Gapped Dirac cones and spin texture in thin film topological insulator

Peter Thalmeier1 and Alireza Akbari2,3,*

1Max Planck Institute for Chemical Physics of Solids, D-01187 Dresden, Germany
2Max Planck POSTECH Center for Complex Phase Materials, POSTECH, Pohang 790-784, Korea
3Department of Physics, POSTECH, Pohang, Gyeongbuk 790-784, Korea

(Received 27 January 2020; accepted 11 June 2020; published 1 July 2020)

The protected surface states of topological insulators (TIs) form gapless Dirac cones corresponding nondegenerate eigenstates with helical spin polarization. The presence of a warping term deforms the isotropic cone of the most simple model into snowflake Fermi surfaces as in Bi2Se3 and Bi2Te3. Their features have been identified in STM quasiparticle interference (QPI) experiments on isolated surfaces. Here we investigate the QPI spectrum for the TI thin film geometry with a finite tunneling between the surface states. This leads to a dramatic change of spectrum due to gapping and a change in spin texture that should leave distinct signatures in the QPI pattern. We consider both normal and magnetic exchange scattering from the surface impurities and obtain the scattering $t$ matrix in Born approximation as well as the general closed solution. We show the expected systematic variation of QPI snowflake intensity features by varying film thickness and study, in particular, the influence on backscattering processes. We predict the variation of QPI spectrum for Bi2Se3 thin films using the observed gap dependence from ARPES results.

DOI: 10.1103/PhysRevResearch.2.033002

1. INTRODUCTION

The most dramatic manifestation in a topological insulator (TI) is the presence of protected surface states which are helical spin locked nondegenerate eigenstates with a gapless Dirac dispersion. This has been manifestly demonstrated, e.g., by countless (spin-polarized) ARPES experiments [1–3] and STM-QPI [4–10] investigations on isolated surfaces, in particular, in the canonical examples Bi2Se3 and Bi2Te3. The theory of QPI spectra on single surfaces has been presented in numerous effective model investigations [10–14] and recently also with an ab initio approach [15]. Due to their helical nature these surface states have forbidden backscattering by normal (scalar) impurities leading to the well-known weak antilocalization (WAL) effects in the surface magnetoconductance experiments [16,17]. However the strict WAL due to the destructive interference caused by a Berry phase $\mp \pi$ applies only to isolated surfaces. When we consider thin films there will be a tunneling between opposite surface states due to their finite decay lengths into the bulk. The tunneling mixes the surfaces states leading to a finite gap at the Dirac point and it lifts their protection as witnessed by the suppression of helical polarization and modification of the Berry phase for states close to the gap energy [18]. This has been directly observed in magnetotransport experiments in Bi2Se3 where the weak antilocalization as indicator of the ungapped states rather suddenly breaks down below a thickness of five quintuple layers (QL) [16] (In Bi2Se3, 1 QL ≈ $c/3 = 9.55$ Å, $c$ is the lattice constant.) The gapping of the Dirac cones in thin films has also been observed directly with the ARPES technique [19,20].

Here we want to investigate theoretically what kind of influence the intersurface tunneling in TI thin films may have on the QPI pattern in STM experiments. In the known QPI experiments on single surfaces, the Fourier transformed scanning image is dominated by the scattering vectors that interconnect the characteristic points of the real “snowflake” type Fermi surfaces, with those of the backscattering characteristically missing. Therefore our investigation demands that we take into account not only the intersurface tunneling but also include the higher-order intrasurface warping terms that deform the isotropic Dirac cones with Fermi surface (FS) circles into the observed snowflake geometry. The QPI pattern of the isolated surfaces should be dramatically modified for small bias voltages such that $\omega = eV \approx \pm |t|$, where $V$ is the STM bias voltage and $t$ (which may be assumed positive) is the intersurface tunneling leading to a gap energy $2|t|$. Furthermore the QPI spectrum due to scattering from magnetic impurities should also be sensitive to the change of spin texture close to the gap threshold which is induced by the tunneling consisting primarily in a suppression of helical polarization for small energies. Most importantly, the intensity variation of QPI peaks on the constant frequency cuts should carry an imprint of the reemergence of backscattering for the topologically trivial surface states close to the hybridization gap edge. In this work, we intend to give a detailed theoretical analysis of the tunneling effects in TI thin films that one may expect to see as a guidance to QPI experiments with systematic variation of film thickness. We also present the
TABLE I. Dirac cone parameters for Bi$_2$Se$_3$ and Bi$_2$Te$_3$ (Refs. [2,27,28]). Here, $v$ and $\lambda$ are determined from experiment, e.g., by ARPES. They also fix the intrinsic momentum and energy scales $k_c$ and $E^*$.

| TI compound      | quintuple layer | Fermi velocity | intrinsic momentum scale | intrinsic energy scale | warping parameter |
|------------------|-----------------|----------------|--------------------------|------------------------|------------------|
|                  | 1QL             | $v$            | $k_c = \sqrt{v/k}$       | $E^* = v k_c$          | $\lambda = v^2/E^* = E^*/k_c^2$ |
| Bi$_2$Se$_3$ [2] | 9.55 Å          | 3.55 eV Å      | 0.17 Å$^{-1}$ ($0.60$QL)$^{-1}$ | 0.59 eV               | 1.28 $\times 10^2$ eV Å$^3$ |
| Bi$_2$Te$_3$ [27]| 10.16 Å         | 2.55 eV Å      | 0.10 Å$^{-1}$ ($1$QL)$^{-1}$  | 0.25 eV               | 2.5 $\times 10^2$ eV Å$^3$  |

The expected variation of the QPI spectrum using the thickness dependent surface state gap obtained from ARPES.

These issues are central to understand properties of thin film topological surface states. Our theoretical approach to them is entirely analytical $t$-matrix theory [21,22] because we want to study in detail the combined effect of in-plane warping and intersurface hybridization on QPI images. This allows a transparent interpretation of their spin texture, Berry phase, and gradual change of backscattering characteristics with intersurface tunneling which are at the heart of the TI thin film problem. Naturally this demands a restriction to indispensable ingredients of the TI thin film model, e.g., to only four basis states and a momentum independent intralayer impurity scattering. Otherwise a fully numerical approach would be necessary which is not the aim of this work. We will derive the QPI spectrum in Born approximation as well as in closed fully analytical $t$-matrix form, using both normal and magnetic example of (momentum-isotropic) impurity scattering.

The paper is organized as follows, in Secs. II and III, we define the model of warped and intersurface coupled thin film TI states and discuss their spin texture in Sec. IV. In the main formal part of Sec. V, we derive the full $t$ matrix for the assumed impurity scattering potential and discuss the simplified Born approximation. These results are then applied to calculate the QPI spectrum in Sec. VI in the normal and magnetic scattering cases with full $t$-matrix approach. An extended discussion of the numerical results as a function of bias voltage and film thickness is presented in Sec. VII. Finally Sec. VIII gives the conclusions.

II. MODEL FOR WARPED TOPOLOGICAL INSULATOR DIRAC CONES

The protected surface states in topological insulators like, e.g., Bi$_2$Se$_3$ and Bi$_2$Te$_3$ may be derived from a $k \cdot p$-type Hamiltonian model for the 3D bulk states [23,24] using four strongly spin-orbit coupled basis states with angular momentum (pseudospin) components $m_j = \pm \frac{1}{2}$ and parity $P = \pm$. Because the crystal-field splitting is much larger than the spin-orbit coupling these states are dominated by $p_z$ orbitals from Se (4p) and Bi (6p) and therefore the $m_j$ pseudospin is proportional to the real electron spin $\sigma_z$ [25,26] which will be used for simplicity in the following. The surface states are obtained by solving the 1D Dirac equation along the surface normal direction ($\hat{z}$) under appropriate boundary conditions [23,24]. The effective low energy Hamiltonian for the 2D topological surface states can then be parametrized in the well-known form [11,27–29]

$$H = \sum_k \hbar \epsilon_k \left( \begin{array}{cc} \cos k \theta & -ie^{i\phi} \sin k \theta \\ ie^{i\phi} \sin k \theta & -\hbar^2 \sin^2 k \theta \end{array} \right) \cdot \left( \begin{array}{c} \hat{z} \cos \theta \hbar \sigma_z \end{array} \right).$$

Here the first term describes the massless Dirac states forming an isotropic cone dispersion in the 2D surface Brillouin zone (SBZ) where the velocity $v$ is the slope of the cone and its position is the SBZ projection of a bulk time-reversal invariant (TRI) point of the bulk Brillouin zone, e.g., the $\Gamma$ point for the above compounds. The second “warping term” distorts the cone anisotropically in accordance with crystal symmetry in such a way that the 2D Fermi surfaces (cuts through the cone) evolve from circular shapes to hexagons to six-prong “snowflakes” [27,30] when the Fermi energy increases. The amount of distortion is determined by the strength $\lambda$ of the warping term. The spin degeneracy of states at the projected TRI points is lifted due to spin-orbit coupling when moving away from them. Furthermore the surface wave vector $k = (k_x, k_y) = \tilde{k} = k (\cos \theta_k, \sin \theta_k)$ has the polar presentation with a magnitude $k = (k_x^2 + k_y^2)^{1/2}$ and an azimuthal angle $\theta_k = \tan^{-1}(k_y/k_x)$.

The above $2 \times 2$ Hamiltonian can be explicitly written as

$$\hbar \epsilon_k = \epsilon_k \left( \begin{array}{cc} k_x \cos \theta_k & -ie^{i\phi} \sin \theta_k \\ ie^{i\phi} \sin \theta_k & -\hbar^2 \sin^2 k \theta \end{array} \right),$$

where $\epsilon_k = E^* \tilde{k} = vk$ is the isotropic ($\lambda = 0$) Dirac dispersion. An overall parabolic term $\epsilon_{0k} = k^2/2m_0$ that breaks particle hole symmetry for larger energies [28] will be neglected. We also introduced a momentum scale $k_c = \sqrt{v/\pi}$ and energy scale $E^* = vk_c$ to define dimensionless momentum variables $\tilde{k} = k/k_c$ and energies $\tilde{\epsilon}_k = \epsilon_k/E^* = \tilde{k}$, see also Table I. The single surface Hamiltonian may be diagonalized by a unitary transformation according to

$$\tilde{h}_k = S_k^\dagger \hat{h}_k S_k = E_k \hat{\epsilon}_k,$$

with the warped Dirac cone energy now given by

$$E_k = [(vk)^2 + (\lambda k^3 \cos \theta)^2]^{1/2} = E^* \tilde{k} [1 + (\tilde{k}^2 \cos \theta k)^2]^{1/2},$$

or dimensionless $\tilde{E}_k = E_k/E^*$. The dispersion of Eq. (4) shown in Figs. 1(i) and 1(j) results in a warped constant energy that surfaces have the shape of six-prong “snowflakes” (Fig. 1). The unitary transformation to the eigenstates is given by the matrix

$$S_k = \left( \begin{array}{cc} \cos \frac{\phi}{2} & i \sin \frac{\phi}{2} e^{-i\theta_k} \\ i \sin \frac{\phi}{2} e^{i\theta_k} & \cos \frac{\phi}{2} \end{array} \right).$$
The columns of $S_k$ are the eigenvectors $|\psi_k\rangle$ (in spin $\uparrow\downarrow$ basis) to the eigenvalues $\pm i\hbar E_k$ ($k = \pm$). These are called “chiral” or “helical” basis states because in the isotropic case $\lambda = 0$, they are also eigenstates to the chirality operator defined by $\kappa_z = (\sigma \times \mathbf{k}) \cdot \mathbf{z} = \pm 1$. The spin mixing angle $\phi_k$ is given by

$$\tan \phi_k = k^2/(k^2 \cos 3\theta_k) = 1/(k^2 \cos 3\theta_k). \quad (6)$$

In the full circle $0 \leq \theta_k < 2\pi$, $\cos 3\theta_k$ changes sign six times at $\theta_k = (2n + 1)\pi/3$; $n = 0, 1, 2, 3$. To obtain a continuous variation of $\phi_k(\theta_k)$ at these boundaries and to have a well defined uniform convergence to the isotropic limit $\phi_k \to \pi/2$ for vanishing warping ($\lambda \to 0$), we define $\phi_k = \tan^{-1}(1/k^2 \cos 3\theta_k)$ for $\cos 3\theta_k > 0$ and $\phi_k = \tan^{-1}(1/k^2 \cos 3\theta_k) + \pi$ for $\cos 3\theta_k < 0$. This amounts to taking the second (upper) branch of $\tan^{-1}$ in the latter case. The $k$ and $\theta_k$ dependence of the spin mixing angle $\phi_k$ is shown in Fig. 2, and it indeed varies continuously with $\theta_k$ around the (half)circle centered at the isotropic limit $\phi_k = \pi/2$.

For calculation of the QPI spectrum, one needs the Green’s function in spin representation. For the single surface problem it is given by

$$G_k = (i\omega_n - \hbar k)^{-1} = \begin{pmatrix} i\omega_n - \epsilon_k k^2 \cos 3\theta_k & \epsilon_k e^{i\theta_k} \\ i\epsilon_k e^{i\theta_k} & i\omega_n + \epsilon_k k^2 \cos 3\theta_k \end{pmatrix}^{-1}. \quad (7)$$

Defining the warping energy by

$$\Delta_k = \epsilon_k(k^2 \cos 3\theta_k) = \pm (E_k^2 - \epsilon_k^2)^{1/2},$$

the Green’s function is obtained as

$$G_k = \frac{1}{(i\omega_n)^2 - \hbar^2 k^2} \begin{pmatrix} i\omega_n + \Delta_k & -i\epsilon_k e^{i\theta_k} \\ i\epsilon_k e^{i\theta_k} & i\omega_n - \Delta_k \end{pmatrix}. \quad (8)$$

In the isotropic Dirac cone case, one simply has to set $\Delta_k = 0$ and $E_k = \epsilon_k$.

### III. Splitting of Dirac Cones by Intersurface Tunneling in TI Thin Films

The warped cones of single TI surfaces have been abundantly demonstrated by ARPES experiments [31,32]. As outlined in the introduction, thin films of the tunneling of helical Dirac states between the two surfaces introduce a gap in the spectrum signifying the transition to topologically trivial surface states close to the gap threshold as a function of thickness. This has been observed again directly with ARPES [19,20] in agreement with magnetotransport results [16]. Theoretically it has been studied in detail [23,24,33] by starting from the bulk thin film states and introducing the appropriate boundary conditions. A nonmonotonic, even oscillating behavior of
surface state tunneling and gap size as a function of thickness \( d \) is possible although this has so far not been observed in photoemission (Sec. VII).

For the purpose of investigating QPI signatures of the topological transition in thin films, we employ a phenomenological model starting from the independent helical states on the two isolated surfaces. They are coupled by the thickness dependent tunneling matrix element \( t(\alpha) \) of top (T) and bottom (B) surface states with equal helicities (Fig. 3). The \( 4 \times 4 \) thin-film Hamiltonian in the space with basis \( |\sigma\alpha\rangle \) (helicities \( \kappa = \pm 1 \) and surface index \( \alpha = \text{T}, \text{B} \)) may be written as \( 2 \times 2 \) block matrix in T,B space according to

\[
H_k = \begin{pmatrix} \hbar k & i\sigma_0 \\ i\sigma_0 & -\hbar k \end{pmatrix} \rightarrow \tilde{H}_k = \begin{pmatrix} E_k\kappa_\tau & i\kappa_0 \\ i\kappa_0 & -E_k\kappa_\tau \end{pmatrix},
\]

in the spin space with basis \( |\sigma\rangle \) (spin up \( \uparrow \), down \( \downarrow \)) (left) and helicity space with basis \( |\kappa\rangle \) (helicity \( \kappa = \pm 1 \)). In both representations \( \alpha = \text{T, B} \) is the surface index. The minus sign in the last element appears because states of equal helicities on B,T surfaces belong to opposite half-cones (see Fig. 3). We use the convention that \( \sigma, \kappa, \alpha \) denote the vector of 2 \( \times 2 \) Pauli matrices in spin, helicity and surface layer (T,B) space, respectively. Furthermore, the unit in each space is denoted by \( \sigma_0, \kappa_0, \alpha_0 \). The film Hamiltonian may be diagonalized by two transformations for the \( \kappa = \pm 1 \) helicities separately with the simple form

\[
U_k^\pm = \begin{pmatrix} \cos \psi_k & \mp \sin \psi_k \\ \pm \sin \psi_k & \cos \psi_k \end{pmatrix}; \quad \tan 2\psi_k = \frac{t}{E_k} = \frac{i}{\tilde{E}_k},
\]

where \( \tilde{t} = t/E^* \). The total transformation matrix from spin single surface representation to helicity-film eigenstate representation that diagonalizes \( H_k \) is then given by \( W_k \). The column vectors of this \( 4 \times 4 \) matrix are the helicity eigenstates of the film (explicitly given in Appendix A). They have the energies resulting from the diagonalization \( W_k^\dagger H_k W_k = E_k\kappa_\tau \otimes \tau_\tau \) with

\[
\tilde{E}_k = [E_k^2 + t^2]^\frac{1}{2} = E^*[k^2(1 + (\hat{k}^2 \cos 3\theta_k)^2) + \hat{t}^2]^\frac{1}{2}.
\]

Therefore, at the Dirac point \( k = 0 \), a gap of size \( \Delta = 2|\hat{t}| \) opens between T,B hybridized film states \( \tau = 1, 2 \) each twofold degenerate due to time-reversal symmetry with helicities \( \kappa = \pm 1 \).

IV. MIXED TOPOLOGICAL SURFACE STATES AND THEIR SPIN TEXTURE AND BERRY PHASE

The explicit form of the surface states is given by the columns of Eq. (A1). For the upper and lower split half-cones with energies \( E_{1,2}(k) = \pm \tilde{E}_k \), the pairs of doubly degenerate \((\kappa = \pm \uparrow)\) upper and lower \((\tau = 1, 2)\) band states \( |\tilde{\psi}_{\tau\kappa}\rangle \) can be written, respectively, as

\[
|\tilde{\psi}_{1\kappa}\rangle = \cos \psi_k |\psi_{\tau\uparrow}\rangle + \sin \psi_k |\psi_{\tau\downarrow}\rangle : E_{1\kappa} = +\tilde{E}_k,
\]

\[
|\tilde{\psi}_{2\kappa}\rangle = \cos \psi_k |\psi_{\tau\downarrow}\rangle - \sin \psi_k |\psi_{\tau\uparrow}\rangle : E_{2\kappa} = -\tilde{E}_k.
\]

in terms of the isolated surface chiral states \( |\psi_{\alpha\tau}\rangle \). These eigenstates are related by combined time reversal \( \mp \leftrightarrow \mp \) represented by \( \Theta = i\sigma_3 \hat{K} \) (K complex conjugation) and interchange of surfaces \( T,B \leftrightarrow B,T \) represented by the reflection \( \Sigma = \sigma_1 \). The pairwise degeneracy of hybridized states in Eq. (12) is due to the symmetry \((\Theta \Sigma)|\tilde{\psi}_{\tau\kappa}\rangle = |\tilde{\psi}_{\tau\kappa}\rangle \). For states close to the gap with \( \tilde{E}_k \gg |t| (k \ll \hat{k}_s) \), one has \( \psi_k \approx \frac{\pi}{2} \) meaning these are true film states with equal amplitudes on B,T surfaces. On the other hand, far from the gap with \( \tilde{E}_k \ll |t| (k > k_s) \), we have \( \psi_k \approx \frac{\pi}{4} - \frac{\pi}{2} < \frac{\pi}{4} \) and consequently the above states are mostly localized on either T or B surface with only small amplitude on the opposite B or T surfaces. Therefore, for large thickness when \( t \rightarrow 0 \), the state in Eq. (12) decouple to isolated surface states on T,B such that \( |\tilde{\psi}_{1\kappa}\rangle \) becomes states with opposite chirality on opposite surfaces with positive energies and likewise \( |\tilde{\psi}_{2\kappa}\rangle \) with negative energies. From these states we may now calculate the surface spin textures \( |\sigma\rangle^T_{\tau\kappa} \) on T or B of the model as a function of the tunneling matrix element \( t \) between the isolated helicity surface states. The spin textures of states 1,2 will be opposite on the same surface and because of \( |\sigma\rangle^T_{1\kappa} = -|\sigma\rangle^T_{2\kappa} \) identical on opposite surfaces. Therefore, restricting to the top surface we have for the spin expectation value of each pair corresponding to each half-cone:

\[
|\sigma\rangle^T_{1,2\kappa} = \cos^2 \psi_k (|\psi_{\tau\uparrow}\rangle |\sigma\rangle |\psi_{\tau\downarrow}\rangle) + \sin^2 \psi_k (|\psi_{\tau\downarrow}\rangle |\sigma\rangle |\psi_{\tau\uparrow}\rangle) = \pm (\sin \phi_k \sin \theta_k, \sin \phi_k \cos \theta_k, \cos \phi_k) \cos 2\psi_k.
\]

Using the previous expressions for the mixing angles \( \phi_k, \psi_k \) in terms of the polar angle \( \theta_k \) we get explicitly for in-
out-of-plane spin polarization:

\[
\langle \sigma_\parallel \rangle_T = \frac{k}{\sqrt{k^2 + (k^2 \cos 3\theta_k)^2 + t^2}}; \\
\langle \sigma_\perp \rangle_T = \frac{k^3 \cos \theta_k}{\sqrt{k^2 + (k^2 \cos 3\theta_k)^2 + t^2}}.
\]  (14)

Here, \(\langle \sigma_\parallel \rangle_T = \langle \sigma_\parallel \rangle_T^2 + \langle \sigma_\perp \rangle_T^2\) and therefore the total length of the spin on \(T\) is

\[
\langle \sigma_{\text{total}} \rangle_T^2 = \left( \langle \sigma_\parallel \rangle_T^2 + \langle \sigma_\perp \rangle_T^2 \right) \cos^2 2\psi_k = \left( \frac{E_k}{E_{\text{m}}} \right)^2.
\]  (15)

When the tunneling between surfaces becomes negligible \(\psi_k \to 0\) and \(\langle \sigma_{\text{total}} \rangle_T^2 \to 1\). Contour plots of \(\langle \sigma_\parallel \rangle_T\) and \(\langle \sigma_\perp \rangle_T\) are shown in Fig. 4 for isolated and coupled surfaces. In particular, the in-plane component which is maximum (\(\leq 1\)) at the \(\Gamma\) point and along the tips (white region) in Fig. 4(c) is suppressed to zero by the tunneling at the \(\Gamma\) point in (d) (dark green area). We note that \textit{ab initio} calculations of the spin textures for thick slabs have also been performed [34].

The Dirac point of topological surface states may be viewed as a monopole in momentum space. It is connected with a nonvanishing Berry-phase when encircling the Dirac point on a closed path \(C\) containing the origin. First we consider the case of warped Dirac cone of a single (T or B) surface. Then the Berry phase is given by \((\kappa = \pm)\) [35,36]

\[
\gamma_k = \oint_C dk \cdot i\langle \psi_k^* | \nabla_k | \psi_k \rangle,  
\]  (16)

where the Berry connection described by

\[
i\langle \psi_k^* | \nabla_k | \psi_k \rangle = \kappa [\cos \phi_k - 1],
\]  depends on the warping through \(\phi_k\). Because of the antiperiodic property \(\cos [\phi_k(0)] = - \cos [\phi_k(\pi)]\), the contour integral over the cosine vanishes and we get \(\gamma_k = \mp \pi\). This means that the topologically nontrivial Berry phase for the Dirac cone states is not influenced by the warping effect since the latter does not destroy the \(k = 0\) singularity. Therefore, when considering the effect of the intersurface tunneling \(t\) on the Berry phase, we may safely neglect the warping for simplicity. The tunneling leads to film states determined by the TB intersurface mixing angle \(\psi_k\) [Eq. (12)]. The Berry connection for the isotropic \((\phi_k = \pm \pi)\) film states \((t = 1, 2, \kappa = \pm)\), \(\langle \psi_k^* \rangle\) with a given bare Dirac cone energy \(E = E_k = vk\) may be calculated as

\[
\gamma_{\text{D}} = \oint_C dk \cdot i\langle \psi_k^* | \nabla_k | \psi_k \rangle = (-1)^{1+t} \kappa \pi \cos 2\psi; \\
\cos 2\psi = \frac{E}{(E^2 + t^2)^{1/2}} = \left( 1 - \frac{1}{2}(\frac{E}{t})^2 \right)^{1/2} E \ll t (k \ll t/v) \\
E \gg t (k \gg t/v).
\]  (17)

Therefore the Berry phase of states at the gap threshold which have parabolic dispersion vanishes while for energies much larger than the gap it approaches the values \(\gamma_{\text{D}} = (-1)^{1+t} \kappa \pi\), \((t = 1, 2, \kappa = \pm 1)\) of the isolated surface Dirac cones. Here, \(\kappa = \pm\) correspond now to degenerate pairs (they become the T+, B− states for \(t = 1\) and B+, T− states for \(t = 2\) when \(t, \psi_k \to 0\)). The reduction of the Berry phase close to the mass gap of Dirac electrons leads to a violation of topological protection. Therefore WAL due to destructive interference caused by the Berry phase \(\mp \pi\) is suppressed and breaks down completely for large enough mass gap in ultrathin films with \(d \leq 5\)QL [16]. Furthermore the acquired degeneracy of film states opens the backscattering channel for the gapped states reducing the surface mobility [16] and also influencing the QPI signatures.

V. SELF-CONSISTENT \(\Gamma\)-MATRIX THEORY FOR IMPURITY SCATTERING

The STM-QPI method measures the electronic density fluctuations on the surface caused the interference of scattered and ingoing waves at an impurity site. It must be stressed that this is a single impurity effect [21] although the amplitude of the density fluctuations will be proportional to the number of impurities. In the situation of a thin film, another aspect is important. Due to the tunneling, the film states are eigenstates composed of surface states on both top and bottom surface. Therefore, even if the density fluctuations are measured on the top surface, they will also be influenced by the scattering on the bottom surface due to the tunneling. This effect, which is illustrated in Fig. 3, has to be included in the calculation. For the impurity scattering potential in spin \(\sigma\) and TB surface space \((\alpha)\), we assume the generic form

\[
\hat{V} = (V_{\sigma_0} + V_{\sigma_0} \sigma_{\alpha_0}) \alpha_0 = V \alpha_0; \\
V = \begin{pmatrix} V_{\uparrow} & 0 \\ 0 & V_{\downarrow} \end{pmatrix}; \\
V_{\uparrow, \downarrow} = V_{c} \pm V_{m}.
\]  (18)

Here, \(V_{c}\) denotes charge and \(V_{m}\) exchange scattering by normal and magnetic impurities, respectively, which does \textit{not} depend...
on the momentum transfer \( \mathbf{q} = \mathbf{k}' - \mathbf{k} \) of the scattering process. In this case a closed solution for the scattering \( t \)-matrix necessary to compute the QPI spectrum is possible in spin representation which then will have to be transformed to the helical eigenstate basis of the surfaces. The second exchange scattering term in Eq. (18) corresponds to a frozen impurity spin (along \( z \) direction), neglecting any spin dynamics such as treated in Ref. [37]. This spin polarization can be achieved either by tiny magnetic field or, since the impurity concentration is finite, may be the result of long-range exchange interactions (long relaxation times) for each impurity spin. For the potential in Eq. (18) we assume only intrasurface scattering (\( \sim \alpha_0 \)), i.e., cross scattering between opposite T,B surfaces is neglected. This is a reasonable simplification because, firstly the scattering in QPI experiments happens mostly via gas impurity atoms or molecules adsorbed above the surfaces. Therefore their perturbing potential should be constrained to the top and bottom surfaces. Since the overlap of T,B surface state wave functions is small at each surface no sizable to the top and bottom surfaces. Since the overlap of T,B surface state wave functions is small at each surface no sizable intersurface (T,B) scattering \( V' \) for \( d > 2QL \) is expected. Secondly, for zero momentum transfer \( \mathbf{q} \) a scattering between equal helicity states on opposite surfaces vanishes because of opposite spins for initial and final states (due to the reversed helicity sense on both surfaces). This means that the scattering \( V' \) has to be odd in \( \mathbf{q} \) whereas the intraplanar scattering \( V \) is even, taken as constant. Because the Dirac cones have small radius (\( k_c < \pi/a \)), all scattering processes on the cone have small momentum transfer and therefore \( V' \) will also be small for them.

### A. Closed solution in spin basis

For a momentum-independent scattering potential, the \( t \)-matrix equation in spin basis may be solved as

\[
\begin{align*}
\hat{t}_{kk}(i\omega_n) &= i(i\omega_n) = [1 - \hat{V}_D(i\omega_n)]^{-1}\hat{V} \\
&= R^{-1}\hat{V} = t_T \oplus t_B.
\end{align*}
\] (19)

We made the approximation that impurities on T,B do not scatter between the surfaces but only within each of them, i.e., we approximate \( R^{-1} = R_T^{-1} \oplus R_B^{-1} \) and therefore \( t_{TB} = t_{BT} = 0 \). However the propagation, i.e., tunneling between surfaces is fully taken into account by the nondiagonal Green’s function elements given below. Furthermore we assume that the type and scattering potential of impurities on T,B are the same and therefore set \( \hat{V}_T = \hat{V}_B = \hat{V} \). Then

\[
t = \begin{pmatrix}
t_{++} & t_{+\downarrow} \\
t_{+\uparrow} & t_{\downarrow\downarrow}
\end{pmatrix}; \quad V = \begin{pmatrix}
V_v + V_m & 0 \\
0 & V_v - V_m
\end{pmatrix};
\]

\[
\hat{g}(i\omega_n) = \frac{1}{N} \sum_k \tilde{G}_k(i\omega_n).
\] (20)

The spin flip amplitudes \( t_{i\downarrow}, t_{i\uparrow} \) that appear formally vanish identically (Appendix C) so that we only have to consider the diagonal elements \( t_{i\uparrow}, t_{i\downarrow} \) in the following. Ultimately this is due to the absence of spin-flip terms in the magnetic scattering term of Eq. (18). Furthermore the film Green’s function \( \tilde{G}_k(i\omega_n) \) in spin basis is given by

\[
\tilde{G}_k(i\omega_n) = [i\omega_n - H_k]^{-1}; \quad H_k = \begin{pmatrix}
h_k & t\sigma_0 \\
t\sigma_0 & h_k
\end{pmatrix}
\] (21)

with \( h_k \) defined in Eqs. (1) and (2). The individual \( t \)-matrix elements \( t_{i\sigma} \) may be evaluated (see Appendix C) from Eq. (19). It is also useful to introduce (anti)symmetric combinations by \( t_{i,a} = \frac{1}{2}(t_{i\uparrow} \pm t_{i\downarrow}) \). Furthermore we need the determinants \( D = \det[R_{T,B}] \), which are given by

\[
D = 1 - 2V_c\tilde{g}_D + (V_c^2 - V_m^2)\tilde{g}_D^2.
\] (22)

The momentum-integrated Green’s function in the above expressions which depend only on frequency \( i\omega_n \) are defined by [see Eq. (11)]:

\[
\tilde{g}_D(i\omega_n) = \frac{1}{N} \sum_k \frac{i\omega_n}{[i\omega_n^2 - E_k^2]};
\] (23)

the (anti) symmetric \( t \)-matrix elements are then obtained as

\[
t_{i\uparrow}(i\omega_n) = \frac{[V_c - (V_c^2 - V_m^2)\tilde{g}_D(i\omega_n)]}{1 - 2V_c\tilde{g}_D(i\omega_n) + (V_c^2 - V_m^2)\tilde{g}_D(i\omega_n)^2};
\]

\[
t_{i\downarrow}(i\omega_n) = \frac{V_m}{1 - 2V_c\tilde{g}_D(i\omega_n) + (V_c^2 - V_m^2)\tilde{g}_D(i\omega_n)^2}.
\] (24)

Obviously the antisymmetric amplitudes require a magnetic scattering and for normal scattering (\( V_m = 0 \), \( \theta_0 = 0 \)), we only obtain the symmetric scattering amplitude \( t_{i\downarrow} = V_c/(1 + V_c\tilde{g}_D) \). For this calculation, we used the spin basis because then the scattering potential is isotropic allowing for a closed solution. However, for later calculation of the QPI spectrum, we now have to transform back to the helical eigenbasis of the surfaces. The dynamics of the scattering matrix is determined by that of the momentum-integrated \( \tilde{g}_D(i\omega_n) \) which is shown in Fig. 6. The transparent closed form of the \( t \) matrix obtained in Eq. (24) which is essential for the following discussions was
only possible due to the restrictions of the model mentioned in the introduction. A more general t-matrix formalism including momentum dependent scattering and the realistic layer dependence of surface state wave functions would require a high-dimensional real-space representation of the Green’s functions and t matrix as, e.g., in Refs. [15,37,38]. This would require to solve the t-matrix equation and calculation of the resulting QPI spectrum within a fully numerical treatment that is beyond the present approach and intention of this work.

B. Transformation to helical basis

The transformation to helical basis is accomplished by the unitary matrix $S_k$ in Eq. (5) which is identical for T,B. The transformed scattering matrix

$$\tilde{t}_{kk} = S_k^\dagger t_{kk} S_k,$$

is then explicitly given in terms of the spin based solution Eq. (24) by its elements for both T,B blocks as

$$\tilde{t}_{++} = \alpha_{++}^{+} t_{++} + \alpha_{++}^{-} t_{--}; \quad \tilde{t}_{--} = \alpha_{--}^{-} t_{--} + \alpha_{--}^{+} t_{++};$$

$$\tilde{t}_{+-} = \alpha_{+-}^{+} t_{+-} + \alpha_{+-}^{-} t_{−+}; \quad \tilde{t}_{−+} = \alpha_{−+}^{−} t_{−+} + \alpha_{−+}^{+} t_{+-}.$$

The (T,B-independent) form factors $\alpha_j^{i}(i, j = \pm)$ that are functions of $(\phi_k, \theta_k, \phi_k', \theta_k')$ originating from the helical eigenstates are given explicitly in Appendix B. Before proceeding to the QPI spectrum, it is useful to discuss the general solution for the t matrix in the simple case of the Born approximation (BA) corresponding to single scattering from the impurity, i.e., to first order in $V_c$ and $V_m$.

C. Born approximation and selection rules for scattering

In BA, we have $t_c = V_c$ and $t_m = V_m$. Treating normal ($V_c$) and magnetic ($V_m$) scattering separately, we get the simple complementary results, respectively [cf. Eq. (B2)]:

$$\tilde{t}_{kk} = V_c \tilde{\alpha}_{kk} = V_c \left( \begin{array}{cc} \alpha_{++}^{+} & \alpha_{++}^{-} \\ \alpha_{--}^{+} & \alpha_{--}^{-} \end{array} \right)$$

and

$$\tilde{t}_{kk} = V_m \tilde{\alpha}_{kk} = V_m \left( \begin{array}{cc} \alpha_{++}^{+} & \alpha_{--}^{+} \\ \alpha_{--}^{-} & \alpha_{++}^{-} \end{array} \right).$$

First we consider the limiting case of isotropic Dirac cones, i.e., vanishing warping term $\lambda \rightarrow 0$. Then $\phi_k = \pi/2$ and $\cos \theta_k = \sqrt{2}/2$ and similar for $\theta_k'$. Explicitly we have

$$\tilde{t}_{kk} = V_c \left( \begin{array}{cc} 1 + e^{i(\theta_k - \theta_k')} & -ie^{i(\theta_k' - \theta_k)} \\ -i(e^{-i\theta_k' - \theta_k}) & 1 + e^{-i(\theta_k' - \theta_k)} \end{array} \right);$$

$$\tilde{t}_{kk} = \frac{V_m}{2} \left( \begin{array}{cc} 1 - e^{i(\theta_k' - \theta_k)} & ie^{i(\theta_k + \theta_k')} \\ -i(e^{-i\theta_k - \theta_k'}) & 1 - e^{-i(\theta_k' - \theta_k)} \end{array} \right).$$

In both cases, the symmetry relation $\tilde{t}_{kk}^\dagger = \tilde{t}_{kk}$ is fulfilled. It is instructive to look at the special examples of forward scattering $(f) \theta_k = \theta_k'$ and backward scattering $(b) \theta_k = \theta_k \pm \pi$. For charge scattering, this means

$$(f) : \tilde{t}_{kk} = V_c \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right)$$

and

$$(b) : \tilde{t}_{kk} = V_m \left( \begin{array}{cc} 0 & -ie^{i\theta_k} \\ ie^{-i\theta_k} & 0 \end{array} \right).$$

while for purely magnetic scattering the complementary result is

$$(f) : \tilde{t}_{kk} = V_m \left( \begin{array}{cc} 0 & ie^{i\theta_k} \\ -ie^{-i\theta_k} & 0 \end{array} \right)$$

and

$$(b) : \tilde{t}_{kk} = V_m \left( \begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right).$$

This demonstrates that backward scattering between same helicity states is forbidden for normal impurity scattering while it is allowed in the case of magnetic impurities. This will have direct influence on the QPI spectrum (Sec. VI).

It is also useful to derive the explicit expressions of form factors for forward $(f; \theta_k = \theta_k')$ and backward $(b; \theta_k = \theta_k + \pi)$ scattering including the effect of finite warping ($\lambda \neq 0$) when $\phi_k \neq 0$ in general. Using $\phi_k = \phi_k'$ in case $(f)$ and the identities $\sin \phi_k = \cos \phi_k'; \cos \phi_k = \sin \phi_k'$ in case $(b)$, we nevertheless get for normal scattering $V_c$ ($V_m = 0$) the identical result as in Eq. (29). Thus the spin mixing angle $\phi_k$ of the warped case does not enter in the $f,b$ amplitudes and the result is identical to the isotropic Dirac case. In particular, this means that the backscattering for equal helicity states remains forbidden even for the eigenstates in the warped Dirac cone. This is natural since the normal scattering does not react to the changed spin texture caused by the warping term.

On the other hand, for exchange scattering $V_m$ ($V_c = 0$), the change in spin texture will be important and therefore the result for the scattering matrix will depend on the spin mixing angle $\phi_k$ caused by warping. In BA, we obtain from the only
nonvanishing second term in Eq. (B3)

\[ f : \tilde{t}_{kk} = V_m \left( \begin{array}{ccc} \cos \phi_k & ie^{i\phi_k} \sin \phi_k \\ -ie^{-i\phi_k} \sin \phi_k & \cos \phi_k \end{array} \right) \]

and

\[ b : \tilde{t}_{kk} = V_m \left( \begin{array}{ccc} \sin \phi_k & -ie^{i\phi_k} \cos \phi_k \\ ie^{-i\phi_k} \cos \phi_k & -\sin \phi_k \end{array} \right). \]  (31)

In the isotropic cone limit (\( \lambda \rightarrow 0, \phi_k = \pi/2 \)), we recover the result of Eq. (30). We note that in the warped case when \( \phi_k \neq \pi/2 \) in general the forward scattering for equal helicity states no longer vanishes as in Eq. (30) due to the presence of the perpendicular spin component \( \sim \cos \phi_k \) of the eigenstates [Eq. (13)].

VI. QUASIPARTICLE INTERFERENCE SPECTRUM

The spectral Fourier component of the surface density modulation visible by STM-QPI is given by [21,39,40]

\[ \delta N_T(q, \omega) = -\frac{1}{\pi} \text{Im}[\Lambda_T(q, i\omega_n)]_{\omega_n \rightarrow -\omega + i\delta}, \]

\[ \Lambda_T(q, i\omega_n) = \frac{1}{N} \sum_k \text{tr}_\sigma [\tilde{G}_k/\tilde{t}_{kk} \tilde{G}_k]_{TT} \]

\[ = \frac{1}{N} \sum_k \text{tr}_\sigma [X_{kk}]_{TT}. \]  (32)

Here, \( \omega \equiv eV \) with \( V \) denoting the variable STM-tip bias voltage (Fig. 3). It is assumed that the tip is placed to the top (T) surface, therefore only the spin trace over the T block of \( t \) matrix and Green’s function product has to be performed. Explicitly we have two contributions given by

\[ [X_{kk}]_{TT} = \tilde{G}_k^{TT} \tilde{t}_{kk} \tilde{G}_k^{TT} + \tilde{G}_k^{TB} \tilde{t}_{kk} \tilde{G}_k^{BT}. \]  (33)

As it stands everything is still written on spin basis. Since the Green’s functions are diagonal in the helical bases one should transform to the latter, using Eq. (25) and the similar transformation for the Green’s function:

\[ \tilde{G}^{\alpha \alpha'}_{kk'} = S_{kk} \tilde{G}^{\alpha \alpha'}_{kk'} S_{kk'}. \]  (34)

where \( \alpha, \alpha' = \text{T}, \text{B} \). For the \( t \) matrix, we restricted to \( \text{T}, \text{B} \) diagonal elements only, neglecting the impurity scattering between the bottom and top surfaces. This cannot be done for the Green’s functions because the intersurface tunneling or hybridization is essential for the low energy surface states of the TI film. Therefore the \( \text{T}, \text{B} \) nondiagonal Green’s function elements have to be kept. This leads to two terms in the kernel Eq. (33) of the QPI spectrum involving scattering at the top (first term), and via hybridization of surface states also at the bottom (second term) surface. They are schematically shown in Fig. 3.

The \( \text{T}, \text{B} \) Green’s function blocks are all diagonal in the helical eigenstate basis. Therefore it is economic to transform Eq. (32) to this basis using Eqs. (25) and (34): The latter leads to

\[ \tilde{G}_k^{TT, BB} = \left( \begin{array}{ccc} A_{kk} & 0 \\ 0 & A_{kk} \end{array} \right) \]

and

\[ \tilde{G}_k^{TB} = \tilde{G}_k^{BT} = \left( \begin{array}{ccc} B_{kk} & 0 \\ 0 & B_{kk} \end{array} \right). \]  (35)

Where \( \text{T}, \text{B} \) correspond to upper and lower signs, respectively. Furthermore

\[ A_{kk} = \frac{i\omega_n \pm E_k}{(i\omega_n)^2 - E_k^2}; B_{kk} = \frac{t}{(i\omega_n)^2 - E_k^2}. \]  (36)

Using the diagonal Green’s functions the kernel in Eq. (33) may be transformed to helical basis as

\[ [X_{kk}]_{TT} = S_k [G_k^{TT} \tilde{t}_{kk} G_k^{TT}]_{\omega} + S_k [G_k^{TB} \tilde{t}_{kk} G_k^{BT}]_{\omega} \]

\[ = [X_{kk}]_{TT}^{(1)} + [X_{kk}]_{TT}^{(2)}. \]  (37)

Evaluating the two terms and inserting into Eq. (32), we obtain, correspondingly for (1) and (2):

\[ \Lambda_T^{(1)}(q, i\omega_n) = \frac{1}{N} \sum_k \left[ \alpha_+^+ \tilde{t}_{kk} \tilde{t}_{kk} \alpha_+^+ - \alpha_+^- \tilde{t}_{kk} \tilde{t}_{kk} \alpha_+^- \right] \]

\[ - \alpha_-^+ \tilde{t}_{kk} \tilde{t}_{kk} \alpha_-^- - \alpha_-^- \tilde{t}_{kk} \tilde{t}_{kk} \alpha_-^+; \]

\[ \Lambda_T^{(2)}(q, i\omega_n) = \frac{1}{N} \sum_k \left[ B_{kk} \tilde{B}_{kk} \right] \]

\[ \times \left[ \alpha_+^+ \tilde{t}_{kk} + \alpha_-^+ \tilde{t}_{kk} - \alpha_-^- \tilde{t}_{kk} - \alpha_+^- \tilde{t}_{kk} \right], \]  (38)

with the total result

\[ \Lambda_T(q, i\omega_n) = \Lambda_T^{(1)}(q, i\omega_n) + \Lambda_T^{(2)}(q, i\omega_n). \]

Here, \( \tilde{t}_{kk}(i\omega_n) \) is given by Eqs. (26) and (24).

A. Born approximation for QPI

To obtain a better insight in the general expression for the QPI spectrum we first simplify to the case of Born approximation with only single scattering events at the impurity included. In this case we have \( t_{kk'}^{\text{T},\text{B}} = V_c \) and \( t_{kk'}^{\text{T},\text{B}} = V_m. \) Again we treat normal (c) and magnetic (m) scattering separately. Inserting into Eq. (39) and using the explicit form factors and expressions for \( \tilde{A}_{kk} \), we obtain for normal scattering, after considerable algebra:

\[ \Lambda_c(q, i\omega_n) = \frac{2V_c}{N} \sum_k m_c^c(kk') \left[ (i\omega_n)^2 - E_k^2 \right] E_k E_{k'}. \]  (39)

with

\[ m_c^c(kk') = 1, \]

\[ m_c^m(kk') = \cos \phi_k \cos \phi_{k'} + \sin \phi_k \sin \phi_{k'} \cos(\theta_k - \theta_{k'}). \]
and likewise for magnetic scattering

\[
\Lambda_m(q, \omega_0) = \frac{2V_m}{N} \sum_k \frac{i \omega_0 [m^e_m(k) E_k + m^m_m(k') E_{k'}] + m^m_m(kk') E_k E_{k'}}{[(i \omega_0)^2 - E_k^2][(i \omega_0)^2 - E_{k'}^2]},
\]

(40)

with

\[
m^e_m(k) = \cos \phi_k, \quad m^m_m(k') = \cos \phi_{k'}, \quad m^m_m(kk') = -i \sin \phi_k \sin \phi_{k'} \sin (\theta_k - \theta_{k'}).\]

Note that in these expression the gapped cone energies \( \tilde{E}_k \) of the film appear in the denominator while the ungapped energies \( E_k \) of isolated surfaces remain in the numerator. It is instructive to consider the limit of the isotropic Dirac cone when \( \lambda \to 0 \) and then \( \phi_k = \pi/2 \). Inserting this into the form factors \( m^e_m(k, k') \) the above formulas then simplify to

\[
\Lambda_e(q, \omega_0) = \frac{2V_c}{N} \sum_k \frac{(i \omega_0)^2 + t^2 + \cos(\theta_k - \theta_{k'}) E_k E_{k'}}{[(i \omega_0)^2 - E_k^2][(i \omega_0)^2 - E_{k'}^2]},
\]

\[
\Lambda_m(q, \omega_0) = i \frac{2V_m}{N} \sum_k \frac{\sin(\theta_k - \theta_{k'}) E_k E_{k'}}{[(i \omega_0)^2 - E_k^2][(i \omega_0)^2 - E_{k'}^2]}.
\]

(41)

For the decoupled case \( t = 0 \), \( \tilde{E}_k = E_k \), these expressions can also be directly obtained from Eq. (32) by using the single-surface Green’s functions of Eq. (8) for the isotropic case. Due to the sign change of the numerator of \( \Lambda_m \) under \( k \leftrightarrow k' \) which results from the helical spin locking, we have \( \Lambda_m \equiv 0 \) in Born approximation. For a finite result, one has to use the full \( t \)-matrix theory.

**B. Full \( t \)-matrix expressions QPI spectrum**

For computational convenience we also give here the rather lengthy explicit expression for the QPI spectrum including the general \( t \)-matrix elements that will mostly be used in the numerical calculations. The total QPI spectrum, equal to the sum \( \Lambda_T(q, \omega_0) = \Lambda_T^{(1)}(q, \omega_0) + \Lambda_T^{(2)}(q, \omega_0) \) from Eq. (39) is then obtained as

\[
\Lambda_T(q, \omega_0) = \frac{1}{N} \sum_k \frac{1}{[(i \omega_0)^2 - E_k^2][(i \omega_0)^2 - E_{k'}^2]}
\]

\[
\times \left[ \tilde{f}_{\alpha-}((i \omega_0 + E_k)(i \omega_0 + E_{k'}) + t^2) + \tilde{f}_{\alpha+}((i \omega_0 - E_k)(i \omega_0 - E_{k'}) + t^2)
\right.
\]

\[
- \tilde{f}_{\alpha-}((i \omega_0 + E_k)(i \omega_0 - E_{k'}) + t^2)
\]

\[
- \tilde{f}_{\alpha+}((i \omega_0 - E_k)(i \omega_0 + E_{k'}) + t^2)\right]
\]

(42)

In each term, the first part describes the QPI contribution on the top surface due to scattering on top surface while the second part proportional to \( t^2 \) represents the QPI contribution on the top surface due to scattering on the bottom surface. Obviously this term can only be present when there is tunneling (described by \( G_{TB}, G_{BT} \) in Fig. 3) between the T,B surface states, therefore it vanishes for decoupled surfaces. In these equations the expressions for the full \( t \)-matrix elements

**FIG. 7.** QPI spectrum in Born approximation [Eq. (40)] for normal scattering \( V_t \equiv 1 \) only (\( V_c = 0 \)). The images are for various bias voltages \( eV = \omega \) (rows) and intersurface tunneling strength \( t \) (columns). When the latter is increased the typical QPI structure appear only for \( \omega > |t| \) due to the vanishing of constant energy surfaces for smaller frequencies (Fig. 1). The details of the QPI images are explained in Sec. VII.

\( \tilde{f}_{\alpha-}((i \omega_0) \) in helical presentation are obtained from Eq. (26). They are composed of the irreducible \( t_{\alpha\alpha'} \) matrix elements in Eq. (24) and the form factors \( \alpha_{\alpha\alpha'} \) of Appendix B.

**VII. NUMERICAL RESULTS AND DISCUSSION**

We now discuss the numerical results for the QPI spectrum under systematic variation of bias voltage \( eV = \omega \), intersurface tunneling \( t \) of the thin film and normal (\( V_t \)) and magnetic (\( V_m \)) impurity scattering potential (assumed identical on both surfaces). We mostly use the full \( t \)-matrix approximation as given in closed form by Eq. (42) in this section except when stated otherwise. In Fig. 7, we show an overview over the images for normal \( V_t \)-type scattering where the intersurface tunneling \( t \) varies along the columns and the bias voltage \( eV = \omega \) (measured from the Dirac point) along the rows. Here, for once, the Born approximation [Eq. (40)] is employed for comparison. Generally in QPI images the Fermi surface for a given \( \omega \) is reproduced, with a doubling of the FS radius. However, the intensities at wave vectors connecting special Fermi surface points like tips and dents may be strongly enhanced or depressed. The overall extension of the QPI images increases with \( \omega \), the distance from the Dirac point, according to the diameter of the cut \( \omega = \tilde{E}_k \) (or \( \omega = \tilde{E}_k \)) through the warped cone which gives the snowflake FS. The shape of the pattern and its increasing radius with \( \omega \) is clearly seen in the first row representing the isolated surfaces \( t = 0 \). When the tunneling is turned on (second and third rows) the low-energy spectrum is gapped and therefore around \( \omega \approx t \) (if we use \( \omega, t \approx 0 \)) the QPI image will be strongly modified: the radius shrinks and the anisotropic “snowflakes” character is reduced leading to a more isotropic image. This is completely
in accordance with the evolution of the Fermi surfaces in Fig. 1. Therefore QPI investigation of thin films can give full information on the thickness (tunnelling) dependence of low energy quasiparticles close to the gap threshold. For larger energies (voltages) $\omega \gg |t|$ above the gap, the QPI image approaches that of the isolated surfaces (cf. the two top right figures).

A detailed comparison of Fermi surface shape [first row (a) and (d)] and according QPI image is presented in Figs. 8(a) to 8(f) for normal [second row (b) and (e)] as well as magnetic [third row (c) and (f)] scattering mechanism. Similar to Refs. [5,11] we can identify characteristic wave vectors $\mathbf{q}_i$ ($i = 1-7$) in Figs. 8(a) and 8(d) connecting special points where the azimuthal group velocity vanishes (tips and dents of the snowflake) and should therefore figure prominently in the QPI image. Indeed most of the $\mathbf{q}_i$ vectors (with the exception of $\mathbf{q}_3$) can be identified in Figs. 8(b) and 8(c) for the isolated surface corresponding either to large (dark red) intensity or low and vanishing (deep blue) intensity areas with arrows representing the characteristic wave vectors pointing to them. Some of the latter in Fig. 8 correspond to the forbidden backscattering as they are connected by $\mathbf{q}_1$, $\mathbf{q}_4$ scattering vectors that pass through the origin such that $\mathbf{k}' = -\mathbf{k}$. These regions are, however, quite narrow because of the vicinity (of equivalent) closely allowed $\mathbf{q}_6$, $\mathbf{q}_7$ vectors that are associated with large intensities. The $\mathbf{q}_2$, $\mathbf{q}_5$ scattering vectors can also be identified though less prominently. The overall pattern for magnetic scattering in (c) looks quite similar although the pointwise intensities are largely different. In particular for magnetic impurities the backscattering vectors $\mathbf{q}_1$, $\mathbf{q}_4$ are allowed (Secs. II C and VI A) and are also associated with finite QPI spectral intensity.

There is another weak, though noteworthy peculiar feature of the QPI image: The single snowflake of the FS experiences a doubling in the QPI image. The outer one has the same orientation as the FS and is associated mostly with scattering vectors close to the group $\mathbf{q}_1$, $\mathbf{q}_4$, $\mathbf{q}_6$, $\mathbf{q}_7$ and their equivalents while the inner smaller image is rotated by $\pi/3$ with respect to the FS and is mostly associated with scattering vectors close to $\mathbf{q}_2$, $\mathbf{q}_5$ and equivalents.

When the intersurface tunneling $t$ is turned on and the gap is opened at the Dirac point, at first for $\omega \gg |t|$, there is no drastic change in QPI image characteristics as seen from the upper right corner of Fig. 7, except that the characteristic scattering vector lengths $|\mathbf{q}_i|$ are shrinking. This becomes more dramatic when $\omega$ approaches the (half-) gap size $|t|$ and the FS structure becomes more circular (i). Furthermore the tunneling opens backscattering channels close to the gap edge increasing the relative size of the high intensity areas (dark red). As a result the QPI snowflake image gradually “melts” into a circular drop shape where only the most prominent scattering vectors $\mathbf{q}_6$, and $\mathbf{q}_7$ can still be discerned. Their length is now close to $2k_F$ of the nearly spherical Fermi surface image [Figs. 7(d), 7(i) and 8(e), 8(f)] corresponding to the parabolic surface dispersion close to the gap edge (cf. Fig. 1). This gradual transformation of the QPI image from the six-prong (double) snowflakes at zero or small tunneling (isolated surfaces) to the more isotropic, almost circular structures at large tunneling (few quintuple layer films) should be worthwhile to investigate experimentally with STM-QPI technique for a number of reasons. Firstly it has not yet been observed in ARPES experiments [19]. Furthermore in QPI this transition contains direct visual evidence for the reappearance of backscattering, as a function of increasing $t$, i.e., decreasing film thickness $d$ as clearly seen in Figs. 7 and 8. This observation is central to the change of surface state topology with thickness [witnessed also by the Berry phase of Eq. (17)]. For $d > 5$ QL, the backscattering is largely forbidden and QPI intensity is strongly suppressed for backscattering vectors. For $d < 5$ QL, it reappears and as a consequence the QPI image becomes more circular (due to the gap opening) with more isotropic intensity distribution. Therefore this change demonstrates pointwise in $k$ space how the backscattering reappears as a function of film thickness. This is not directly possible by ARPES which probes just single particle spectral function and also not by magnetotransport which can only indirectly conclude (via destruction of weak antilocalization [16]) on this mechanism.

Instead of looking at the QPI image of $|\Lambda(q_x, q_y, \omega)|$ in the 2D SBZ as before we may consider the QPI “quasiparticle
dispersion” defined by the complementary image $|\Lambda(q_x, q_y, \omega)|$ in the $(q_x, q_y)$ plane for fixed momentum direction $\mathbf{q} \parallel \Gamma K$ or $\Gamma M$ in the SBZ which should directly demonstrate the change of surface quasiparticle spectrum with tunneling strength. The dispersion results for normal scattering are shown in Fig. 9. For isolated surfaces ($t = 0$), the Dirac cone dispersion can clearly be identified by the envelope high intensity region. Furthermore the interior of the structure acquires a closing of the gap for intermediate $\omega$ and $\beta$ surfaces connected by the tunneling, the gap opens progressively and the parabolic dispersion for $\omega > |t|$ is again seen in by a sharp prominent high intensity envelope. The destruction of the Dirac cone by the gap opening via intersurface tunneling has so far been observed directly only in ARPES experiments for Bi$_2$Se$_3$ thin films between 1–6 QL thickness [19] and it would be highly desirable to investigate this also by the complementary QPI method as proposed here. The existence of the Dirac cone dispersion for the isolated surfaces ($t = 0$ in Fig. 9) has indeed been demonstrated before with the QPI method [6–8].

In an alternative constant-$\omega$ presentation of QPI images $|\Lambda(q_x, q_y, \omega = eV)|$, we again keep the bias voltage fixed and change the tunneling $t$ and scattering strengths $V_c$, and $V_m$. This is presented in Fig. 10, where we show the images for a constant $\omega = 1.25$ and as a function of $t$ (columns) and relative normal ($V_c$) to magnetic ($V_m$) scattering strength (rows). While the overall image features are similar, the substructure of the large intensity regions (dents) depends in a subtle way on the relative size of $V_c$, and $V_m$. Finally, in Fig. 11, we show an example of the evolution of QPI spectrum for normal scattering from very low (BA) to very strong scattering strength. The global structure of the image is preserved but details show again subtle changes: The high intensity spots for small scattering merge into a ring for large $V_c$, but still keeping low intensity at the backscattering vectors. Furthermore the interior of the structure acquires more intensity presumably due to the increasing importance of multiple scattering processes with growing $V_c$ contained in the full $t$-matrix approach.

So far the tunneling matrix element $t$ has been treated as an arbitrary but a fixed parameter. It is, however, an effective (film-) surface states parameter that derives from the solution of the true boundary value problem [23,24] starting from the valence (VB) and conduction band (CB) states of the bulk Hamiltonian given in $k \cdot p$ parametrized form [33,41]. Using these bulk parameters it was shown in Ref. [24] that the dependence of $t(d)$ on film thickness $d$ should be given by
t(d) = t_0 \exp \left(-\frac{d}{d_0}\right) \sin \left(\frac{d}{d_0}'\right), \quad (43)
decl where the reference energy $t_0[E^+]$ and length $d_0$ [QL], $d_0'$ [QL] scales are given in terms of the $k \cdot p$ parameters of the true bulk [33,41]. The $t(d)$ dependence contains an overall exponential decrease with increasing $d$ characterized by the decay length $d_0$ but also an oscillatory term governed by $d_0'$ (oscillation period $2\pi d_0$). The oscillatory term leads to a closing of the gap for intermediate $d$ values [24] $d = n\pi d_0(n = 1, 2, \ldots)$. From Ref. [24], we derive the following set of theoretically predicted energy $E^+$ and length [QL] scales for Bi$_2$Se$_3$: $(t_0, d_0, d_0')_{b} = (2.68, 0.64, 0.9)$ and Bi$_2$Te$_3$: $(t_0, d_0, d_0')_{b} = (0.80, 1.79, 0.3)$. For the first

![FIG. 9. The dispersion characteristics of QPI spectrum along symmetry directions of the SBZ $\Gamma \rightarrow \bar{M} \rightarrow \bar{\Gamma}$ as a function of intersurface tunneling strength. The gap opening of the Dirac cones can clearly be followed in the QPI dispersion.](image)

![FIG. 10. QPI spectrum of thin film for frequency $\omega = 1.25$ for various tunneling strengths $t$ (columns) and impurity scattering strengths $(V_c, V_m)$ (rows).](image)

![FIG. 11. The evolution of QPI pattern from small to large normal scattering potential $V_c$ at the energy $\omega = 0.75$ for the case of $t = 0.4$.](image)
compound, this would lead to a vanishing of \( t(d) \) for intermediate \( d \approx 3 \) due to a single oscillation in the range of sizable \( t(d) \). This is not observed in ARPES experiments in Bi\(_2\)Se\(_3\) which show no gap oscillation or closing [19] but rather a monotonous decrease of \(|t(d)|\) in the range \( 2QL < d < 7QL \) where it vanishes on the upper value, restoring the isolated Dirac surface states. This is shown by the symbols connected by a red line in Fig. 12. There is a natural source for the origin of the discrepancy. The bulk \( k \cdot p \) parameters which depend on the position of bulk CB and VB edge are assumed as independent of thickness \( d \) for the film. However, this is unrealistic because band-bending effects due to both surfaces lead to thickness-dependent CB and VB edges as indeed observed in ARPES [5] in addition to a strongly \( d \)-dependent energy position of the Dirac point on surface states. Therefore one has to expect that realistic parameters for \( t(d) \) may be quite different from the theoretical ones given above.

In a more practical strategy, we take the observed monotonous exponential decay of \( t(d) \) as real and derive the parameter set by fitting to the ARPES results for Bi\(_2\)Se\(_3\) [5] corresponding to the blue line in Fig. 12. This leads to the partial experimental set for Bi\(_2\)Se\(_3\): \((t_0, d_0)_{\text{exp}} = (1.0, 1.46)\). Assuming the theoretical ratio \( d_0'/d_0 = 1.41 \) we then obtain \( d_0' = 2.05 \). Because \( d_0' \) is now considerably larger than the purely theoretical value the first zero of \( t(d) \) due to the oscillatory term is shifted to \( d \) values where the exponential term has already reduced the gap close to zero. This means the oscillatory dependence of \( t(d) \) is no longer observable. Therefore we suggest there is no fundamental contradiction between ARPES results and theoretical model for the inter-surface tunneling but rather the effective theoretical parameter values do not correspond to the real ones, possibly due to the effect of the observed strong band bending effects found for VB and CB in this compound.

For Bi\(_2\)Te\(_3\) so far no systematic ARPES results for thin films are available that could be used to constrain the parameter set for \( t(d) \). The theoretical values given above predict a much smaller oscillation period \( 2\pi d_0' \) of \( t(d) \) for the Bi\(_2\)Te\(_3\) compound which would lead to many zeros of the tunneling gap in the thickness range \( d = 2–8 \) as discussed below. Therefore even if the real parameter set is again different the chances to observe oscillatory behavior should be much higher in Bi\(_2\)Te\(_3\).

Now we discuss the expected QPI pattern dependence on thickness which may show dramatic effects depending on the hybridization gap or \( t(d) \) behavior. Let us assume for each thickness the bias voltage is kept at a constant value with respect to the Dirac point in each film because the absolute value of the latter itself depends on \( d \) as shown in Ref. [5]. Then, as the film thickness is changed the gap strongly varies and the according QPI pattern should change between relatively isotropic circular pattern for large \( |t| \) and strongly six-prong snowflake patterns for small \( |t| \). One may have the two principal cases.

(i) For monotonous gap dependence as in Bi\(_2\)Se\(_3\), the QPI pattern should also change monotonically from the almost circular shape with isotropic intensity to the snowflake shape with sixfold peaked intensities due to the forbidden backscattering in the latter.

(ii) For oscillatory behavior of the gap, as possibly realized in Bi\(_2\)Te\(_3\), the QPI images can also oscillate between more isotropic and snowflake type patterns. This depends sensitively on the precise positions of the zeros and maxima of \( t(d) \). If the former happen to lie close to an integer value of \( d \) (the ones achievable for real films), the gap will be zero and the snowflake structure will appear although \( t_0 \) is finite. If, on the other hand, the maxima of \( t(d) \) come to lie close to one of these \( d \) values, the QPI pattern shows more isotropic form. In the general case when zeros and maximum positions of \( t(d) \) shift with respect to the integer \( d \) values oscillatory behavior of QPI pattern between isotropic and snowflake type may be expected. Observation of such sequence in STM-QPI experiments should be taken as direct proof for the gap oscillation.

First we discuss case (i) which is realized in Bi\(_2\)Se\(_3\) (Fig. 12). The symbols present the tunneling \(|t|\) or gap size \((\Delta = 2|t|)\) as a function of thickness \( d[QL] \) as determined by ARPES [5]. The red line is a guide to the eye and the blue line is a fit with Eq. (43) taking \( t_0 \), \( d_0 \) as free parameters and assuming the theoretical value \( d_0'/d_0 = 1.41 \) for Bi\(_2\)Se\(_3\). This shifts the zero due to the oscillatory term to a thickness \( d = 6–7 \) (arrow) where \(|t|\) has already been exponentially suppressed and cannot be observed. For the theoretical values of \( t_0 \), \( d_0 \) the zero would have appeared at \( d = 2.8 \). In the
FIG. 13. Oscillatory (half-) gap \(|t(d)|\) (units \(E^*\)) dependence on film thickness \(d\) (units of \(QL\)) obtained from Eq. (43) corresponding to a case of rapid oscillation of \(|t(d)|\). Concomitant oscillations in QPI spectra (upper row inset, \(\omega = 0.5\)) between snowflake \((b), (d),\) and \((c)\) and isotropic shape \(\{(a)\}\) and \(\{(e)\}\) depends critically on the positions of zeros and maxima with respect to the possible film thickness \(d = 2, 3, \ldots\) Red circles correspond to actual values of \(|t|\) in the row images. The central image \((f)\) shows the distinct intensity \(|A(q, \omega)|\) behavior for two directions due to different backscattering rules as a function of thickness. For the highest \(t = 0.6 > 0\), it collapses to a bell shape curve because there are no more scattering processes on the equal-energy surface \(E_k = \omega\).

detailed analysis of the QPI intensity given in the central inset provides even more insight. Here \(|A(q, \omega = 0.5)|\) is plotted for two directions \(q = (q_x, 0)\) and \(q = (0, q_y)\). From Fig. 8(b), one can see that \(q_x\) direction corresponds to allowed scattering vectors \(q_0\) whereas \(q_y\) direction corresponds to (equivalent) \(q_1\) vectors forbidden due to backscattering. Therefore, for small \(|t|\) in the inset (red curves), the intensity along \(q_x\) shows a steep drop when crossing the \(2k_F\) value while it has a sharp spike along \(q_y\). For thickness values with larger \(|t|\), the forbidden backscattering reappears and therefore the spike is now seen also for \(q_y\) direction (blue curve). Therefore the detailed intensity analysis of QPI pattern as a function of film thickness can give a pointwise \((in k\) space) information on the backscattering variation with intersurface tunneling strength and the associated topological character of near-gap hybridized surface states.

VIII. SUMMARY AND CONCLUSION

In this work, we proposed a theory of quasiparticle interference for topological insulator thin films. We start from a model of single Dirac cone surface states with a warping term included to reproduce the realistic snowflake FS shape of the isolated TI surface in compounds like Bi$_2$Se$_3$ and Bi$_2$Te$_3$. The main physical effect of thin film geometry is implemented within a model where equal helicity states can tunnel from one surface to the other, described by the tunneling energy \(t\). This has two consequences: a gap opening is observed due to the (momentum-dependent) mixing of top/bottom states and the helical spin texture and Berry phase of Dirac states is strongly modified close to the gap edge at \(\omega = |t|\), leading to a vanishing of the in-plane spin component and Berry phase.

The warped surface cones have been previously found for isolated surfaces both by ARPES and QPI methods. The gap opening of surface states in thin films as a function of thickness has so far only been directly seen with the former method. Here we investigated what is to be expected for QPI experiments. To calculate the QPI signal we use a completely analytical \(t\)-matrix calculation with a full summation to infinite order in the scattering strength as well as the Born approximation. To carry out this approach, it is necessary to restrict to a tractable model of impurity scattering and four surface basis states connected by the intersurface tunneling. For the former, we employ a generic potential consisting of momentum-independent intraplane scalar and exchange scattering whose strengths are model parameters. We find a closed solution of the \(t\) matrix including the effect of interplane propagation. The impurity scattering and the tunneling to the STM tip is expressed in the spin basis whereas the eigenstates of the warped cones have a complicated helical spin texture. Therefore the analytical expressions involve a summation over terms with different azimuthal form factors which also control the behavior under backward scattering.

The QPI images can be interpreted in terms of characteristic scattering vectors connecting tips/dents in the snowflake FS for a given bias voltage. The allowed scattering and forbidden backscattering can be identified for normal impurities as long as the interplane tunneling is moderate. For constant \(\omega\) the latter, when the gap size increases, shrinks the dimension of the QPI image and its anisotropic features. This turns the
six-prong snowflakes of QPI with strong intensity anisotropy into a more isotropic circular 2kF image of the Fermi surface of the gapped Dirac cones, also due to reopening of backscattering. Once \( \omega < |t| \) falls below the (half-) gap size all specific QPI features are suppressed. Such a pointwise (in \( k \) space) observation of the changing backscattering properties, which characterize the topological nature of hybridized gap threshold states evidenced by the Berry phase, is a unique ability of the QPI method and would certainly be worthwhile to discover. A further strong incentive to carry out such an investigation is suggested by previous surface state calculations in \( Bi_2Se_3 \) and \( Bi_2Te_3 \) that have shown \( t(d) \) to behave nonmonotonically or even oscillate with film thickness \( d \). This would imply that as thickness is reduced continuously the QPI image of TI surface state will oscillate between anisotropic snowflake and nearly isotropic circular patterns. Previous ARPES experiments in \( Bi_2Se_3 \) have, however, found only a monotonous reduction of the gap with increasing film thickness. As discussed in Sec. VII this advocates that the empirical parameters that describe the tunneling energy \( t(d) \) are different from theoretical ones such that the oscillatory period \( t(d) \) is much longer (Fig. 12), possibly due to band bending effects. However, \( Bi_2Te_3 \) may be a more favorable case because it is predicted to have a higher number of \( t(d) \) oscillations, and some of them might survive even when again the empirical parameters for the tunneling are different from the calculated ones. As shown in Fig. 13, the sequence of QPI patterns for various \( d = 2, 3, \ldots \) depends sensitively on the oscillation period and the possible coincidence of film thickness with the zeros or maxima of \( t(d) \). Therefore QPI method would be able to identify the predicted oscillatory behavior of tunneling induced gap in \( Bi_2Te_3 \). Furthermore the dispersion of surface state energies may be extracted from QPI thin film images in \( (q, \omega) \) plane for fixed direction of scattering vector \( \hat{q} \). It shows the breaking of the Dirac cone by gapping and appearance of quadratic low energy dispersion which is anisotropic due to the effect of warping. Their observation would give a further instructive comparison to ARPES results as a function of film thickness. Finally we want to mention that a first indication of the thin film effects discussed in this work may have recently been observed in a different context [42]. The gap opening of Dirac cones was clearly identified by QPI method in the multilayer \( MnBi_3Te_7 \) where it is induced by hybridization of surface states with adjacent inequivalent layers rather than the opposite surface as is the case in the thin films discussed here.

ACKNOWLEDGMENTS

A.A. acknowledges financial support from the National Research Foundation (NRF) funded by the Ministry of Science of Korea (Grants No. 2016K1A4A01922028, No. 2017R1D1A1B03033465, and No. 2019R1H1A2039733).

APPENDIX A: EXPLICIT HELICITY EIGENSTATES OF THE FILM

The eigenvectors for coupled surface states given in Eq. (12) form the columns of the unitary transformation \( W_k \) that diagonalizes the coupled layer 4 spin Hamiltonian \( H_k \) according to

\[
E_k = W_k^\dagger H_k W_k = E_k^\tau \otimes k_0 \quad \text{where} \quad \tau = 1, 2 \quad \text{is the upper/lower band index and} \quad k = \pm \quad \text{the helicity. It is given explicitly by}
\]

\[
W_k = \begin{pmatrix}
\cos \psi_k \cos \frac{\phi_k}{2} + i \sin \psi_k \sin \frac{\phi_k}{2} e^{-i\theta_k} & i \cos \psi_k \sin \frac{\phi_k}{2} - \sin \psi_k \cos \frac{\phi_k}{2} e^{-i\theta_k} \\
-\sin \psi_k \cos \frac{\phi_k}{2} + i \cos \psi_k \sin \frac{\phi_k}{2} e^{-i\theta_k} & \cos \psi_k \sin \frac{\phi_k}{2} + i \sin \psi_k \cos \frac{\phi_k}{2} e^{-i\theta_k}
\end{pmatrix}
\]

where the sequence of columns corresponds to \( W_k = \left( |\psi_{1+}\rangle, |\psi_{1-}\rangle, |\psi_{2+}\rangle, |\psi_{2-}\rangle \right) \) ordering of coupled surface eigenstates \( |\psi_{kk}\rangle \) with energies \( E_k, E_k', -E_k, -E_k' \), respectively.

APPENDIX B: FORM FACTORS FOR THE t-MATRIX

In this Appendix, we list the explicit expressions for the form factors \( a_{\alpha\alpha'}^k \), where \( \kappa (\kappa') \) denotes the helicity quantum numbers. We have

\[
\begin{align*}
\alpha_{++}^+ &= \cos \frac{\phi_k}{2} \cos \frac{\phi_k'}{2} \pm \sin \frac{\phi_k}{2} \sin \frac{\phi_k'}{2} e^{i(\theta_k - \theta_k')}, \\
\alpha_{+-}^- &= \cos \frac{\phi_k}{2} \cos \frac{\phi_k'}{2} \pm \sin \frac{\phi_k}{2} \sin \frac{\phi_k'}{2} e^{-i(\theta_k - \theta_k')}, \\
\alpha_{+-}^+ &= i \left( \cos \frac{\phi_k}{2} \sin \frac{\phi_k'}{2} e^{i\theta_k'} \pm \cos \frac{\phi_k}{2} \sin \frac{\phi_k'}{2} e^{-i\theta_k'} \right), \\
\alpha_{++}^- &= -i \left( \cos \frac{\phi_k}{2} \sin \frac{\phi_k'}{2} e^{-i\theta_k'} \pm \cos \frac{\phi_k}{2} \sin \frac{\phi_k'}{2} e^{i\theta_k'} \right).
\end{align*}
\]

Here, \( \theta_k \) is the azimuthal angle of the \( k \) vector and \( \phi_k \) is defined in Eq. (6). In these expressions, the upper indices \( \pm \) on the left-hand side are correlated with the sign choice on the right-hand side. We can arrange the form factors in \( 2 \times 2 \) matrices \( \bar{a}_k, \bar{\alpha}_k \) according to

\[
\bar{a}_{\alpha\alpha'} = \begin{pmatrix} \alpha_{++}^+ & \alpha_{+-}^- \\ \alpha_{+-}^+ & \alpha_{++}^- \end{pmatrix}, \quad \bar{\alpha}_{\alpha\alpha'} = \begin{pmatrix} \alpha_{+-}^- & \alpha_{++}^+ \\ \alpha_{++}^- & \alpha_{+-}^+ \end{pmatrix}.
\]

In terms of these form factor matrices, the scattering matrix of Eq. (26) is then given in shorthand notation by

\[
\bar{t}_{kk'} = \begin{pmatrix} \bar{t}_{++} & \bar{t}_{+-} \\ \bar{t}_{-+} & \bar{t}_{--} \end{pmatrix} = t_s(i\omega_0)\bar{a}_{kk'} + t_s(i\omega_0)\bar{\alpha}_{kk'},
\]

where the (equal) top (T) and bottom (B) (anti) symmetrized \( t \)-matrix elements \( t_{\alpha\alpha'}(i\omega_0) \) that depend only on frequency are given in Eq. (24) for the general case.
APPENDIX C: FORMS OF THE $R$ AND $t$ MATRICES IN SPIN REPRESENTATION

For the calculation of the $t$-matrix elements in spin basis we need to calculate $R^{-1} = [1 - V \hat{g}(i\omega_n)^{-1}]$ in Eq. (19). Restricting to intrasurface ($T \leftrightarrow T, B \leftrightarrow B$) scattering (which is exact in Born approximation) we only have to consider the diagonal blocks (T/B correspond to upper/lower signs, respectively). Using the auxiliary functions defined in Eq. (C3), we obtain for the $R$ matrix

$$R_{T,B} = \begin{pmatrix} [1 - V_1 \hat{g}_0 \pm V_1 f_\epsilon] & -V_1 f_\epsilon^* \\ -V_1 f_\epsilon & [1 - V_1 \hat{g}_0 \pm V_1 f_\epsilon] \end{pmatrix},$$  \hspace{1cm} \text{(C1)}$$

where we defined $V_{1 \pm} = V_1 \pm V_2$. The (complex) auxiliary function $f_\epsilon, f_\epsilon^*$ are defined by

$$f_\epsilon(i\omega_n) = i \frac{1}{N} \sum_k F_\epsilon(i\omega_n, \theta_k) e^{-i\theta_n};$$  \hspace{1cm} \text{(C2)}$$

$$F_\epsilon(i\omega_n, \theta_k) = \frac{\sin \phi_k \tilde{E}_k}{(i\omega_n)^2 - \tilde{E}_k^2};$$

and

$$f_\epsilon(i\omega_n) = \frac{1}{N} \sum_k F(i\omega_n, \theta_k);$$

$$F(i\omega_n, \theta_k) = \frac{\cos \phi_k \tilde{E}_k}{(i\omega_n)^2 - \tilde{E}_k^2}.$$  \hspace{1cm} \text{(C3)}$$

From the expressions for $\phi_k, E_k, \tilde{E}_k$ given in Sec. II, one may show that $F_\epsilon(i\omega_n, \theta_k)$ is periodic with period $\frac{\pi}{2}$, i.e., $F_\epsilon(i\omega_n, \theta_k + \frac{\pi}{2}) = F_\epsilon(i\omega_n, \theta_k)$, whereas $F(i\omega_n, \theta_k)$ is antiperiodic with period $\frac{\pi}{2}$, i.e., $F(i\omega_n, \theta_k + \frac{\pi}{2}) = -F(i\omega_n, \theta_k)$. From this property, it follows immediately that $f_\epsilon(i\omega_n) \equiv 0$, and due to the different periodicities in the integral for $f_\epsilon$ also $f_\epsilon \equiv 0$. Therefore $R^{-1}$ may be easily obtained and finally from Eq. (19) (the diagonal) $t$-matrix elements in spin representation are given by

$$t_{\uparrow \uparrow} = \frac{V_\epsilon + V_m}{1 - (V_\epsilon + V_m)g_0}, \hspace{1cm} t_{\uparrow \downarrow} = \frac{V_\epsilon - V_m}{1 - (V_\epsilon - V_m)g_0},$$  \hspace{1cm} \text{(C4)}$$

where $g_0(i\omega_n)$ is given in Eq. (24). For the $t$ matrix in helical basis representation [Eq. (26)], it is then more convenient to use (anti)symmetric combinations defined by $t_{\alpha \beta} = \frac{1}{2} (t_{\uparrow \uparrow} \pm t_{\uparrow \downarrow})$, which are explicitly written in Eq. (24).
[29] J. Wang and B.-F. Zhu, Chin. Phys. B 22, 067301 (2013).
[30] A. Bansil, H. Lin, and T. Das, Rev. Mod. Phys. 88, 021004 (2016).
[31] M. Z. Hasan and J. E. Moore, Ann. Rev. Condens. Matter Phys. 2, 55 (2011).
[32] Y. Ando, J. Phys. Soc. Jpn. 82, 102001 (2013).
[33] C.-X. Liu, H. J. Zhang, B. Yan, X.-L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. B 81, 041307(R) (2010).
[34] Z.-H. Zhu, C. N. Veenstra, G. Levy, A. Ubaldini, P. Syers, N. P. Butch, J. Paglione, M. W. Haverkort, I. S. Elfimov, and A. Damascelli, Phys. Rev. Lett. 110, 216401 (2013).
[35] D. Xiao, M.-C. Chang, and Q. Niu, Rev. Mod. Phys. 82, 1959 (2010).
[36] A. A. Taskin and Y. Ando, Phys. Rev. B 84, 035301 (2011).
[37] P. G. Derry, A. K. Mitchell, and D. E. Logan, Phys. Rev. B 92, 035126 (2015).
[38] Z.-G. Fu, P. Zhang, Z. Wang, and S.-S. Li, J. Phys. Condens. Matter 24, 145502 (2012).
[39] A. Akbari, P. Thalmeier, and I. Eremin, Phys. Rev. B 84, 134505 (2011).
[40] A. Akbari and P. Thalmeier, Europhys. Lett. 102, 57008 (2013) Eur. Phys. J. B 86, 495 (2013).
[41] M. M. Asmar, D. E. Sheehy, and I. Vekhter, Phys. Rev. B 95, 241115(R) (2017).
[42] X. Wu, J. Li, X.-M. Ma, Y. Zhang, Y. Liu, C.-S. Zhou, J. Shao, Q. Wang, Y.-J. Hao, Y. Feng, E. F. Schwier, S. Kumar, H. Sun, P. Liu, K. Shimada, K. Miyamoto, T. Okuda, K. Wang, M. Xie, C. Chen, Q. Liu, C. Liu, and Y. Zhao, arXiv:2002.00320.