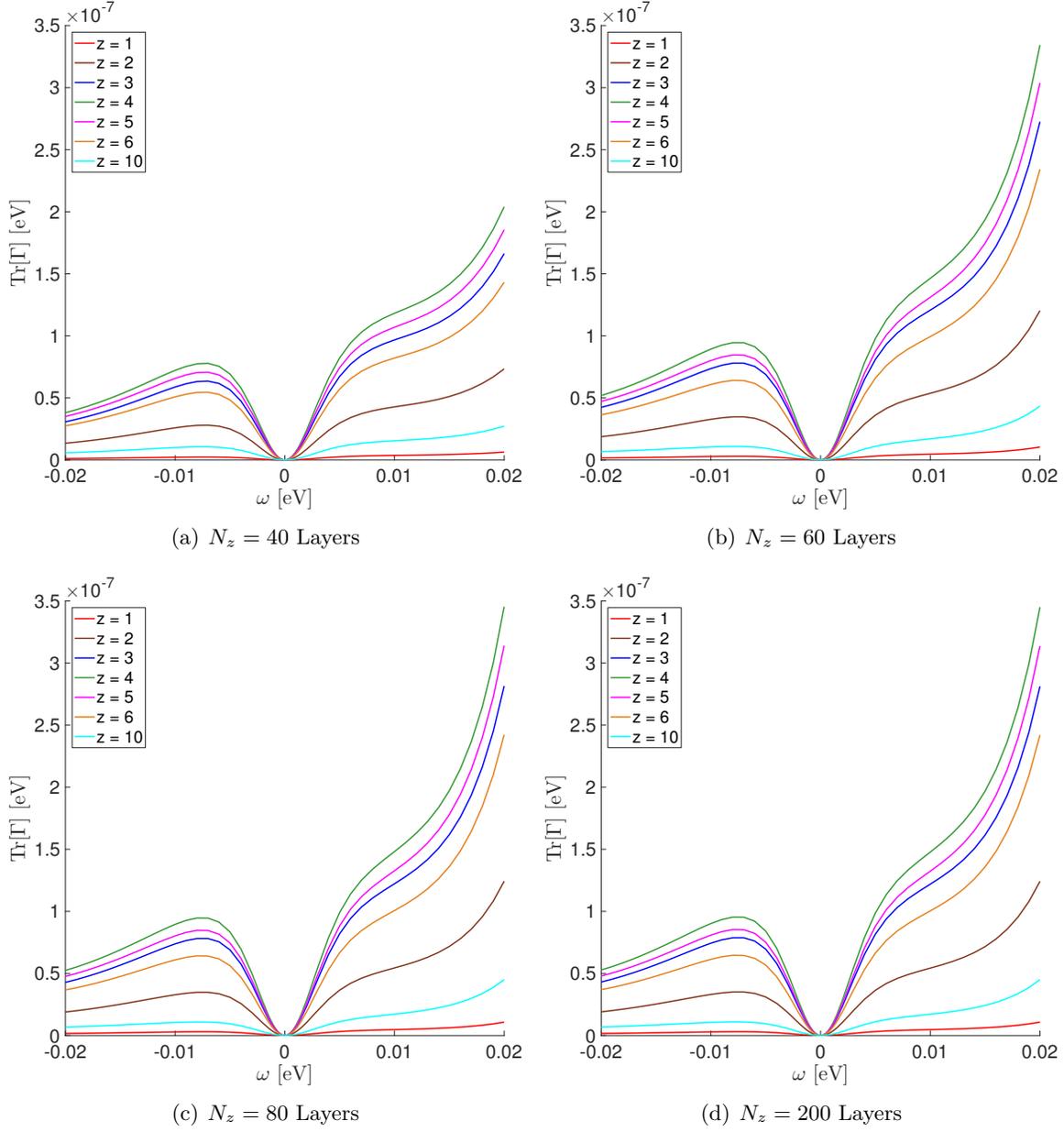For small ω the broadening function is expected to predominantly depend on the dispersion relation near the Fermi surface. This is due to the spectral functions, on which the broadening function depends, only having significant weight near the Fermi surface in the relevant range of energies. For a Fermi liquid at low energies, we expect the broadening function to behave like ω^2 [63].

In Fig. 3.11, we fit the small ω behaviour ($|\omega| < 1$ meV) of the broadening function to ω^2. This is what Song *et al.* measured experimentally ($\omega = E - E_F$) [55]. The two sets of points for each layer index, z, correspond to the positive and negative ω data, which are plotted on the same graph. While the fit is relatively good, we observe a difference in slope for positive and negative frequencies. This is a consequence of the fact that the DC dispersion is in fact not perfectly linear (see Fig. 3.2), and thus positive and negative ω are not strictly equivalent.

Strictly speaking, for a 2D system at very low energies we expect the broadening function to behave as

$$\Gamma \sim \omega^2 \left| \ln \left(\frac{|\omega|}{\mu} \right) \right| \tag{3.85}$$

where μ is measured from the bottom of the band [75]. This expression is for a parabolic band; however, a conventional 2D parabolic dispersion can always be linearized in the vicinity of the Fermi surface. For a Dirac dispersion $E_F = v_F k_F$, while a parabolic band has $\mu = v_F k_F / 2$. Putting this together yields a Γ dependence of

$$\Gamma \sim \omega^2 \left| \ln \left(2 \left| \frac{\omega}{E_F} \right| \right) \right| \tag{3.86}$$

for a Dirac dispersion, where E_F is measured from the DP. The energies required to see this behaviour must be several orders of magnitude less than E_F. If the energies are too large, the sub-leading term, ω^2, dominates the behaviour of Eq. 3.86. Calculating GFs for such energies requires a very small η ($\eta < \omega$ in Eq. 2.35) and would lead to numerical instabilities.

(a) $N_z = 40$ Layers

(b) $N_z = 60$ Layers

(c) $N_z = 80$ Layers

(d) $N_z = 200$ Layers
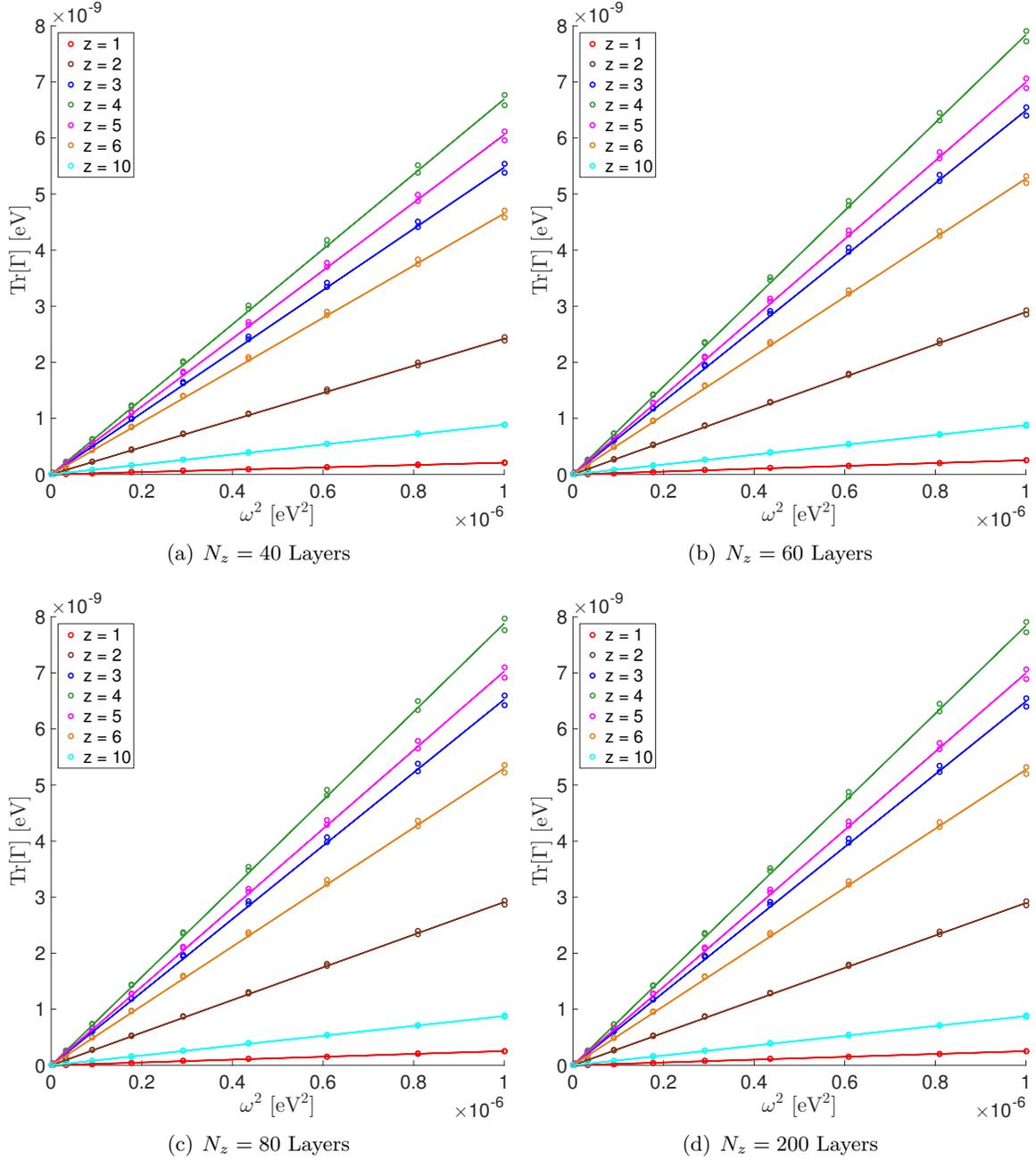
Figure 3.11: Small ω $(-1 \leq \omega \leq 1$ meV$)$ behaviour of Tr$[\Gamma]$.

Chapter 4

Surface States for a Finite Magnetic Field

Electrons in the presence of a magnetic field undergo circular orbits due to the Lorentz force. In a 2D electron gas with a strong perpendicular magnetic field, electrons exhibit quantized circular orbits called LLs. One of the effects of LLs is that they give rise to the existence of the integer quantum Hall effect [6]. In this chapter we review the LL spectrum in both relativistic and non-relativistic systems, and show that the experimental results described in Sec. 1.2.4 indicate a relativistic LL spectrum. We also obtain the LL DOS for a non-interacting TI system using the same Hamiltonian as in Ch. 3. Finally, we adapt the many-body formalism of Ch. 3 to the case of finite magnetic fields and obtain an expression for the second order self-energy in the LL basis.

4.1   Landau Levels Overview

For a non-relativistic electron confined to the xy plane, and a magnetic field, B, applied in the $+z$ direction (here, and in the rest of this chapter, we use the Landau gauge for the vector potential, $\boldsymbol{A} = -By\,\hat{x}$), the LL Hamiltonian can be written as

$$H_{LL} = \frac{(\boldsymbol{p} + e\boldsymbol{A})^2}{2m} = \frac{p_y^2}{2m} + \frac{1}{2}m\omega_c^2(y - y_0)^2 \tag{4.1}$$

where $\boldsymbol{p}$ is the momentum operator, $-e$ is the charge on an electron, m is the mass of an electron, y is the position operator, $\omega_c = \frac{eB}{m}$ is the cyclotron frequency, and $y_0 = \frac{k_x}{eB}$, where k_x is a momentum and is a good quantum number. This has the same form as the 1D quantum harmonic oscillator Hamiltonian. As such, the eigenstates of this Hamiltonian can be written in terms of Hermite polynomials and the energy spectrum is the familiar

harmonic oscillator spectrum:

$$E_n = \omega_c \left(n + \frac{1}{2} \right), \qquad n = 0, 1, 2, \dots \tag{4.2}$$

Here, however, n is the LL index.

For infinite L_x and L_y (the lengths of the slab in the x and y directions, respectively), there is an infinite degeneracy in each LL. However, for a finite L_x and L_y the degeneracy of each LL is

$$g_n = \frac{\text{total flux through area } L_x L_y}{\text{flux quantum}} = \frac{L_x L_y}{2\pi \, \ell_B^2} \tag{4.3}$$

where $\ell_B = (eB)^{-1/2}$ is the magnetic length, and as such the strength of the magnetic field determines the degeneracy [76]. This degeneracy is important in calculating the interaction term in the LL basis (see Sec. 4.3).

For 2D relativistic electrons, the LL spectrum is [56, 57]

$$E_n = E_{DP} + \text{sgn}\,(n) v_F \sqrt{2e|n|B}\,, \qquad n = 0, \pm 1, \pm 2, \dots \tag{4.4}$$

where E_{DP} is the energy of the DP and v_F is the Fermi velocity. This spectrum can be obtained from the Dirac Hamiltonian, $H_{\text{Dirac}} = v_F \, \hat{z} \cdot (\boldsymbol{\sigma} \times \boldsymbol{k})$, where $\boldsymbol{\sigma}$ are Pauli matrices and $\boldsymbol{k}$ denotes momentum. There are some important differences in the LL spectra of Eq. 4.2 and Eq. 4.4. First, Eq. 4.4 describes massless particles; note that there is no mass term present. Next, Eq. 4.4 supports a magnetic field independent ($n = 0$) LL, often referred to as the zeroth LL, which occurs at $E_{n=0} = E_{DP}$. Finally, unlike Eq. 4.2 where $E_n \sim Bn$, in Eq. 4.4, $E_n \sim \sqrt{|n|B}$; thus, the relativistic LLs are not equally spaced.

Although LLs are a fundamentally 2D effect, the existence of the 2D SSs on 3D TIs permits them to emerge on the 3D TI surfaces [76]. Due to the linear Dirac dispersion of TI SS electrons observed in Fig. 1.8, the observation of the zeroth LL in Fig. 1.12, and the $\sqrt{|n|B}$ dependence of the LLs observed in Fig. 1.13, Eq. 4.4 gives the expected LL energy spectrum of TIs.

4.2 Landau Levels in Surface States

We now turn to the inclusion of a magnetic field in our theoretical model for 3D TIs to demonstrate how the LL spectrum emerges from the full bandstructure of 3D TIs, and how the spectrum and associated LL wavefunctions are affected by the slab thickness. In first quantization, the kinetic momentum operator is defined as

$$\boldsymbol{\pi} = \boldsymbol{k} + e\boldsymbol{A} \tag{4.5}$$

where, again, we let $-e$ be the charge on an electron. We add a magnetic field in the $+z$ direction to the non-interacting Hamiltonian, Eq. 3.2, by making the Peierls substitution $(\boldsymbol{k} \to \boldsymbol{\pi})$ in the x and y directions:

$$\boldsymbol{k}_\parallel \to \boldsymbol{\pi}_\parallel = \pi_x\,\hat{x} + \pi_y\,\hat{y} \tag{4.6}$$

Eq. 3.2 therefore becomes

$$h_{\text{eff}}(\boldsymbol{\pi}_\parallel, k_z) = \epsilon_{\boldsymbol{\pi}_\parallel, k_z}\,\mathbb{1} + \mathcal{M}(\boldsymbol{\pi}_\parallel, k_z)\Gamma_5 + B_0\,k_z\Gamma_4 + A_0\,(\pi_y\Gamma_1 - \pi_x\Gamma_2) \tag{4.7}$$

where

$$\epsilon_{\boldsymbol{\pi}_\parallel, k_z} = C_0 + C_1\,k_z^2 + C_2\,\boldsymbol{\pi}_\parallel{}^2 \tag{4.8}$$

$$\mathcal{M}(\boldsymbol{\pi}_\parallel, k_z) = M_0 + M_1\,k_z^2 + M_2\,\boldsymbol{\pi}_\parallel{}^2 \tag{4.9}$$

We can define raising, $\hat{a}^\dagger$, and lowering, $\hat{a}$, operators[1] that raise and lower, respectively, the LL index, n, for a wavefunction, ψ_n, describing an electron in the nth LL. They have

---

[1] Note that we are not in the interaction picture here.

the following properties:

$$\hat{a}\,|\psi_n\rangle = \sqrt{n}\,|\psi_{n-1}\rangle \tag{4.10}$$

$$\hat{a}^\dagger\,|\psi_n\rangle = \sqrt{n+1}\,|\psi_{n+1}\rangle \tag{4.11}$$

$$[\hat{a}, \hat{a}^\dagger] = 1 \tag{4.12}$$

$$\hat{a}^\dagger \hat{a} = \hat{n} \tag{4.13}$$

We can express $\boldsymbol{\pi}_\parallel$ in terms of these operators:

$$\pi_x = \frac{1}{\sqrt{2}\,\ell_B}\,(\hat{a}^\dagger + \hat{a})\,, \qquad \pi_y = \frac{1}{\sqrt{2}\,\ell_B}\,(\hat{a}^\dagger - \hat{a}) \tag{4.14}$$

such that

$$\boldsymbol{\pi}_\parallel{}^2 = \frac{2}{\ell_B^2}\left(\hat{n} + \frac{1}{2}\right) \tag{4.15}$$

They can also be written as

$$\pi_\pm = \pi_x \pm i\,\pi_y \tag{4.16}$$

where

$$\pi_+ = \frac{\sqrt{2}}{\ell_B}\,\hat{a}^\dagger\,, \qquad \pi_- = \frac{\sqrt{2}}{\ell_B}\,\hat{a} \tag{4.17}$$

Substituting Eq. 4.14 and 4.15 into Eq. 4.7 gives the non-interacting Hamiltonian in terms of the raising and lowering operators:

$$h'(k_z, \hat{a}, \hat{a}^\dagger) = \epsilon'_{\boldsymbol{\pi}_\parallel, k_z}\,\mathbb{1} + \mathcal{M}'(\boldsymbol{\pi}_\parallel, k_z)\Gamma_5 + B_0\,k_z\Gamma_4 + \frac{A_0}{\sqrt{2}\,\ell_B}\,((\hat{a}^\dagger - \hat{a})\Gamma_1 - (\hat{a}^\dagger + \hat{a})\Gamma_2) \tag{4.18}$$

where

$$\epsilon'_{\pi_\parallel, k_z} = C_0 + C_1\, k_z^2 + \frac{2}{\ell_B^2} C_2 \left(\hat{n} + \frac{1}{2}\right) \tag{4.19}$$

$$\mathcal{M}'(\pi_\parallel, k_z) = M_0 + M_1\, k_z^2 + \frac{2}{\ell_B^2} M_2 \left(\hat{n} + \frac{1}{2}\right) \tag{4.20}$$

We want to solve this Hamiltonian for n and B in a similar range as those observed in the experimental measurements described in Sec. 1.2.4. That is, for $0 \le n \le 10$ and $0 \le B \le 20$ T. The $n = 0$ and $n > 0$ LLs must be treated separately, but the same Hamiltonian, Eq. 4.18, is used in both approaches.

For $n > 0$, we solve the Schrödinger equation, $h'(k_z, \hat{a}, \hat{a}^\dagger)\, \Psi_{n>0} = E_n\, \Psi_{n>0}$, with the eigenvector

$$\Psi_{n>0} = \begin{pmatrix} c_1\, \psi_{n-1} \\ c_2\, \psi_{n-1} \\ c_3\, \psi_n \\ c_4\, \psi_n \end{pmatrix} \tag{4.21}$$

where the $c_{1\to4}$ are probability amplitudes to be determined and ψ_n is the eigenvector for the nth LL. After acting with $h'(k_z, \hat{a}, \hat{a}^\dagger)$ on $\Psi_{n>0}$, the ψ_n's can be cancelled on both sides of the equation, leaving

$$h^{n>0}(k_z) \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = E_n \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \tag{4.22}$$

where the non-interacting effective Hamiltonian is

$$
h^{n>0}(k_z) = \begin{pmatrix}
\tilde{C}_{-\frac{1}{2}} + \tilde{M}_{-\frac{1}{2}} & -i\, B_0\, k_z & 0 & i\, \frac{\sqrt{2n}}{\ell_B}\, A_0 \\
i\, B_0\, k_z & \tilde{C}_{-\frac{1}{2}} - \tilde{M}_{-\frac{1}{2}} & i\, \frac{\sqrt{2n}}{\ell_B}\, A_0 & 0 \\
0 & -i\, \frac{\sqrt{2n}}{\ell_B}\, A_0 & \tilde{C}_{+\frac{1}{2}} + \tilde{M}_{+\frac{1}{2}} & -i\, B_0\, k_z \\
-i\, \frac{\sqrt{2n}}{\ell_B}\, A_0 & 0 & i\, B_0\, k_z & \tilde{C}_{+\frac{1}{2}} - \tilde{M}_{+\frac{1}{2}}
\end{pmatrix}
\tag{4.23}
$$

and

$$
\tilde{C}_{\pm\frac{1}{2}} = C_0 + C_1\, k_z^2 + \frac{2}{\ell_B^2} C_2 \left(n \pm \frac{1}{2} \right)
\tag{4.24}
$$

$$
\tilde{M}_{\pm\frac{1}{2}} = M_0 + M_1\, k_z^2 + \frac{2}{\ell_B^2} M_2 \left(n \pm \frac{1}{2} \right)
\tag{4.25}
$$

Following the steps described in Sec. 3.1.1, we periodize Eq. 4.23 using Eq. 3.9 and 3.10, and then Fourier transform the result from k_z to z using the form of Eq. 3.15. This yields a tight-binding Hamiltonian of the same form as Eq. 3.16:

$$
\mathcal{H}^{n>0} = \sum_{n=1}^{\infty} \sum_z \left[\hat{c}_{nz\alpha}^\dagger\, h_{\alpha\beta}^{n>0}\, \hat{c}_{nz\beta} - \left(t^{n>0} \right)_{\alpha\beta}^\dagger\, \hat{c}_{n,z+1\alpha}^\dagger\, \hat{c}_{nz\beta} - t_{\alpha\beta}^{n>0}\, \hat{c}_{nz\alpha}^\dagger\, \hat{c}_{n,z+1,\beta} \right]
\tag{4.26}
$$

where the on-site Hamiltonian is

$$
h^{n>0} = \begin{pmatrix}
\bar{C}_{-\frac{1}{2}} + \bar{M}_{-\frac{1}{2}} & 0 & 0 & i\, \frac{\sqrt{2n}}{\ell_B}\, A_0 \\
0 & \bar{C}_{-\frac{1}{2}} - \bar{M}_{-\frac{1}{2}} & i\, \frac{\sqrt{2n}}{\ell_B}\, A_0 & 0 \\
0 & -i\, \frac{\sqrt{2n}}{\ell_B}\, A_0 & \bar{C}_{+\frac{1}{2}} + \bar{M}_{+\frac{1}{2}} & 0 \\
-i\, \frac{\sqrt{2n}}{\ell_B}\, A_0 & 0 & 0 & \bar{C}_{+\frac{1}{2}} - \bar{M}_{+\frac{1}{2}}
\end{pmatrix}
\tag{4.27}
$$

and

$$
\bar{C}_{\pm\frac{1}{2}} = C_0 + 2C_1 + \frac{2}{\ell_B^2} C_2 \left(n \pm \frac{1}{2} \right)
\tag{4.28}
$$

$$
\bar{M}_{\pm\frac{1}{2}} = M_0 + 2M_1 + \frac{2}{\ell_B^2} M_2 \left(n \pm \frac{1}{2} \right)
\tag{4.29}
$$

and the nearest-neighbour hopping in the z direction is

$$
\left(t^{n>0}\right)^{\dagger}_{\alpha\beta} =
\begin{pmatrix}
C_1 + M_1 & -\frac{B_0}{2} & 0 & 0 \\
\frac{B_0}{2} & C_1 - M_1 & 0 & 0 \\
0 & 0 & C_1 + M_1 & -\frac{B_0}{2} \\
0 & 0 & \frac{B_0}{2} & C_1 - M_1
\end{pmatrix}
\tag{4.30}
$$

which is the same as Eq. 3.18.

We solve for the $n = 0$ effective Hamiltonian using the same process as for the $n > 0$ effective Hamiltonian. We solve the Schrödinger equation for the Hamiltonian in Eq. 4.18, $h'(k_z, \hat{a}, \hat{a}^{\dagger}) \, \Psi_{n=0} = E_0 \, \Psi_{n=0}$, with the following eigenvector

$$
\Psi_{n=0} =
\begin{pmatrix}
0 \\
0 \\
c_1 \, \psi_0 \\
c_2 \, \psi_0
\end{pmatrix}
\tag{4.31}
$$

and we obtain

$$
h^{n=0}(k_z)
\begin{pmatrix}
0 \\
0 \\
c_1 \\
c_2
\end{pmatrix}
= E_0
\begin{pmatrix}
0 \\
0 \\
c_1 \\
c_2
\end{pmatrix}
\tag{4.32}
$$

We can write this equation as a 2×2 matrix equation, where

$$
h^{n=0}(k_z) = \left(C_0 + C_1 \, k_z^2 + \frac{1}{\ell_B^2} C_2 \right) \mathbb{1} + \left(M_0 + M_1 \, k_z^2 + \frac{1}{\ell_B^2} M_2 \right) \sigma_3 + B_0 \, k_z \, \sigma_2
\tag{4.33}
$$

The bandstructures of Eq. 4.23 and 4.33 (i.e., for a slab without boundaries) are given in Fig. 4.1. As seen in Ch. 3, this Hamiltonian describes a band insulator. As was done in Ch. 3, only results for Bi_2Se_3 are shown in this chapter. By contrast with the purely 2D case, here the LLs disperse with k_z. These are bulk LLs. In the right panel of Fig. 4.1 we

see that some of the $k_z = 0$ LLs have the "wrong" magnetic field dependence, i.e., some of the conduction band LLs disperse downwards while some of the valence band LLs disperse upwards. This is a consequence of band inversion and has been observed in 2D TIs as well [77].
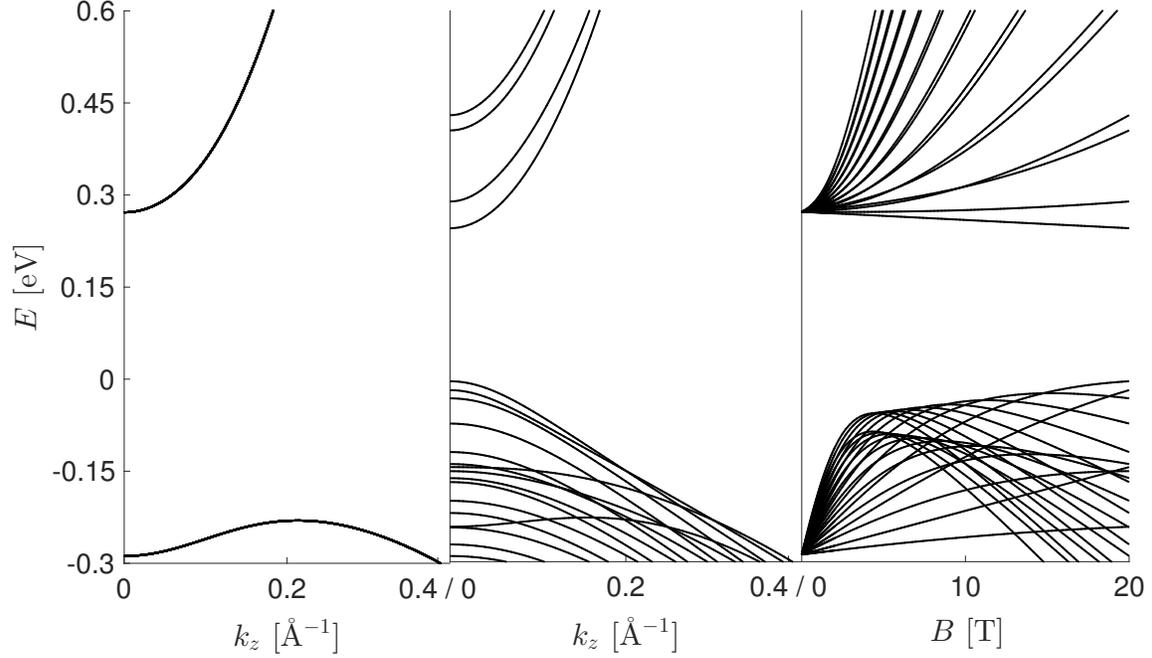
Figure 4.1: The bandstructure for infinite z obtained by diagonalizing Eq. 4.23 and 4.33 at $B = 0$ T (left) and B = 20 T (centre). (Right) At $k_z = 0$ Å^{-1}, the eigenvalues of Eq. 4.23 and 4.33 plotted as a function of the magnetic field. In all three plots, $0 \leq n \leq 10$. This figure reproduces results from Ref. [72].

We periodize Eq. 4.33 using Eq. 3.9 and 3.10, and then Fourier transform using the form of Eq. 3.15. This yields a tight-binding Hamiltonian of the same form as Eq. 3.16:

$$\mathcal{H}^{n=0} = \sum_z \left[\hat{c}^\dagger_{nz\alpha} h^{n=0}_{\alpha\beta} \hat{c}_{nz\beta} - \left(t^{n=0} \right)^\dagger_{\alpha\beta} \hat{c}^\dagger_{n,z+1\alpha} \hat{c}_{nz\beta} - t^{n=0}_{\alpha\beta} \hat{c}^\dagger_{nz\alpha} \hat{c}_{n,z+1,\beta} \right] \tag{4.34}$$

where

$$h^{n=0} = \left(C_0 + 2C_1 + \frac{1}{\ell_B^2} C_2 \right) \mathbb{1} + \left(M_0 + 2M_1 + \frac{1}{\ell_B^2} M_2 \right) \sigma_3 \tag{4.35}$$

and

$$\left(t^{n=0}\right)^{\dagger} = C_1 \, \mathbb{1} + M_1 \, \sigma_3 - i\frac{B_0}{2} \, \sigma_2 \qquad (4.36)$$

The energy spectrum as a function of the magnetic field strength for slabs of finite thickness, N_z, and their corresponding probability densities are shown in Fig. 4.2 and 4.3, respectively, for the tight-binding Hamiltonian in Eq. 4.26 and 4.34. Only LLs for $0 \leq n \leq 10$ are plotted here; however, in practice LLs with a higher n are occupied. In this section, we include results for $N_z = 20$, which does not have well-defined SSs, as a point of comparison for the thicker slabs. We added a small perturbative potential to the surface layer of the Hamiltonian so as to break the reflection symmetry of the slab. This causes the SSs to localize on either the top or bottom surface (see Fig. 4.3).

The states in blue in each figure are SSs, which is easily seen in the probability density plots in Fig. 4.3. For the thinner slabs, a few SSs appear in the bulk conduction band. However, once the slab is sufficiently thick, Fig. 4.2d, we see that all of the SSs are located within the bulk bandgap. As with the bandstructure plots in Fig. 3.2, each state is doubly degenerate: one state for the top surface, and one for the bottom. For smaller N_z this degeneracy is split. This is best illustrated in Fig. 4.2a where the zeroth LLs (the states with a flat dispersion) are approximately 0.15 eV from each other. However, in the $N_z = 200$ plot, the higher energy state has "converged" towards the lower energy state, which occurs at the DP. There are twelve distinct SSs seen in each plot. Although the total number of states increases with slab thickness, the number of SSs does not; the additional states are bulk states. As N_z increases, we also see the bulk states approaching the edges of the bulk valence and conduction bands at $B = 20$ T.

In Fig. 4.3 we see that as the thickness increases, the SSs become more localized. This is evident in two ways. First, the relative probability densities of the blue curves near the surface compared to the bulk layers becomes much greater. Second, in the thinner slabs, the probability densities are "M" shaped. As the thickness increases, the probability density of one of the "humps" decreases (this is best seen in Fig. 4.3c), such that at $N_z = 200$, the

81

state can essentially only exist on one (either the top or the bottom) surface. This is what is seen in Fig. 3.3.

(a) $N_z = 20$ Layers

(b) $N_z = 40$ Layers

(c) $N_z = 60$ Layers

(d) $N_z = 200$ Layers
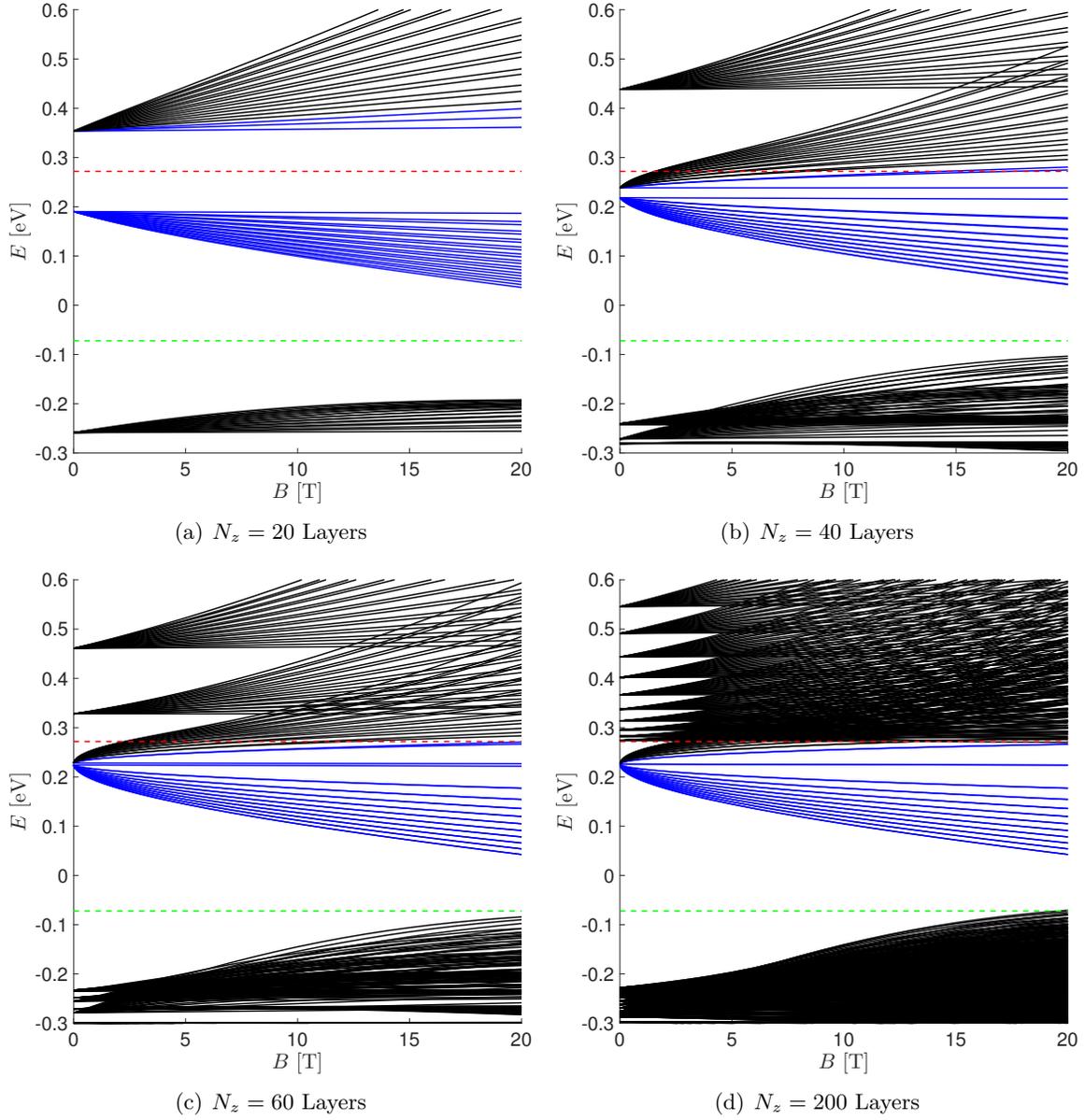
Figure 4.2: The energy spectrum of Eq. 4.26 and 4.34 for various thicknesses, N_z, and for LL index $0 \leq n \leq 10$. The red and green lines denote the bottom of the conduction band and top of the valence band, respectively. The SSs are plotted in blue, while the bulk states are plotted in black.

(a) $N_z = 20$ Layers

(b) $N_z = 40$ Layers

(c) $N_z = 60$ Layers

(d) $N_z = 200$ Layers
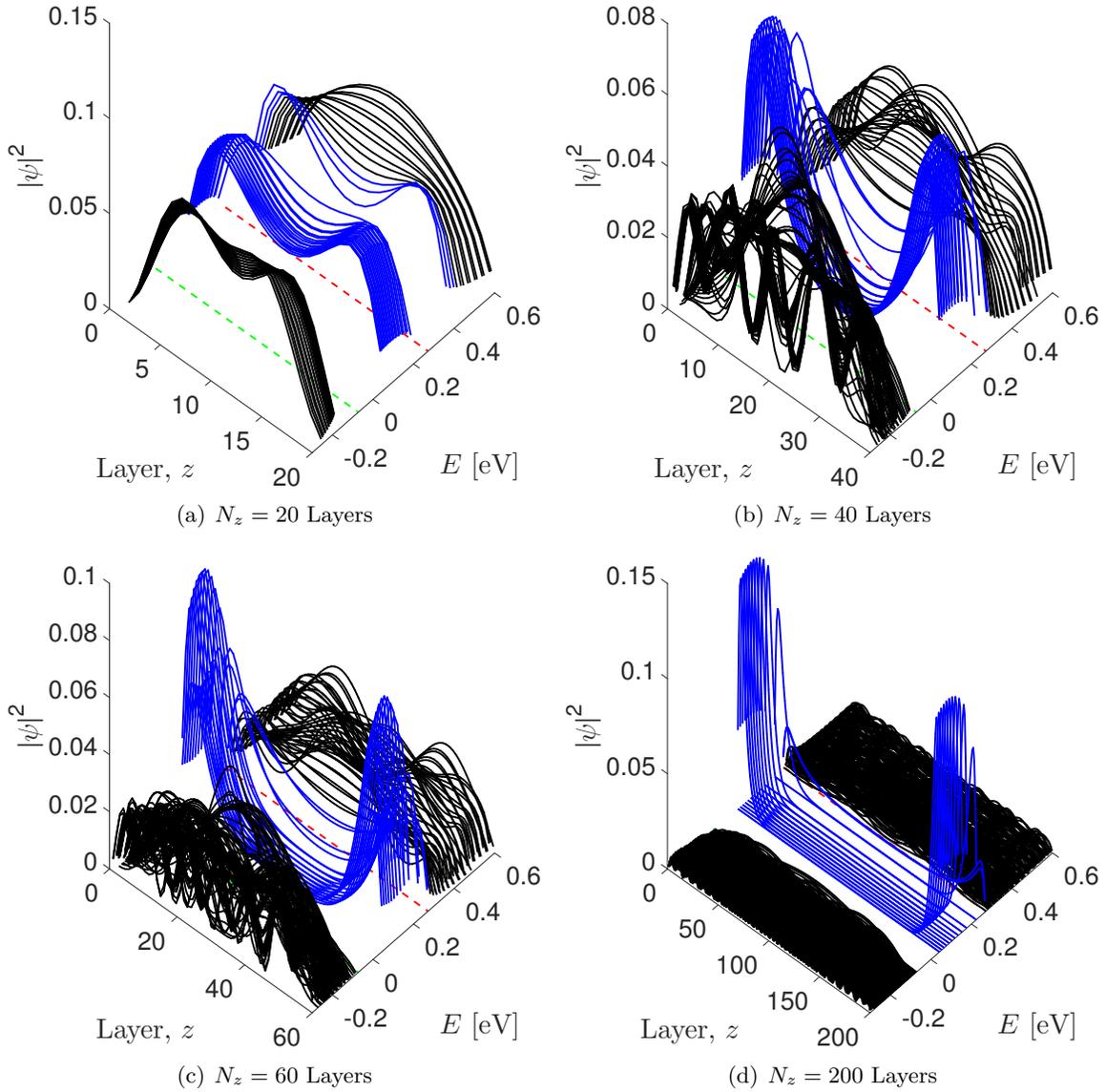
Figure 4.3: The probability densities of the states plotted in Fig. 4.2 for $B = 20$ T. Though difficult to see, the red and green lines denote the bottom of the conduction band and top of the valence band, respectively.

We plot the non-interacting surface DOS in Fig. 4.4, which is given by

$$\rho(\omega) = \sum_{n=0}^{\infty} A^{\alpha\beta}_{z,z'=1}(n,\omega)\,\delta_{\alpha\beta} \qquad (4.37)$$

where $A^{\alpha\beta}_{zz'}(n,\omega)$ is the non-interacting spectral function in the presence of a magnetic field. Although in principle the sum over LL index, n, goes to infinity, for a given magnetic field strength and energy window, only a finite number of LLs contribute and we can truncate the sum (for us, $\max(n) = 35$ was sufficient). As observed in Fig. 1.12a–b (the Bi_2Se_3 samples), the surfaces of 3D TIs support LL peaks in the DOS. However, there are a couple differences between Fig. 4.4 and Fig. 1.12a–b. First, since this is a non-interacting model, there is no broadening of the LLs away from E_F and E_{DP}. By Eq. 2.31, one might expect perfect delta function peaks in the DOS; however, the peaks in Fig. 4.4 have a finite width, η, that arises from the numerical calculation of the non-interacting GF (Eq. 2.35). Second, the negative LLs ($n < 0$) are much easier to distinguish here than in the experimental data. In the experimental data, no negative LLs are observed due to the DP lying near the valence band; thus the bulk states overlapped with the surface LLs [57].

In Fig. 4.4a the minimum near E_C occurs because no DC forms for so few layers (recall the bandstructure in Fig. 3.2a). This is not seen in the other three plots. It is easier to see positive LLs for the thinner samples in Fig. 4.4 because with so few layers, the bulk layers act as stacked quasi-2D systems, and can thus support bulk LL quantization. In addition, there are fewer bulk conduction band states to interfere with the $n > 0$ surface LLs. Conversely, as the slab thicknesses increase they act more and more as true 3D systems and only the SSs can support true LLs. For $N_z \geq 40$, only the first positive LL is easily distinguishable, as the rest of the positive LLs lie in the bulk conduction band. On the other hand, the negative LLs are SSs that lie in the bulk bandgap and thus are not interfered with by bulk states. The LLs are more distinct with increasing magnetic field, regardless of the slab thickness. The DOS in the bulk bandgap does not qualitatively change for $N_z \geq 60$. In fact, the only noticeable difference between the $N_z = 40$ and $N_z > 40$ plots is the location of the $n = 0$ LL. Since the $n = 0$ LL occurs at the DP, and the DP is gapped for $N_z = 40$ and gapless for $N_z = 60$, the DPs, and therefore the $n = 0$ LLs, of these slabs occur at

different energies. The fact that the $B = 5$ T curves have the same general shape as the non-interacting DOS curve in Fig. 3.5 suggests that if we could set $B = 0$ T (which does not provide physical results in this formalism), we would return to the zero-magnetic field results. In Fig. 4.4 and 4.5, results for $N_z \geq 60$ yield qualitatively the same results as the DCs for these thicknesses, and therefore their Fermi surfaces as well, are virtually the same. As with the non-interacting results from Sec. 3.2, the numerical evaluation of the DOS resulted in quadratic scaling with N_z. The number of n and ω values resulted in a linear time scaling.

(a) $N_z = 20$ Layers

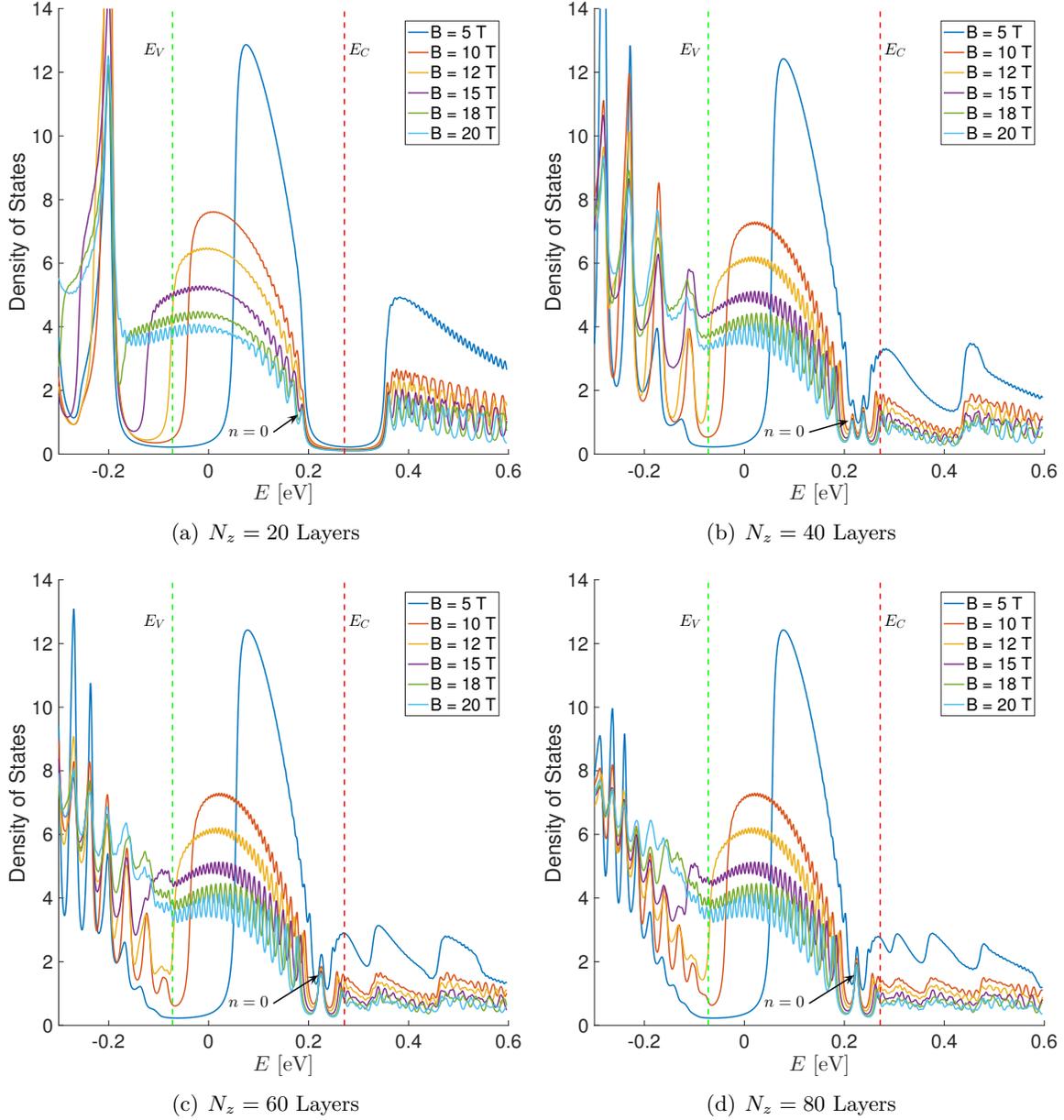(b) $N_z = 40$ Layers

(c) $N_z = 60$ Layers

(d) $N_z = 80$ Layers

Figure 4.4: The surface DOS near the bulk bandgap for $0 \leq n \leq 35$. Clearly, the LLs are most easily distinguishable in the bulk bandgap.

The energies where the LL peaks occur are plotted against $\sqrt{|n|B}$ for $-15 \leq n \leq 0$ in Fig. 4.5. Only LLs in this range are included here due to higher LL peaks being difficult to discern under lower applied magnetic fields. In addition, at larger $\sqrt{|n|B}$ the peaks begin to deviate from the expected linear behaviour. $B < 10$ T data was not used at it is too difficult to accurately locate the LL peaks. The $N_z = 20$ slab is not thick enough to support well-defined SSs, which is the reason for its far more dispersive results. The only noticeable difference between the $N_z = 40$ plot and the $N_z = 60$ plot is the energy of the $n = 0$ LL, which is slightly lower in energy for $N_z = 40$. The slight deviation from linearity is likely due to the Dirac dispersions of the SSs (see Fig. 3.2) not being perfectly linear. Although Jiang *et al.* attributed this non-linearity to a tip-gating effect [58], we propose an alternate explanation that the lack of a perfectly linear Dirac dispersion could cause these results. This is because our deviation from linearity is not caused by any tip-related effects. Note that the zeroth LL (the value at $\sqrt{|n|B} = 0$) does not change with B, as expected for surface LLs. The presence of the zeroth LL and the linear $\sqrt{|n|B}$ dependence are clear indicators that the TI SSs obey a relativistic LL spectrum.

We can extract the Fermi velocity from the plots in Fig. 4.5 using the slope of the linear fit and Eq. 4.4. For $N_z = 80$, we obtain $|v_F| = 1.5064 \times 10^5$ m/s. Comparing to experiments, Xia *et al.* measured the Fermi velocity of Bi_2Se_3 to be 5×10^5 m/s using ARPES [35] and Cheng *et al.* measured $v_F = 3.4 \times 10^5$ m/s using a STM [57]. Liu *et al.* predict that their Hamiltonian (similar to Eq. 3.2), yields $v_F = 5 \times 10^5$ m/s. These different values do not precisely match but are within a factor of four of each other. Furthermore, v_F is not uniquely defined but depends on the Fermi level (which varies from compound to compound), since the dispersion is not strictly linear.

(a) $N_z = 20$ Layers

(b) $N_z = 40$ Layers

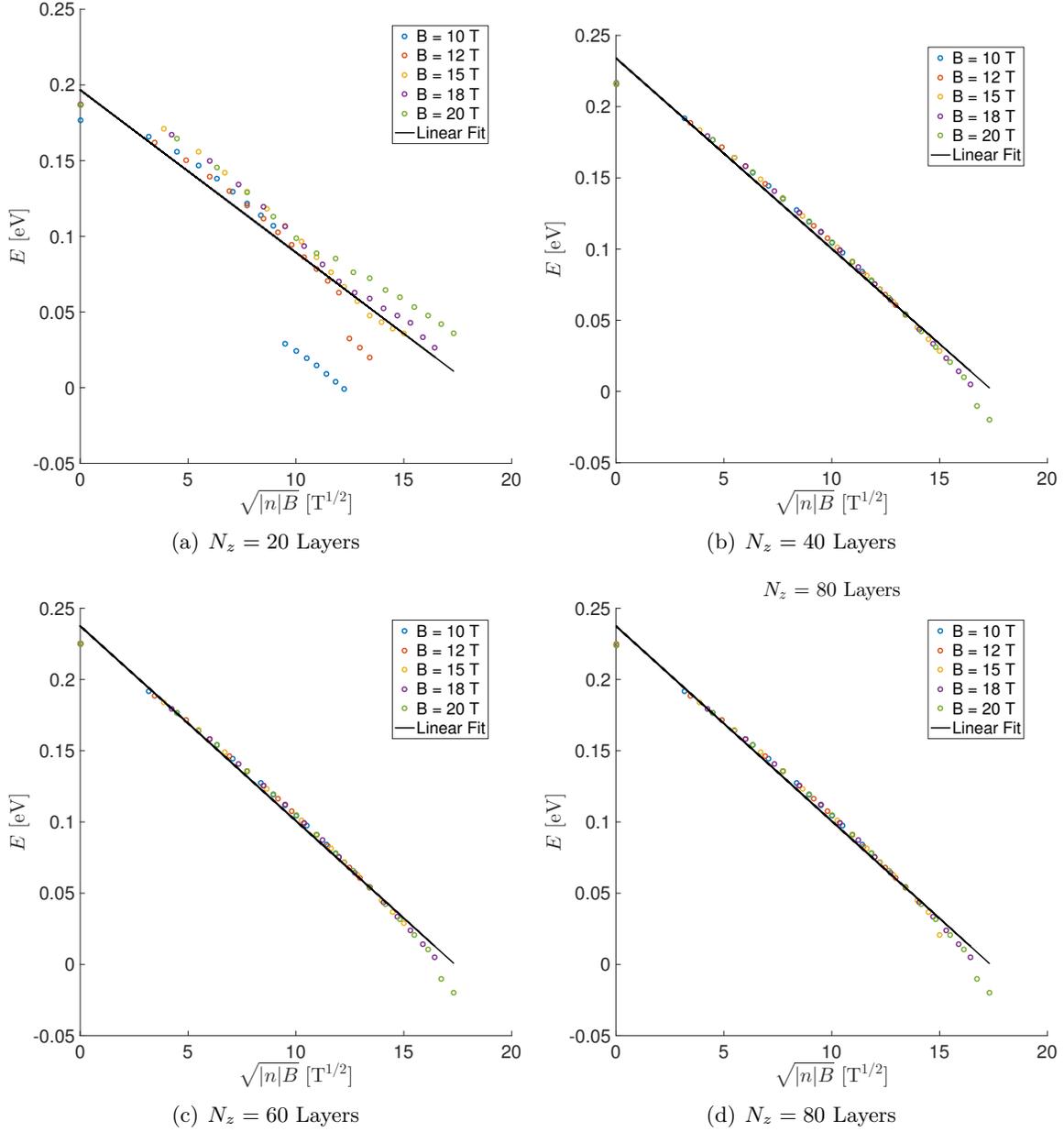(c) $N_z = 60$ Layers

(d) $N_z = 80$ Layers

Figure 4.5: Linear behaviour of the LL peak energies plotted against $\sqrt{|n|B}$ (see Eq. 4.4). The deviation from linearity at larger $\sqrt{|n|B}$ is also seen in Fig. 1.13.

4.3 Interactions for a Finite Magnetic Field

To include interactions we must write Eq. 4.26 and 4.34 in second-quantized form and write the interaction term, $\hat{V}$, in the LL basis. We might expect that we can simply write the second-quantized Hamiltonian as

$$\mathcal{H}_0 = \sum_{zz'} \sum_{n=0}^{\infty} c_{nz\alpha}^{\dagger} \, h_{zz'}^{\alpha\beta}(n) \, c_{nz'\beta} \tag{4.38}$$

which uses the same basis as Eq. 4.26 and 4.34. However, this neglects the degeneracy described in Eq. 4.3. This degeneracy leads to an additional quantum number, k_x, that must be summed over.

We now derive the non-interacting Hamiltonian more carefully. Once again, using the gauge $\boldsymbol{A} = -By\,\hat{x}$, $k = k_x$ is a good quantum number. In second-quantized form, the Hamiltonian is

$$\mathcal{H}_0 = \sum_{zz'} \int d^2\boldsymbol{r}_{\parallel} \, c_{\boldsymbol{r}_{\parallel}z\alpha}^{\dagger} \, h_{zz'}^{\alpha\beta}(\hat{a}, \hat{a}^{\dagger}) \, c_{\boldsymbol{r}_{\parallel}z'\beta} \tag{4.39}$$

where $h_{zz'}^{\alpha\beta}(\hat{a}, \hat{a}^{\dagger})$ is Eq. 4.18 periodized in the z direction and Fourier transformed from k_z to z. We change to the LL basis to allow the $\hat{a}, \hat{a}^{\dagger}$ to act on LL eigenstates:

$$\begin{pmatrix} c_{\boldsymbol{r}_{\parallel}z1} \\[4pt] c_{\boldsymbol{r}_{\parallel}z2} \\[4pt] c_{\boldsymbol{r}_{\parallel}z3} \\[4pt] c_{\boldsymbol{r}_{\parallel}z4} \end{pmatrix} = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{dk}{2\pi} \begin{pmatrix} c_{nkz1} \, \psi_{n-1,k}(\boldsymbol{r}_{\parallel}) \\[4pt] c_{nkz2} \, \psi_{n-1,k}(\boldsymbol{r}_{\parallel}) \\[4pt] c_{nkz3} \, \psi_{nk}(\boldsymbol{r}_{\parallel}) \\[4pt] c_{nkz4} \, \psi_{nk}(\boldsymbol{r}_{\parallel}) \end{pmatrix} \tag{4.40}$$

where

$$\psi_{nk}(\boldsymbol{r}_{\parallel}) = \left(\frac{1}{\pi \, \ell_B^2} \right)^{1/4} \frac{1}{\sqrt{2^n \, n!}} \, e^{ikx} \, e^{-(y-k\ell_B^2)^2/2\ell_B^2} \, H_n \left(\frac{y - k\ell_B^2}{\ell_B} \right) \tag{4.41}$$

and H_n are Hermite polynomials. It can be shown that this forms an orthonormal basis. We define $\psi_{-1,k}$ to be zero. The Hamiltonian can thus be written as

$$\mathcal{H}_0 = \sum_{zz'} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{dk}{2\pi} \, c_{nkz\alpha}^{\dagger} \, h_{zz'}^{\alpha\beta}(n) \, c_{nkz'\beta} \tag{4.42}$$

where $h_{zz'}^{\alpha\beta}(n)$ is either Eq. 4.26 or Eq. 4.34 depending on the LL index, n.

The interaction term, Eq. 3.26, is

$$V = U_{\gamma\delta} \sum_z \int d^2\boldsymbol{r}_{\parallel} \, c_{\boldsymbol{r}_{\parallel}z\gamma}^{\dagger} \, c_{\boldsymbol{r}_{\parallel}z\gamma} \, c_{\boldsymbol{r}_{\parallel}z\delta}^{\dagger} \, c_{\boldsymbol{r}_{\parallel}z\delta} \tag{4.43}$$

Making the same change of basis as in Eq. 4.40, and explicitly writing the sum over γ and δ, we obtain

$$V = \sum_{\gamma\delta} U_{\gamma\delta} \sum_z \sum_{n_1 \ldots n_4} \int_{-\infty}^{\infty} \frac{dk_1}{2\pi} \cdots \frac{dk_4}{2\pi} \, c_{n_1 k_1 z\gamma}^{\dagger} \, c_{n_2 k_2 z\gamma} \, c_{n_3 k_3 z\delta}^{\dagger} \, c_{n_4 k_4 z\delta}$$
$$\times \left[\int d^2\boldsymbol{r}_{\parallel} \, \psi_{n_1 k_1 \gamma}^*(\boldsymbol{r}_{\parallel}) \psi_{n_2 k_2 \gamma}(\boldsymbol{r}_{\parallel}) \psi_{n_3 k_3 \delta}^*(\boldsymbol{r}_{\parallel}) \psi_{n_4 k_4 \delta}(\boldsymbol{r}_{\parallel}) \right]$$

$$\tag{4.44}$$

where we define

$$\psi_{nk\alpha}(\boldsymbol{r}_{\parallel}) = \begin{cases} \psi_{n-1,k}(\boldsymbol{r}_{\parallel}), & \alpha = 1, 2 \\ \psi_{nk}(\boldsymbol{r}_{\parallel}), & \alpha = 3, 4 \end{cases} \tag{4.45}$$

The expression in square brackets in Eq. 4.44 is what must be determined. We define

$$A_{n_1 \ldots n_4}^{\gamma\delta}(k_1, \ldots, k_4) \equiv \int d^2\boldsymbol{r}_{\parallel} \, \psi_{n_1 k_1 \gamma}^*(\boldsymbol{r}_{\parallel}) \psi_{n_2 k_2 \gamma}(\boldsymbol{r}_{\parallel}) \psi_{n_3 k_3 \delta}^*(\boldsymbol{r}_{\parallel}) \psi_{n_4 k_4 \delta}(\boldsymbol{r}_{\parallel}) \tag{4.46}$$

To express this more simply in terms of the standard LL eigenstates, Eq. 4.41, we define a new quantity, $W_{n_1 \ldots n_4}$, that is independent of spin:

$$W_{n_1 \ldots n_4}(k_1, \ldots, k_4) \equiv \int d^2\boldsymbol{r}_{\parallel} \, \psi_{n_1 k_1}^*(\boldsymbol{r}_{\parallel}) \psi_{n_2 k_2}(\boldsymbol{r}_{\parallel}) \psi_{n_3 k_3}^*(\boldsymbol{r}_{\parallel}) \psi_{n_4 k_4}(\boldsymbol{r}_{\parallel}) \tag{4.47}$$

where

$$A^{11,12,21,22}_{n_1,n_2,n_3,n_4} = W_{n_1-1,n_2-1,n_3-1,n_4-1} \tag{4.48}$$

$$A^{13,14,23,24}_{n_1,n_2,n_3,n_4} = W_{n_1-1,n_2-1,n_3,n_4} \tag{4.49}$$

$$A^{31,32,41,42}_{n_1,n_2,n_3,n_4} = W_{n_1,n_2,n_3-1,n_4-1} \tag{4.50}$$

$$A^{33,34,43,44}_{n_1,n_2,n_3,n_4} = W_{n_1,n_2,n_3,n_4} \tag{4.51}$$

Writing Eq. 4.47 out explicity using Eq. 4.41, we obtain

$$
W_{n_1 \ldots n_4}(k_1, \ldots, k_4) = \frac{1}{\pi \ell_B^2} \frac{2^{-(n_1+n_2+n_3+n_4)/2}}{\sqrt{n_1! n_2! n_3! n_4!}} \int_{-\infty}^{\infty} dx \, e^{-i(k_1-k_2+k_3-k_4)x}
$$
$$
\times \int_{-\infty}^{\infty} dy \exp\left(-\sum_{i=1}^{4} \frac{(y - k_i \ell_B^2)^2}{2\ell_B^2} \right) \prod_{i=1}^{4} H_{n_i}\left(\frac{y - k_i \ell_B^2}{\ell_B} \right) \tag{4.52}
$$

Letting $y/\ell_B \to y$, defining $p_i \equiv k_i \ell_B$, and evaluating $\int dx$, we obtain

$$
W_{n_1 \ldots n_4}(k_1, \ldots, k_4) = \frac{1}{\pi \ell_B^2} \frac{2^{-(n_1+n_2+n_3+n_4)/2}}{\sqrt{n_1! n_2! n_3! n_4!}} 2\pi \, \ell_B \, \delta(p_1 - p_2 + p_3 - p_4)
$$
$$
\times \ell_B \int_{-\infty}^{\infty} dy \exp\left(-\frac{1}{2} \sum_{i=1}^{4} (y - p_i)^2 \right) \prod_{i=1}^{4} H_{n_i}(y - p_i) \tag{4.53}
$$

To evaluate the integral over y we express the product of two Hermite polynomials as a sum of higher order single Hermite polynomials, as was done by Na and Marsiglio [78]. This reduces the product of four Hermite polynomials appearing in Eq. 4.53 to a sum of products of only two polynomials, which can be evaluated by making use of the orthogonality of Hermite polynomials. As will become clearer later, it is most useful to cast this expansion in the form

$$
H_n(y - p) \, H_m(y - q) = \sum_l f_{nml}(p, q) \, H_l\left[\sqrt{2}\left(y - \frac{p+q}{2} \right) \right] \tag{4.54}
$$

where the argument of H_l is chosen such that we can exploit the orthogonality of Hermite polynomials. We must find the coefficients $f_{nml}(p, q)$. The procedure is somewhat different from what is done in Ref. [78], as the arguments of the two Hermite polynomials on the

left-hand side of Eq. 4.54 are different in our case, but the basic idea of the method is the same.

First, we write the Hermite polynomial generating function

$$e^{-t^2+2tx} = \sum_{n=0}^{\infty} H_n(x)\,\frac{t^n}{n!} \tag{4.55}$$

Since we want to solve the product of two Hermite polynomials with arguments $(y-p)$ and $(y-q)$, we write the product of two generating functions:

$$e^{-t^2+2t(y-p)}\,e^{-s^2+2s(y-q)} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} H_n(y-p)\,H_m(y-q)\,\frac{t^n}{n!}\frac{s^m}{m!} \tag{4.56}$$

The left-hand side (LHS) of this expression can be written as

$$\text{LHS} = e^{2\left(\frac{t+s}{\sqrt{2}}\right)\sqrt{2}\left(y-\frac{p+q}{2}\right)-\left(\frac{t+s}{\sqrt{2}}\right)^2}\,e^{-\frac{1}{2}(t-s)^2-(t-s)(p-q)} \tag{4.57}$$

where the first exponential is a generating function with different arguments. Using Eq. 4.55, we obtain

$$\text{LHS} = \sum_{l=0}^{\infty} H_l\left[\sqrt{2}\left(y-\frac{p+q}{2}\right)\right]\left(\frac{t+s}{\sqrt{2}}\right)^l \frac{1}{l!}\,e^{-\frac{1}{2}(t-s)^2-(t-s)(p-q)} \tag{4.58}$$

We must now expand this expression in powers of t and s. We use the binomial theorem to write

$$\left(\frac{t+s}{\sqrt{2}}\right)^l = \frac{1}{2^{l/2}}\sum_{u=0}^{l}\binom{l}{u}s^u\,t^{l-u} \tag{4.59}$$

Completing the square of the exponential in Eq. 4.58, yields

$$e^{-\frac{1}{2}(t-s)^2-(t-s)(p-q)} = e^{\frac{1}{2}(p-q)^2}\,e^{-\frac{1}{2}(t-s+p-q)^2} \tag{4.60}$$

and then expanding the second exponential in the above expression in a power series of $(t - s)$, gives

$$e^{-\frac{1}{2}(t-s+p-q)^2} = \sum_{r=0}^{\infty} g_r(q - p) \, (t - s)^r \tag{4.61}$$

The LHS of this equation can be written as a generating function of Hermite polynomials:

$$e^{-\frac{1}{2}(t-s+p-q)^2} = e^{-\frac{1}{2}(q-p)^2} \sum_{r=0}^{\infty} H_r \left(\frac{q - p}{\sqrt{2}} \right) \frac{(t - s)^r}{2^{r/2} \, r!} \tag{4.62}$$

and thus it is easily seen, by comparing Eq. 4.61 and 4.62,

$$g_r(q - p) = \frac{e^{-\frac{1}{2}(q-p)^2}}{2^{r/2} \, r!} H_r \left(\frac{q - p}{\sqrt{2}} \right) \tag{4.63}$$

Expanding $(t - s)^r$ using the binomial theorem, and putting it all together, we can write Eq. 4.56 as

$$\sum_{l=0}^{\infty} H_l \left[\sqrt{2} \left(y - \frac{p + q}{2} \right) \right] \frac{e^{\frac{1}{2}(p-q)^2}}{2^{l/2} \, l!} \sum_{r=0}^{\infty} g_r(q - p) \sum_{u=0}^{l} \sum_{v=0}^{r} \binom{l}{u} \binom{r}{v} (-1)^v t^{l+r-u-v} s^{u+v}$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} H_n(y - p) \, H_m(y - q) \frac{t^n}{n!} \frac{s^m}{m!} \tag{4.64}$$

Matching the powers of t and s yields

$$n = l + r - u - v \qquad m = u + v \tag{4.65}$$

We can use these two expressions to remove the sums over r and v. In addition, Eq. 4.65 gives more stringent bounds on the sums on the LHS of Eq. 4.64. Comparing Eq. 4.54 to Eq. 4.64, we see that

$$f_{nml}(q - p) = \frac{n! m!}{2^{l/2} \, l!} e^{\frac{1}{2}(q-p)^2} g_{n+m-l}(q - p) \, \eta_{nml} \tag{4.66}$$

93

which is an explicit expression for $f_{nml}(p,q)$, and

$$\eta_{nml} = \sum_{u=\max(0,l-n)}^{\min(m,l)} (-1)^{m-u} \binom{l}{u} \binom{n+m-l}{m-u} \tag{4.67}$$

Now, we make use of the results just obtained to determine the matrix elements, $W_{n_1\ldots n_4}(k_1,\ldots,k_4)$, in Eq. 4.53. We begin by substituting the product of two Hermite polynomials as sums of single Hermites, as in Eq. 4.54, where $f_{nml}(q-p)$ is given by Eq. 4.66:

$$H_{n_1}(y-p_1)\,H_{n_3}(y-p_3) = \sum_{l=0}^{n_1+n_3} f_{n_1 n_3 l}(p_3-p_1)\,H_l\left[\sqrt{2}\left(y-\frac{p_1+p_3}{2}\right)\right] \tag{4.68}$$

$$H_{n_2}(y-p_2)\,H_{n_4}(y-p_4) = \sum_{j=0}^{n_2+n_4} f_{n_2 n_4 j}(p_4-p_2)\,H_j\left[\sqrt{2}\left(y-\frac{p_1+p_3}{2}\right)\right] \tag{4.69}$$

where we have used $p_1+p_3 = p_2+p_4$ coming from the momentum-conserving delta function in Eq. 4.53. By completing the square, we can write the Gaussian in Eq. 4.53 as

$$\exp\left(-\frac{1}{2}\sum_{i=1}^{4}(y-p_i)^2\right) = e^{-2\left(y-\frac{p_1+p_3}{2}\right)^2}\,e^{-\frac{1}{4}(p_1-p_3)^2}\,e^{-\frac{1}{4}(p_2-p_4)^2} \tag{4.70}$$

Thus, using Eq. 4.69 and 4.70, we can write the the integral over y in Eq. 4.53 as

$$e^{-\frac{1}{4}(p_1-p_3)^2}\,e^{-\frac{1}{4}(p_2-p_4)^2} \sum_{l=0}^{n_1+n_3}\sum_{j=0}^{n_2+n_4} f_{n_1 n_3 l}(p_3-p_1)\,f_{n_2 n_4 j}(p_4-p_2)$$

$$\times \int dy\, e^{-2\left(y-\frac{p_1+p_3}{2}\right)^2} H_l\left[\sqrt{2}\left(y-\frac{p_1+p_3}{2}\right)\right] H_j\left[\sqrt{2}\left(y-\frac{p_1+p_3}{2}\right)\right] \tag{4.71}$$

Defining $x = \sqrt{2}\left(y-\frac{p_1+p_3}{2}\right)$, we can write the integral above as

$$\frac{1}{\sqrt{2}}\int_{-\infty}^{\infty} dx\, e^{-x^2} H_l(x)\,H_j(x) = \frac{1}{\sqrt{2}}\,2^l\,l!\,\sqrt{\pi}\delta_{lj} \tag{4.72}$$

where we have used the orthogonality relation of Hermite polynomials. Finally, substituting the coefficient in front of the integral in Eq. 4.71, and Eq. 4.72 into Eq. 4.53, we arrive at

our expression for $W_{n_1,\ldots,n_4}(k_1,\ldots,k_4)$ in Eq. 4.47, which is

$$W_{n_1,\ldots,n_4}(k_1,\ldots,k_4) = \frac{1}{\sqrt{2\,\pi}} \frac{2^{-(n_1+n_2+n_3+n_4)/2}}{\sqrt{n_1!n_2!n_3!n_4!}} 2\pi\delta(p_1+p_3-p_2-p_4)$$

$$\times\, e^{-\frac{1}{4}(p_1-p_3)^2}\, e^{-\frac{1}{4}(p_2-p_4)^2} \sum_{l=0}^{\min(n_1+n_2+n_3+n_4)} 2^l\, l!\, f_{n_1 n_3 l}(p_3-p_1)\, f_{n_2 n_4 l}(p_4-p_2) \quad (4.73)$$

We write the spin-dependent matrix element, $A^{\gamma\delta}_{n_1,\ldots,n_4}(k_1,\ldots,k_4)$, as

$$A^{\gamma\delta}_{n_1,\ldots,n_4}(k_1,\ldots,k_4) = 2\pi\,\delta(p_1+p_3-p_2-p_4)\, M^{\gamma\delta}_{n_1,\ldots,n_4}(p,p',q) \qquad (4.74)$$

where

$$M^{\gamma\delta}_{n_1,\ldots,n_4}(p,p',q) = \frac{1}{\sqrt{2\,\pi}} \frac{2^{-(n_1+n_2+n_3+n_4)/2}}{\sqrt{n_1!n_2!n_3!n_4!}}\, e^{-\frac{1}{4}(p_1-p_3)^2}\, e^{-\frac{1}{4}(p_2-p_4)^2}$$

$$\times \sum_{l=0}^{\min(n_1+n_2+n_3+n_4)} 2^l\, l!\, f_{n_1 n_3 l}(p_3-p_1)\, f_{n_2 n_4 l}(p_4-p_2) \quad (4.75)$$

is a form factor that contains information about the LL wavefunctions of the interacting particles. The delta function leaves three unconstrained momenta, p, p', q, which can be chosen as $p_1 = p' + q$, $p_2 = p'$, $p_3 = p - q$, $p_4 = p$, and we obtain

$$M^{\gamma\delta}_{n_1,\ldots,n_4}(p,p',q) = \frac{1}{\sqrt{2\,\pi}} \frac{2^{-(n_1+n_2+n_3+n_4)/2}}{\sqrt{n_1!n_2!n_3!n_4!}}\, e^{-\frac{1}{4}(p'-p+2q)^2}\, e^{-\frac{1}{4}(p'-p)^2}$$

$$\times \sum_{l=0}^{\min(n_1+n_2+n_3+n_4)} 2^l\, l!\, f_{n_1 n_3 l}(p-p'-2q)\, f_{n_2 n_4 l}(p-p') \quad (4.76)$$

Substituting Eq. 4.74 into Eq. 4.44 gives

$$V = \sum_{\gamma\delta} U_{\gamma\delta} \sum_{z} \sum_{n_1\ldots n_4} \int_{-\infty}^{\infty} \frac{dp}{2\pi}\frac{dp'}{2\pi}\frac{dq}{2\pi}\, M^{\gamma\delta}_{n_1,\ldots,n_4}(p,p',q)\, c^{\dagger}_{n_1,p'+q,z,\gamma}\, c_{n_2 p' z\gamma}\, c^{\dagger}_{n_3,p-q,z,\delta}\, c_{n_4 p z\delta}$$

$$(4.77)$$

In the interaction picture we obtain

$$\hat{V}(\tau_1) = \sum_{\gamma\delta} U_{\gamma\delta} \sum_z \sum_{n_1...n_4} \int_{-\infty}^{\infty} \frac{dp}{2\pi} \frac{dp'}{2\pi} \frac{dq}{2\pi} \, M_{n_1,...,n_4}^{\gamma\delta}(p,p',q)$$

$$\times \, \hat{c}_{n_1,p'+q,z,\gamma}^{\dagger}(\tau_1) \, \hat{c}_{n_2 p' z \gamma}(\tau_1) \, \hat{c}_{n_3,p-q,z,\delta}^{\dagger}(\tau_1) \, \hat{c}_{n_4 p z \delta}(\tau_1) \quad (4.78)$$

which, aside from the form factor, the sum over LLs, and the momenta being only in 1D, is essentially of the same form as Eq. 3.29. Thus, we can carry out Wick contractions for Eq. 4.78 and follow the steps laid out in Sec. 3.4 to obtain the second order self-energy for a finite magnetic field.

4.3.1 Second Order Self-Energy for a Magnetic Field

The non-interacting Matsubara GF in the LL basis can be written as

$$\mathcal{G}_{zz',nn'}^{(0)\alpha\beta}(k,k',\tau-\tau') = -\left\langle T_\tau \left[\hat{c}_{nkz\alpha}(\tau) \, \hat{c}_{n'k'z'\beta}^{\dagger}(\tau') \right] \right\rangle \quad (4.79)$$

and due to translational invariance in the x direction of the Hamiltonian in the Landau gauge, similar to Eq. 3.31, the GF is diagonal in k:

$$\mathcal{G}_{zz',nn'}^{(0)\alpha\beta}(k,k',\tau-\tau') = \mathcal{G}_{zz',nn'}^{(0)\alpha\beta}(k,\tau-\tau') \, 2\pi \, \delta(k-k') \quad (4.80)$$

and since the non-interacting Hamiltonian, Eq. 4.42, is also diagonal in n:

$$\mathcal{G}_{zz',nn'}^{(0)\alpha\beta}(k,\tau-\tau') = \mathcal{G}_{zz'}^{(0)\alpha\beta}(n,k,\tau-\tau') \, \delta_{nn'} \quad (4.81)$$

Moving to the Matsubara frequency domain, we can calculate the non-interacting Matsubara GF from

$$\mathcal{G}_{zz'}^{(0)\alpha\beta}(n,ik_m) = (ik_m - h(n))_{z\alpha,z'\beta}^{-1} \quad (4.82)$$

where, once again, $h(n)$ is either Eq. 4.26 or Eq. 4.34 depending on the LL index, n. Note that in this basis the non-interacting Matsubara GF has no dependence on k. As such, the spectral functions in the self-energy will not have any dependence on k either.

The second order correction to the GF in the LL basis is

$$\mathcal{G}_{zz'}^{(2)\alpha\beta}(n,\tau) = \frac{-1}{2!} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \left\langle T_\tau \left[\hat{V}(\tau_1)\hat{V}(\tau_2)\, \hat{c}_{nkz\alpha}(\tau)\, \hat{c}_{nkz'\beta}^\dagger(0) \right] \right\rangle_0 \qquad (4.83)$$

which is similar to Eq. 3.44. As mentioned in the previous section, the only differences between this equation and Eq. 3.44 are the $\boldsymbol{k}_\parallel$ are "replaced" by the quantum numbers appropriate in a magnetic field, n and k, there is a sum over LLs, and there is an additional factor, the form factor, in the interaction terms. Following the same steps as those described in Sec. 3.4.1 and 3.4.2, and using the properties in Eq. 4.80 and 4.81, we can obtain the second order self-energy expressions for a finite magnetic field. We only give the final results here due to the similarities in calculating the self-energy with and without a magnetic field. The pair-bubble retarded second order self-energy for a finite magnetic field is (analogous to Eq. 3.61)

$$\Sigma_{z_1 z_2}^{R(2a)\gamma\mu}(n,\omega) =$$
$$4\, U_{\gamma\delta}\, U_{\mu\nu} \sum_{n_1 n_2 n_3} \int \frac{dp}{2\pi} \int \frac{dq}{2\pi}\, M_{n,n_1,n_2,n_3}^{\gamma\delta}(p,k-q,q)\, M_{n_1,n,n_3,n_2}^{\mu\nu}(p-q,k,-q)$$
$$\times \int \frac{d\epsilon_1}{2\pi} A_{z_1 z_2}^{(0)\gamma\mu}(n_1,\epsilon_1) \int \frac{d\epsilon_2}{2\pi} A_{z_1 z_2}^{(0)\delta\nu}(n_3,\epsilon_2) \int \frac{d\epsilon_3}{2\pi} A_{z_2 z_1}^{(0)\nu\delta}(n_2,\epsilon_3)$$
$$\times \left(\frac{(n_F(\epsilon_2) - n_F(\epsilon_3))(n_F(\epsilon_1) + n_B(\epsilon_3 - \epsilon_2))}{\omega + i\delta - \epsilon_1 - \epsilon_2 + \epsilon_3} \right) \quad (4.84)$$

and for the double exchange diagram it is (analogous to Eq. 3.62)

$$\Sigma_{z_1 z_2}^{R(2b)\gamma\mu}(n,\omega) =$$
$$-4\, U_{\gamma\delta}\, U_{\mu\nu} \sum_{n_1 n_2 n_3} \int \frac{dp}{2\pi} \int \frac{dq}{2\pi} M_{n,n_1,n_2,n_3}^{\gamma\delta}(p+q,k-q,q)\, M_{n_1,n_2,n_3,n}^{\mu\nu}(k,p,k-p-q)$$
$$\times \int \frac{d\epsilon_1}{2\pi} A_{z_1 z_2}^{(0)\gamma\nu}(n_1,\epsilon_1) \int \frac{d\epsilon_2}{2\pi} A_{z_2 z_1}^{(0)\nu\delta}(n_2,\epsilon_2) \int \frac{d\epsilon_3}{2\pi} A_{z_1 z_2}^{(0)\delta\mu}(n_3,\epsilon_3)$$
$$\times \left(\frac{(n_F(\epsilon_3) - n_F(\epsilon_2))(n_B(\epsilon_2 - \epsilon_3) + n_F(\epsilon_1))}{\omega + i\delta - \epsilon_1 + \epsilon_2 - \epsilon_3} \right) \quad (4.85)$$

It is possible to evaluate $\int \frac{dp}{2\pi} \int \frac{dq}{2\pi}$ analytically as we can write the product of the form factors in terms of Hermite polynomials and a Gaussian, from which we can exploit the orthogonality property of the Hermite polynomials. Performing this calculation yields

$$
\int \frac{dp}{2\pi} \int \frac{dq}{2\pi} \ M^{\gamma\delta}_{n,n_1,n_2,n_3}(p, k-q, q) \, M^{\mu\nu}_{n_1,n,n_3,n_2}(p-q, k, -q) =
$$

$$
\frac{1}{(2\pi)^2} \frac{n! n_1! n_2! n_3!}{2^{(n+n_1+n_2+n_3)}} \sum_{j=0}^{\max(n+n_1+n_2+n_3)} \frac{\eta_{nn_2 j} \, \eta_{n_1 n_3 j}}{(j!)^2 (n+n_2-j)!(n_1+n_3-j)!} \quad (4.86)
$$

for the form factors in Eq. 4.84 and $\eta_{nn_2 j}, \eta_{n_1 n_3 j}$ are defined as in Eq. 4.67. Coincidentally, this is the same result for $\int \frac{dp}{2\pi} \int \frac{dq}{2\pi} M^{\gamma\delta}_{n,n_1,n_2,n_3}(p+q, k-q, q) \, M^{\mu\nu}_{n_1,n_2,n_3,n}(k, p, k-p-q)$ in Eq. 4.85. Note that this expression is independent of momentum.

Using Eq. 4.86, we can write the broadening function for the pair-bubble diagram as

$$
\Gamma^{(a)\gamma\mu}_{z_1 z_2}(n, \omega) =
$$

$$
\frac{2\, U_{\gamma\delta}\, U_{\mu\nu}}{\pi^2} \sum_{n_1 n_2 n_3} \frac{n! n_1! n_2! n_3!}{2^{(n+n_1+n_2+n_3)}} \sum_{j=0}^{\max(n+n_1+n_2+n_3)} \frac{\eta_{nn_2 j} \, \eta_{n_1 n_3 j}}{(j!)^2 (n+n_2-j)!(n_1+n_3-j)!}
$$

$$
\times \int \frac{d\epsilon_1}{2\pi} A^{(0)\gamma\mu}_{z_1 z_2}(n_1, \epsilon_1) \int \frac{d\epsilon_2}{2\pi} A^{(0)\delta\nu}_{z_1 z_2}(n_3, \epsilon_2) \, A^{(0)\nu\delta}_{z_2 z_1}(n_2, \epsilon_1 + \epsilon_2 - \omega)
$$

$$
\times \, (n_F(\epsilon_2) - n_F(\epsilon_1 + \epsilon_2 - \omega))(n_F(\epsilon_1) + n_B(\epsilon_1 - \omega)) \quad (4.87)
$$

and for the double exchange diagram we have

$$
\Gamma^{(b)\gamma\mu}_{z_1 z_2}(n, \omega) =
$$

$$
-\frac{2\, U_{\gamma\delta}\, U_{\mu\nu}}{\pi^2} \sum_{n_1 n_2 n_3} \frac{n! n_1! n_2! n_3!}{2^{(n+n_1+n_2+n_3)}} \sum_{j=0}^{\max(n+n_1+n_2+n_3)} \frac{\eta_{nn_2 j} \, \eta_{n_1 n_3 j}}{(j!)^2 (n+n_2-j)!(n_1+n_3-j)!}
$$

$$
\times \int \frac{d\epsilon_1}{2\pi} A^{(0)\gamma\nu}_{z_1 z_2}(n_1, \epsilon_1) \int \frac{d\epsilon_2}{2\pi} A^{(0)\nu\delta}_{z_2 z_1}(n_2, \epsilon_2) \, A^{(0)\delta\mu}_{z_1 z_2}(n_3, \omega - \epsilon_1 + \epsilon_2)
$$

$$
\times \, (n_F(\omega - \epsilon_1 + \epsilon_2) - n_F(\epsilon_2))(n_F(\epsilon_1) + n_B(\epsilon_1 - \omega) \quad (4.88)
$$

As was mentioned in Sec. 3.4.4, numerically evaluating the expressions for the broadening function give the qualitative behaviour of the quasiparticle lifetime. In order to obtain a quantitative measure of LL broadening, the Hermitian part of the self-energy is also needed

so as to calculate the interacting spectral function. Once this is calculated, the DOS can be plotted and the width of the LL peaks can be extracted.

Chapter 5

Conclusion

In this final chapter we summarize the work that has been accomplished and suggest a few directions for future research. Before examining the effects of electron-electron interactions and magnetic fields on the SSs of 3D TIs, which was motivated by a number of recent experiments, we needed to understand the SSs in the non-interacting limit. First, we plotted the bandstructures for various slab thicknesses. We observed that a DC forms for thicknesses greater than approximately 40 layers, and the gap closes at the DP at approximately 60 layers. Given that for our choice of effective lattice constant 60 layers corresponds roughly to six physical QLs, this is in good qualitative agreement with experiments on TI thin films (see Fig. 1.9). In addition, the states on the DC are localized to the surface, while all other states are found anywhere in the TI slab. Moreover, the spectral function was found to closely resemble the TI bandstructure and the surface DOS was approximately linear for energies near the DP, as expected.

To include electron-electron interactions in the model, we first calculated the first order self-energy. As predicted by theory, it does not affect the quasiparticle lifetime. However, the second order self-energy does due to it having an anti-Hermitian part. It was too computationally intensive to calculate the full self-energy; therefore, we could not obtain a quantitative value for the quasiparticle lifetime. Nevertheless, we calculated the anti-Hermitian part of the self-energy (the broadening function, Γ) which allowed us to examine the qualitative behaviour of the quasiparticle lifetime near the Fermi level. As expected, the lifetime is infinite at the Fermi level and becomes finite as we move away from it. Our numerical results for the broadening function were found to fit reasonably well to an ω^2 dependence

on ω for small ω, in qualitative agreement with Fermi liquid theory and in agreement with experiment (see Fig. 1.10).

In examining the effects of magnetic fields on TI SSs, we first returned to the non-interacting regime. Here, we found that the number of surface LLs does not increase with the slab thickness. We calculated the surface DOS and plotted the LL spectrum. The zeroth LL was easily distinguishable and the LL peaks were linear in $\sqrt{|n|B}$, except for the thinnest slab considered. These two signatures are clear indicators that for sufficiently thick slabs the TI SSs obey a relativistic LL spectrum. Once again, this is expected for states on a massless DC. Next, we added interactions to our finite magnetic field model by deriving second order self-energy expressions. We did not numerically evaluate these expressions; however, given more time it would be possible to do so and would be an interesting property to examine.

5.1   Future work

There is still much work to be done in terms of obtaining a better understanding of how electron-electron interactions and magnetic fields affect SSs on 3D TIs. Some new results can be obtained within the framework of this model through only minor effort, while other results will require substantial work and time.

Since the Hamiltonian used in this thesis describes Bi_2Te_3 and Sb_2Te_3 in addition to Bi_2Se_3, it is quite easy in principle to conduct the same numerical work done in this thesis for these materials as well as no change to any analytical work is needed. Another easy property to examine is the temperature dependence of the broadening function. As the Matsubara formalism was used in calculating the GFs, all that is needed to incorporate temperature effects is to keep the full Fermi and Bose functions in Eq. 3.73 and 3.74. Calculating the broadening function for a larger/different range of energies, for momenta away from the Fermi surface, and for different chemical potentials are also easy to do as the numerical algorithms for such calculations need not be changed.

One of the more involved calculations would be to numerically evaluate the integrals for the calculation of the Hermitian part of the self-energy, Λ. This would be very computationally intensive. As such, it would only be realistic to calculate Λ for smaller slab thicknesses

(i.e., $N_z \approx 40$). This, along with the calculation of Γ, would allow for the calculation of the interacting spectral function, and thus a quantitative result for the quasiparticle lifetime by measuring the width of the peaks in the spectral function. To determine the quasiparticle lifetime in a finite magnetic field, a numerical calculation of the self-energy expressions at the end of Sec. 4.3.1 is needed. The interacting spectral function and then the DOS could then be calculated and the quasiparticle lifetime extracted from the broadened LL peaks. Real TI samples are not perfect crystals; they have defects/disorder that break translational symmetry. Our model could be made more realistic by considering such disorder effects. This is possible by periodizing our Hamiltonian in all three directions, Fourier transforming fully to real space, and adding a randomly varying potential on each lattice site to mimic disorder. Moreover, we neglected the cubic terms in $\boldsymbol{k}$ in our Hamiltonian. Including such terms would allow for an examination of how interactions renormalize DC warping, and conversely, how DC warping affects the quasiparticle lifetime in the presence of interactions. Finally, we used a contact interaction here. More accurate results may be obtained by using a screened Coulomb interaction.

Bibliography

[1] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).

[2] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[3] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).

[4] M. Z. Hasan and J. E. Moore, Annu. Rev. Condens. Matter Phys. **2**, 55 (2011).

[5] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).

[6] K. v. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).

[7] S.-Q. Shen, *Topological Insulators: Dirac Equation in Condensed Matters: 174 (Springer Series in Solid-State Sciences)* (Springer, 2013).

[8] C.-Z. Chang, J. Zhang, X. Feng, J. Shen, Z. Zhang, M. Guo, K. Li, Y. Ou, P. Wei, L.-L. Wang, Z.-Q. Ji, Y. Feng, S. Ji, X. Chen, J. Jia, X. Dai, Z. Fang, S.-C. Zhang, K. He, Y. Wang, L. Lu, X.-C. Ma, and Q.-K. Xue, Science **340**, 167 (2013).

[9] A. Bestwick, E. Fox, X. Kou, L. Pan, K. L. Wang, and D. Goldhaber-Gordon, Phys. Rev. Lett. **114**, 187201 (2015).

[10] C.-Z. Chang, W. Zhao, D. Y. Kim, H. Zhang, B. A. Assaf, D. Heiman, S.-C. Zhang, C. Liu, M. H. W. Chan, and J. S. Moodera, Nat. Mater. **14**, 473 (2015).

[11] G. Jotzu, M. Messer, R. Desbuquois, M. Lebrat, T. Uehlinger, D. Greif, and T. Esslinger, Nature **515**, 237 (2014).

[12] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).

[13] B. A. Bernevig and S.-C. Zhang, Phys. Rev. Lett. **96**, 106802 (2006).

[14] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).

[15] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).

[16] R. Roy, Phys. Rev. B **79**, 195321 (2009).

[17] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).

[18] J. Maciejko, T. L. Hughes, and S.-C. Zhang, Annu. Rev. Condens. Matter Phys. **2**, 31 (2011).

[19] J. J. Sakurai and J. J. Napolitano, *Modern Quantum Mechanics* (Pearson, 2014).

[20] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).

[21] R. D. Tscheuschner, S. Hoch, E. Leschinsky, C. Meier, S. Theis, and A. D. Wieck, Int. J. Mod. Phys. B **12**, 1147 (1998).

[22] I. Knez and R.-R. Du, Front. Phys. **7**, 200 (2012).

[23] A. Roth, C. Brüne, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Science **325**, 294 (2009).

[24] C. Brüne, A. Roth, H. Buhmann, E. M. Hankiewicz, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Nat. Phys. **8**, 485 (2012).

[25] K. C. Nowack, E. M. Spanton, M. Baenninger, M. König, J. R. Kirtley, B. Kalisky, C. Ames, P. Leubner, C. Brüne, H. Buhmann, L. W. Molenkamp, D. Goldhaber-Gordon, and K. A. Moler, Nat. Mater. **12**, 787 (2013).

[26] E. M. Spanton, K. C. Nowack, L. Du, G. Sullivan, R.-R. Du, and K. A. Moler, Phys. Rev. Lett. **113**, 026804 (2014).

[27] L. Du, I. Knez, G. Sullivan, and R.-R. Du, Phys. Rev. Lett. **114**, 096802 (2015).

[28] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).

[29] J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306 (2007).

[30] R. Roy, Phys. Rev. B **79**, 195322 (2009).

[31] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).

[32] S. K. Kushwaha, I. Pletikosi?, T. Liang, A. Gyenis, S. H. Lapidus, Y. Tian, H. Zhao, K. S. Burch, J. Lin, W. Wang, H. Ji, A. V. Fedorov, A. Yazdani, N. P. Ong, T. Valla, and R. J. Cava, Nat. Commun. **7**, 11456 (2016).

[33] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature **452**, 970 (2008).

[34] D. Hsieh, Y. Xia, L. Wray, D. Qian, A. Pal, J. H. Dil, J. Osterwalder, F. Meier, G. Bihlmayer, C. L. Kane, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Science **323**, 919 (2009).

[35] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nat. Phys. **5**, 398 (2009).

[36] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Nat. Phys. **5**, 438 (2009).

[37] Y. L. Chen, J. G. Analytis, J.-H. Chu, Z. K. Liu, S.-K. Mo, X. L. Qi, H. J. Zhang, D. H. Lu, X. Dai, Z. Fang, S. C. Zhang, I. R. Fisher, Z. Hussain, and Z.-X. Shen, Science **325**, 178 (2009).

[38] D. Hsieh, Y. Xia, D. Qian, L. Wray, F. Meier, J. H. Dil, J. Osterwalder, L. Patthey, A. V. Fedorov, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Phys. Rev. Lett. **103**, 146401 (2009).

[39] M. Neupane, A. Richardella, J. Sánchez-Barriga, S. Xu, N. Alidoust, I. Belopolski, C. Liu, G. Bian, D. Zhang, D. Marchenko, A. Varykhalov, O. Rader, M. Leandersson, T. Balasubramanian, T.-R. Chang, H.-T. Jeng, S. Basak, H. Lin, A. Bansil, N. Samarth, and M. Z. Hasan, Nat. Commun. **5**, ncomms4841 (2014).

[40] D. Hsieh, Y. Xia, D. Qian, L. Wray, J. H. Dil, F. Meier, J. Osterwalder, L. Patthey, J. G. Checkelsky, N. P. Ong, A. V. Fedorov, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature **460**, 1101 (2009).

[41] P. Roushan, J. Seo, C. V. Parker, Y. S. Hor, D. Hsieh, D. Qian, A. Richardella, M. Z. Hasan, R. J. Cava, and A. Yazdani, Nature **460**, 1106 (2009).

[42] T. Zhang, P. Cheng, X. Chen, J.-F. Jia, X. Ma, K. He, L. Wang, H. Zhang, X. Dai, Z. Fang, X. Xie, and Q.-K. Xue, Phys. Rev. Lett. **103**, 266803 (2009).

[43] Z. Alpichshev, J. G. Analytis, J.-H. Chu, I. R. Fisher, Y. L. Chen, Z. X. Shen, A. Fang, and A. Kapitulnik, Phys. Rev. Lett. **104**, 016401 (2010).

[44] J. Xiong, A. C. Petersen, D. Qu, Y. S. Hor, R. J. Cava, and N. P. Ong, Physica E **44**, 917 (2012).

[45] J. Xiong, Y. Khoo, S. Jia, R. J. Cava, and N. P. Ong, Phys. Rev. B **88**, 035128 (2013).

[46] K. Yang, W. Setyawan, S. Wang, M. Buongiorno Nardelli, and S. Curtarolo, Nat. Mater. **11**, 614 (2012).

[47] Y. Zhang, K. He, C.-Z. Chang, C.-L. Song, L.-L. Wang, X. Chen, J.-F. Jia, Z. Fang, X. Dai, W.-Y. Shan, S.-Q. Shen, Q. Niu, X.-L. Qi, S.-C. Zhang, X.-C. Ma, and Q.-K. Xue, Nat. Phys. **6**, 584 (2010).

[48] C.-L. Song, Y.-L. Wang, Y.-P. Jiang, Y. Zhang, C.-Z. Chang, L. Wang, K. He, X. Chen, J.-F. Jia, Y. Wang, Z. Fang, X. Dai, X.-C. Xie, X.-L. Qi, S.-C. Zhang, Q.-K. Xue, and X. Ma, Appl. Phys. Lett. **97**, 143118 (2010).

[49] G. Wang, X. Zhu, J. Wen, X. Chen, K. He, L. Wang, X. Ma, Y. Liu, X. Dai, Z. Fang, J. Jia, and Q. Xue, Nano Res. **3**, 874 (2010).

[50] Y. Yan, Z.-M. Liao, Y.-B. Zhou, H.-C. Wu, Y.-Q. Bie, J.-J. Chen, J. Meng, X.-S. Wu, and D.-P. Yu, Sci. Rep. **3**, srep01264 (2013).

[51] J. Wang, A. M. DaSilva, C.-Z. Chang, K. He, J. K. Jain, N. Samarth, X.-C. Ma, Q.-K. Xue, and M. H. W. Chan, Phys. Rev. B **83**, 245438 (2011).

[52] M. Liu, C.-Z. Chang, Z. Zhang, Y. Zhang, W. Ruan, K. He, L.-l. Wang, X. Chen, J.-F. Jia, S.-C. Zhang, Q.-K. Xue, X. Ma, and Y. Wang, Phys. Rev. B **83**, 165440 (2011).

[53] Y. Jing, S. Huang, K. Zhang, J. Wu, Y. Guo, Z. Peng, Hailin and, and H. Q. Xu, Nanoscale **8**, 1879 (2016).

[54] S. R. Park, W. S. Jung, G. R. Han, Y. K. Kim, C. Kim, D. J. Song, Y. Y. Koh, S. Kimura, K. D. Lee, N. Hur, J. Y. Kim, B. K. Cho, J. H. Kim, Y. S. Kwon, J. H. Han, and C. Kim, New J. Phys. **13**, 013008 (2011).

[55] C.-L. Song, L. Wang, K. He, S.-H. Ji, X. Chen, X.-C. Ma, and Q.-K. Xue, Phys. Rev. Lett. **114**, 176602 (2015).

[56] T. Hanaguri, K. Igarashi, M. Kawamura, H. Takagi, and T. Sasagawa, Phys. Rev. B **82**, 081305 (2010).

[57] P. Cheng, C. Song, T. Zhang, Y. Zhang, Y. Wang, J.-F. Jia, J. Wang, Y. Wang, B.-F. Zhu, X. Chen, X. Ma, K. He, L. Wang, X. Dai, Z. Fang, X. Xie, X.-L. Qi, C.-X. Liu, S.-C. Zhang, and Q.-K. Xue, Phys. Rev. Lett. **105**, 076801 (2010).

[58] Y. Jiang, Y. Wang, M. Chen, Z. Li, C. Song, K. He, L. Wang, X. Chen, X. Ma, and Q.-K. Xue, Phys. Rev. Lett. **108**, 016401 (2012).

[59] C. Pauly, C. Saunus, M. Liebmann, and M. Morgenstern, Phys. Rev. B **92**, 085140 (2015).

[60] S.-Q. Shen, arXiv:0909.4125 [cond-mat] (2009), arXiv: 0909.4125.

[61] Z. Yang and J. H. Han, Phys. Rev. B **83**, 045415 (2011).

[62] P. Schwab and M. Dzierzawa, Phys. Rev. B **85**, 155403 (2012).

[63] H. Bruus and K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics: An Introduction (Oxford Graduate Texts)* (Oxford University Press, 2004).

[64] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems (Dover Books on Physics)* (Dover Publications, 2012).

[65] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Brooks Cole, 1976).

[66] J. Hubbard, Proc. R. Soc. A **276**, 238 (1963).

[67] Nat. Phys. **9**, 523 (2013).

[68] H. Tasaki, J. Phys. Condens. Matter **10**, 4353 (1998).

[69] R. D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem: Second Edition (Dover Books on Physics)* (Dover Publications, 2012).

[70] J. Schrieffer, J. R. Nat. Stand. Sec. A **74A**, 537 (1969).

[71] S. Y. Wu, J. Cocks, and C. S. Jayanthi, Phys. Rev. B **49**, 7957 (1994).

[72] C.-X. Liu, X.-L. Qi, H. Zhang, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. B **82**, 045122 (2010).

[73] L. Fu, Phys. Rev. Lett. **103**, 266801 (2009).

[74] T. Zhang, N. Levy, J. Ha, Y. Kuk, and J. A. Stroscio, Phys. Rev. B **87**, 115410 (2013).

[75] C. Hodges, H. Smith, and J. W. Wilkins, Phys. Rev. B **4**, 302 (1971).

[76] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Phys. Rev. B **78**, 195424 (2008).

[77] M. König, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, J. Phys. Soc. Jpn **77**, 031007 (2008).

[78] M. Na and F. Marsiglio, arXiv:1604.01661 [cond-mat, physics:quant-ph] (2016), arXiv: 1604.01661.