# Electron beam splitting at topological insulator surface states and a proposal for electronic Goos-Hänchen shift measurement

Hassan Ghadiri[1] and Alireza Saffarzadeh[2,3,*]

[1]*Department of Physics, North Tehran Branch, Islamic Azad University, 16511-53311, Tehran, Iran*
[2]*Department of Physics, Payame Noor University, P. O. Box 19395-3697 Tehran, Iran*
[3]*Department of Physics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6*

The hexagonal warping effect on transport properties and Goos-Hänchen (GH) lateral shift of electrons on the surface of a topological insulator with a potential barrier is investigated theoretically. Due to the warped Fermi surface for incident electron beams, we can expect two propagating transmitted beams corresponding to the occurrence of double refraction. The transmitted beams have spin orientations locked to their momenta so one of the spin directions rotates compared to the incident spin direction. Based on a low-energy Hamiltonian near the Dirac point and considering Gaussian beams, we derive expressions for calculating lateral shifts in the presence of warping effect. We study the dependence of transmission probabilities and GH shifts of transmitted beams on system parameters in detail by giving an explanation for the appearance of large peaks in the lateral shifts corresponding to their transmission peaks. It is shown that the separation between two transmitted beams through their different GH shifts can be as large as a few micrometers, which is large enough to be observed experimentally. Finally, we propose a method to measure the GH shift of electron beams based on the transverse magnetic focusing technique in which, by tuning an applied magnetic field, a detectable resonant path for electrons can be induced.

## I. INTRODUCTION

Topological insulators (TIs) are nonmagnetic insulators with conducting surface states as a consequence of the non-trivial topology of their bulk band structure, which in turn results from strong spin-orbit interaction. The surface states contain an odd number of spin-helical Dirac cones and are protected against any disturbance that maintains time-reversal symmetry [1–3]. In the vicinity of the Dirac point, the electron states can be well-described by massless Dirac equation, whereas at energies far enough away from the Dirac point, a distortion induced by surface spin-orbit coupling deforms the Fermi surface into a hexagonal snowflake shape [4–6]. This effect, called hexagonal warping, has been confirmed by angle-resolved photoemission spectroscopy [5]. Since the surface states close to the Fermi level play a decisive role in the electronic properties of two-dimensional (2D) materials, the hexagonal warping can affect transport properties on the surface of TIs. Therefore, the warping effect and also, topologically, protection of surface states may lead to a variety of interesting properties which are important from the viewpoint of fundamental physics as well as device applications [7–18].

It is well-known that the totally reflected light beam from a dielectric interface undergoes a lateral displacement from the position predicted by the geometrical optics. The study of this phenomenon, which is known as the Goos-Hänchen (GH) shift [19], has been developed to partial reflections and also transmitting configurations [20–23]. When an electron beam is incident on a boundary separating two regions of different densities, the reflected/transmitted beam undergoes a GH shift similar to a light beam crossing a boundary between materials with different optical indices. Accordingly, the GH shift of electrons in condensed matter systems [24–28], especially in Dirac materials [29–46], has been extensively studied.

The GH shift and transverse displacement, called the Imbert-Fedorov (IF) shift, of a light beam on the surface of some Dirac materials, such as graphene and Weyl semimetals, have also been investigated [47,48]. The results showed that the optical beam shifts provide a possible scheme for direct measurement of the parameters in these materials [47,48]. Moreover, it was shown that the electronic IF shift can be utilized to characterize Weyl semimetals [42]. On the other hand, the results of electronic beam shifts (EBSs) suggest a generation of nanoelectronic devices based on transition metal dichalcogenides [36–38] and Weyl semimetals [42,43]. Therefore, the study of EBSs on TIs along with the ability of measuring EBSs can potentially provide applications in characterizing the parameters of TIs as well as the fabrication of new TI-based nanodevices.

In this paper, we investigate the propagation of electrons through a square potential barrier on the surface of a TI by considering the hexagonal warping effect. We show that due to the warped Fermi surface, an electron wave impinging onto the barrier can have two transmitted waves, propagating with different momenta and hence in different directions, much like the double refraction of light in anisotropic crystals, demonstrating another opticslike property of electrons. We derive a formula for calculation of GH shifts of two transmitted

*asaffarz@sfu.ca

beams in the presence of hexagonal warping. We show that the beams can be separated spatially due to their different GH shifts, while at the same time they have different spin orientations due to their different momenta. However, due to the difficulty in producing well-collimated electron beams, the GH shift in electronic systems has not been measured so far [49].

We present a proposal for experimentally measuring the GH shift based on the transverse magnetic focusing (TMF) technique in which, by applying a transverse magnetic field, one can focus the motion of electrons/holes in the ballistic regime [50–52]. The TMF has been used to study the shape of Fermi surfaces [52], Andreev reflection [52,53], spin-orbit interaction [54], the angle-resolved transmission probability in graphene [55], and imaging electron trajectories [53,56–58], as well as proposing a method for measuring warping strength in TIs [16].

In this proposal, by applying a transverse magnetic field on the incident region, the impact point and also incident angle of electrons at the first interface are controlled, similar to the experiment of Ref. [55] and also the proposal in Ref. [16]. The variation of transverse magnetic field applied on the transmission region induces a resonant conduction path (measured as a voltage peak) by which the entry point of electrons at the second interface is determined, and hence the GH shift can be measured.

The paper is organized as follows. We introduce our model and formalism for obtaining transmission probability in Sec. II, where the scattering wave functions in the presence of warping effect and the transmission properties of incident electron waves are discussed in detail. In Sec. III, we calculate the GH shift of electron beams and the spatial beam separation is investigated. In Sec. IV, we present our proposal for measuring the electronic GH shift based on the TMF phenomenon. We conclude our findings in Sec. V.

## II. THEORETICAL MODEL AND DOUBLE REFRACTION

Surface states of TIs with a single Dirac cone are generally described by the Hamiltonian ($\hbar = 1$),

$$H = \upsilon_F \mathbf{k} \cdot (\hat{z} \times \boldsymbol{\sigma}) + \frac{\lambda}{2}(k_+^3 + k_-^3)\sigma_z, \quad (1)$$

where $\mathbf{k} = (k_x, k_y)$ is the wave vector of the electron, $\boldsymbol{\sigma}$ is the Pauli matrix vector, $k_\pm = k_x \pm ik_y$, and $\hat{z}$ is a unit vector normal to the surface. $\upsilon_F$ and $\lambda$ are the Fermi velocity and warping parameter, respectively. The linear term in $\mathbf{k}$ being similar to that of graphene, with the exception that $\boldsymbol{\sigma}$ represents the real spin of electron and cubic terms in the Hamiltonian, are responsible for the hexagonal warping effect. This Hamiltonian, which neglects the multiorbital structure of surface states, is considered as a minimal model preserving given $C_{3v}$ symmetries [4,18]. Among different TIs, Bi$_2$Te$_3$ is found to have strong warping effect with Fermi velocity $\upsilon_F = 2.55$ eV Å and the hexagonal warping parameter $\lambda = 250$ eV Å$^3$ that we consider here in the calculations [4,10]. The eigenvalues of Eq. (1) give the upper and lower bands in $\mathbf{k}$ space as

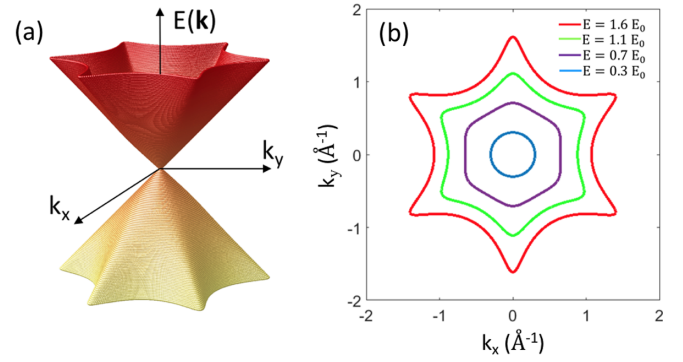$$E_\pm(\mathbf{k}) = \pm\sqrt{(\upsilon_F k)^2 + w^2(\mathbf{k})}, \quad (2)$$



FIG. 1. (a) The Dirac cone of fermions with Fermi energy $E_F = 1.6E_0$ on the surface of Bi$_2$Te$_3$ by including the hexagonal warping effect. (b) The constant-energy contours of Dirac cone at different energies measured from the Dirac point.

where $w(\mathbf{k}) = \lambda k_x(k_x^2 - 3k_y^2)$. We have depicted the Dirac cone of fermions [Eq. (2)] and several constant energy contours (CECs) in Figs. 1(a) and 1(b), respectively. The energy of the CEC is expressed in terms of $E_0 = \sqrt{\upsilon_F^3/\lambda}$, which is the characteristic energy introduced by hexagonal warping. At $E \ll E_0$, the warping effect is negligible and the CEC exhibits a circular shape. As the energy exceeds the critical value $E_c = \frac{\sqrt{7}}{6^{\frac{3}{4}}}E_0 \approx 0.69E_0$, the CEC deforms into a hexagonal shape, with inflection points satisfying the relation $(\frac{\partial k_x}{\partial k_y})_E = (\frac{\partial^2 k_x}{\partial^2 k_y})_E = 0$. With further increasing $E$, the rounded tips of the hexagon become sharper and the CEC exhibits a snowflake shape. The eigenspinors of Hamiltonian Eq. (1) can be written as

$$\chi(\mathbf{k}, E_\pm)e^{i\mathbf{k}\cdot\mathbf{r}} = \frac{1}{N_\pm(\mathbf{k})}\begin{pmatrix} \pm(E_\pm(\mathbf{k}) + w(\mathbf{k})) \\ \upsilon_F(ik_x - k_y) \end{pmatrix}e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (3)$$

where $N_\pm(\mathbf{k})$ are the normalization coefficients and the subscript $+(-)$ corresponds to the upper (lower) band in Eq. (2). Note that the interaction between surface states and bulk states can be ignored since the surface Dirac point on TIs such as Bi$_2$Te$_3$ is closer to the bulk valence band than to the bulk conduction band [5].

Now, we consider the propagation of electrons on the surface of TI scattered by a potential barrier $V(y) = V_0$ for $0 \leqslant y \leqslant d$ and $V(y) = 0$ elsewhere (see Fig. 2). Such a barrier can be produced by a gate electrode deposited on top of the TI surface. We note that the electron transport in the $y$ direction is coherent. Also, due to the translational invariance in the $x$ direction, $k_x$ is a good quantum number, and hence the Fermi energy $E_F$ of the electron is conserved in the scattering process. In zero potential regions for given $E_F$ and $k_x$, equation $E_+(\mathbf{k}) = E_F$ is quartic in terms of $k_y$ and its roots determine the $y$ components of electron Fermi momentum. It gives two real symmetric roots and two imaginary symmetric roots in the case of $E_F < E_c$, indicating that an incoming electron wave with an arbitrary incident angle $\theta$ has one propagating reflected wave and one propagating transmitted wave [see Fig. 2(a)] as it is a normal case for most conventional materials. For $E_F > E_c$, the CEC has concave segments and consequently, as shown in Fig. 2(b), there exists a critical incident angle $\theta_c$ beyond which the equation $E_+(\mathbf{k}) = E_F$ has
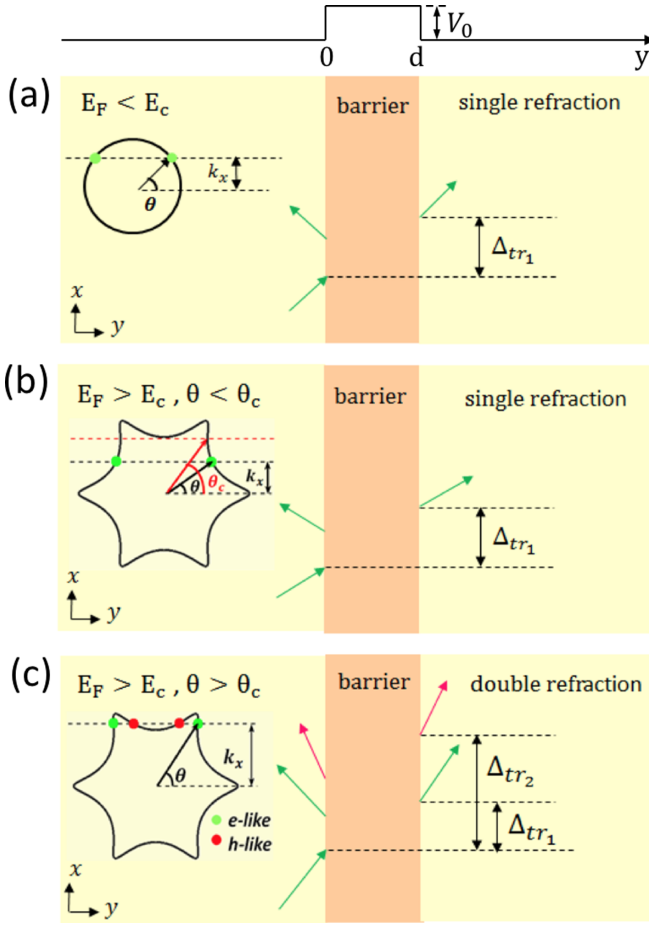
FIG. 2. Scattering processes when an incident electron is reflected and transmitted from the barrier with width $d$ and height $V_0$ as shown on top of (a) in the cases of (a), (b) single and (c) double refractions. The insets show the CEC at (a) $E_F < E_c$, (b) $E_F > E_c$, $\theta < \theta_c$, and (c) $E_F > E_c$, $\theta > \theta_c$. The green (red) circles on CECs indicate propagating electronlike (holelike) states. Also, the GH shift $\Delta_{tr_1}$ of beam 1 with the same momentum as that of the incident beam and GH shift $\Delta_{tr_2}$ of beam 2 with different momentums compared to the incident beam are shown.

four symmetrical real roots. This means that for an incoming electron wave, there exist two propagating reflected waves (double reflection) and simultaneously two propagating transmitted waves (double refraction) [see Fig. 2(c)]. When $\theta < \theta_c$, similar to the case of $E_F < E_c$, single refraction happens as can be seen in Fig. 2(b). We should note that at $\theta > \theta_c$, the Fermi momenta with bigger absolute values along the $y$ axis are parallel to their corresponding group velocities defined by $\upsilon_y = (\frac{\partial E}{\partial k_y})_{k_x}$, therefore the corresponding states are electronlike. In contrast, the Fermi momenta with smaller absolute values are antiparallel to their corresponding group velocities, indicating holelike propagating states. In the barrier region, the roots of $k_y$ are obtained from the equation $E_+(\mathbf{k}) + V_0 = E_F$ depending on the amounts of $E_F, V_0$, and incident angle $\theta$. They can be real, complex, or two real roots and two imaginary roots.

To obtain transmission probability and for future purposes, we write the generic scattering states for given $E_F, V_0$, and

$\theta = \arctan(|\frac{k_x}{k_{y,1}}|)$ as

$$\psi(y) = \begin{cases} \chi(k_{y,1}, E_F)e^{ik_{y,1}y} + r_1\chi(-k_{y,1}, E_F)e^{-ik_{y,1}y} \\ \quad + r_2\chi(k_{y,2}, E_F)e^{ik_{y,2}y}, & y \leqslant 0 \\ \sum_{n=1}^{4} a_n\chi(k'_{y,n}, E_F - V_0)e^{ik'_{y,n}y}, & 0 \leqslant y \leqslant d \\ t_1\chi(k_{y,1}, E_F)e^{ik_{y,1}(y-d)} \\ \quad + t_2\chi(-k_{y,2}, E_F)e^{-ik_{y,2}(y-d)}, & y \geqslant d, \end{cases}$$

(4)

where $\chi(k_{y,1}, E_F)$ is the incident state with Fermi momentum $k_{y,1}$, $r_1(r_2)$ is the reflection amplitude corresponding to the Fermi momentum $-k_{y,1}(k_{y,2})$, and $t_1(t_2)$ is the transmission amplitude corresponding to the Fermi momentum $k_{y,1}(-k_{y,2})$, while $a_n$ is the scattering amplitude corresponding to the momentum $k'_{y,n}$ in the barrier region. As can be seen in Fig. 2, at $E_F > E_c$ and $\theta_c < \theta < \frac{\pi}{3}$ ($\theta > \frac{\pi}{3}$), $k_{y,1}$ is a positive electronlike (negative holelike) momentum and $k_{y,2}$ is a positive holelike (negative electronlike) momentum, while for $\theta < \theta_c$, $k_{y,1}$ is a positive real root and $k_{y,2}$ is a negative imaginary root. In the case of $E_F < E_c$, however, for every incident angle $\theta$, $k_{y,1}$ is a positive real root and $k_{y,2}$ is a negative imaginary root.

The eigenvalue equation $(H + V(y))\psi(y) = E\psi(y)$ corresponding to the Hamiltonian Eq. (1) is a second-order partial differential equation with respect to $y$ due to the warping effect. Therefore, by applying boundary conditions of continuity of $\psi(y)$ and its first derivative with respect to $y$ at the two interfaces $y = 0$ and $y = d$, the reflection and transmission amplitudes, and also the scattering amplitude $a_n$ can be determined. The transmission probability which is defined as the ratio of the $y$ component of the probability current density of the transmitted waves and that of the incident wave can be expressed in terms of the transmission amplitudes and the $y$ component of the corresponding group velocities as $T = T_1 + T_2 = |t_1|^2 + \frac{\upsilon_y(-k_{y,2}, E_F)}{\upsilon_y(k_{y,1}, E_F)}|t_2|^2$ in the case of double refraction and $T = T_1 = |t_1|^2$ in the case of single refraction [59,60]. In these relations, $T$ denotes the total transmission probability, while $T_1$ ($T_2$) represents the transmission probability of the transmitted wave with the same (different) momentum as (from) the incident wave momentum.

We have shown in Fig. 3 the contour plot of the transmission probability as a function of incident angle and the potential barrier height for two different Fermi energies $E_F < E_c$ and $E_F > E_c$. At $E_F < E_c$, there is only one propagating transmitted wave whose probability at the typical Fermi energy of 0.15 eV and the barrier width $d = 500$ Å is shown in Fig. 3(a). As can be seen, there is a region with two boundaries inside where the total internal reflection (TIR) takes place. The boundaries represent the geometrical locations of a critical angle $\theta_{TIR} = \theta_{TIR}(V_0)$ such that when the incident angle reaches $\theta_{TIR}$, all four waves inside the barrier region become evanescent. Therefore, at a sufficiently wide barrier, TIR begins. To obtain an expression for $\theta_{TIR}$, first we consider equation $E_+(\mathbf{k}) + V_0 = E_F$ which is quadratic in terms of $k_y^2$. By solving the discriminant of this equation, the $k_x$ value corresponding to $\theta_{TIR}$ can be obtained. Replacing $k_x$ in equation $E_+(\mathbf{k}) = E_F$, we obtain the corresponding $k_y$ of
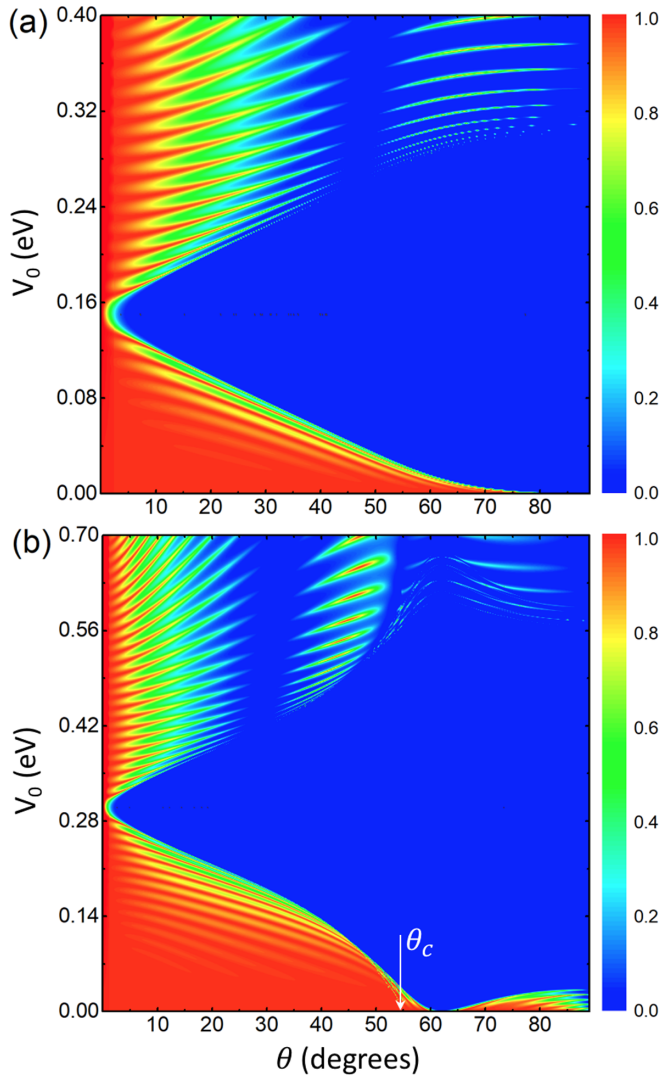
FIG. 3. Calculated transmission probability $T_1$ versus $\theta$ and $V_0$ for incident electrons on a potential barrier of width $d = 500$ Å at Fermi energy (a) $E_F = 0.15$ eV $< E_c$ and (b) $E_F = 0.3$ eV $> E_c$. ($E_c \simeq 0.18$ eV).
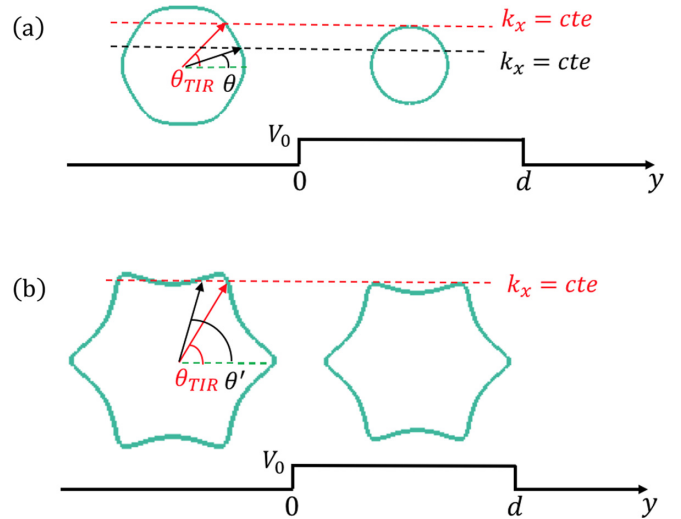


FIG. 4. (a) Solid curves represent CECs in the incident and barrier regions for $E_F = 0.15$ eV and $|V_0 - E_F| = 0.1$ eV. The black (red) dashed line shows the conservation of transverse momentum $k_x$ in the scattering process at incident angle $\theta(\theta_{TIR})$. (b) The CEC in the incident region corresponds to $E_F = 0.3$ eV. The value of $V_0$ is chosen such that the maximum of $k_x$ on the CEC of the barrier region is bigger than the $k_x$ value on CEC of the incident region at $\theta = \pi/2$, i.e., $\theta_{TIR} > \theta_c$. Due to the warped CEC in the incident region, TIR is confined in the interval of $\theta_{TIR} < \theta < \theta'$.

CEC of the barrier region, meaning that all $k'_{y,n}$ are complex and, consequently, TIR begins. With increasing $V_0$, the size of CEC in the barrier region decreases and hence the critical angle becomes smaller. As $V_0$ approaches $E_F$, the critical angle at which all $k'_{y,n}$ become complex approaches zero. However, due to the Klein tunneling effect (see discussion below) which prohibits backscattering near the normal incident angle, TIR cannot start at $\theta = 0°$ [see Fig. 3(a)] [61]. When $V_0$ exceeds $E_F$, the size of CEC in the barrier region increases as well as the critical angle, until $|V_0 - E_F|$ reaches $E_F$. From now on, since the size of CEC in the barrier region becomes larger than that of the CEC in the incident region, TIR does not form at any incident angle.

As can be seen in Fig. 3(a), the barrier remains perfectly transparent at incident angles close to the normal incidence $\theta = 0°$, regardless of the amount of $V_0$. This process is known as Klein tunneling [62] and originates from spin conservation [63], since nonmagnetic barriers cannot change the spin direction of incident electrons in a scattering process. On the other hand, at almost normal incidence, the spin states of incident wave and propagating reflected wave are orthogonal [this can be easily deduced from Eq. (3)]. Therefore, backscattering is forbidden and electrons can transmit perfectly. By increasing $\theta$, the spin states of incident and propagating reflected waves are no longer orthogonal and hence the electron reflection is allowed. Now we consider the case of oblique incidence. When a wave impinges on the barrier at a given $\theta \neq 0$, a part of wave transmits into the barrier and is multiply reflected at the two interfaces $y = 0$ and $y = d$, therefore interference happens. As $V_0$ increases from zero, the size of CEC in the barrier region changes, and hence, the acquired phase $k'_y d$ of propagating waves inside the barrier region will change,

$\theta_{TIR}$, resulting in $\theta_{TIR} = \arctan(k_x/k_y)$. The obtained analytical expression for $\theta_{TIR}$ is complicated. Therefore, we have not presented the resulting expression here. Instead, we explain below how $\theta_{TIR}$ occurs and behaves as a function of $V_0$. At $E_F = 0.15$ eV and the typical value $V_0 = 0.05$ eV, the CECs in the incident and barrier regions are shown in Fig. 4(a). The size of the CEC in the barrier region (determined by $|V_0 - E_F|$) is smaller than the size of the CEC in the incident region. The line $k_x = $ cte, corresponding to the incident angle $\theta$, represents the conservation of $k_x$ in the electron scattering process. As shown in Fig. 4(a), at a small incident angle $\theta$, the line $k_x = $ cte intersects the CEC of the barrier region at two points, representing two real wave numbers of propagating electrons inside the barrier region. With increasing $\theta$, the corresponding line $k_x = $ cte moves upward until it touches the CEC of the barrier region at a single point, indicating that the incident angle $\theta$ reaches the critical angle $\theta_{TIR}$. When $\theta$ exceeds $\theta_{TIR}$, the line of $k_x = $ cte can no longer intersect the

causing oscillations in transmission probability. Whenever the waves interfere constructively, the Fabry-Pérot resonances with $T(\theta \neq 0) = 1$ appears. As $V_0$ reaches a value at which the line $k_x = cte$ becomes tangent to the CEC of barrier region [see Fig. 4(a)], the TIR begins. It continues until the line $k_x = cte$ again becomes tangent to the CEC, but this time at an amount of $V_0 > E_F$. With further increasing $V_0$, oscillations appear again in the transmission spectrum. Moreover, similar to the Schrödinger-type electrons, when the potential height is less than Fermi energy (corresponds to an $n$-$n'$-$n$ junction), the incident electrons generally pass through the potential barrier with larger transmission probabilities, compared to the case of $V_0 > E_F$ (corresponds to an $n$-$p$-$n$ junction). Also, we should mention that for $\theta > 80°$ the group velocity of electrons $v_y = \left(\frac{\partial E}{\partial k_y}\right)_{k_x} = -v_x\left(\frac{\partial k_x}{\partial k_y}\right)_E \simeq 0$ [see CEC in incident region in Fig. 4(a)]. Therefore, the electron transmission becomes very small at $V_0 = 0$. However, at some $V_0$ values, $T$ can be considerably magnified due to the interference effect. It is important to point out that for validity of the minimal continuum model described by Hamiltonian Eq. (1), both $E_F$ and $|V_0 - E_F|$ must be considered less than 0.4 eV [5,10]. However, in Fig. 3(a), we terminate $V_0$ at 0.4 eV as at larger values no more features can be observed.

The CEC of zero-potential regions is warped at $E_F > E_c$, and hence two transmitted waves can propagate. The transmission probability $T_1$ at $E_F = 0.3$ eV and with the same barrier width value as that in Fig. 3(a) is depicted in Fig. 3(b). First, we consider a given $V_0$ at which the size of the CEC in the barrier region is small enough compared to the CEC of the incident region. Similar to Fig. 4(a), as $\theta$ increases from zero, the corresponding line $k_x = cte$ moves upward and, consequently, the acquired phase $k_y'd$ of propagating electron waves inside the barrier region will change, resulting in oscillations in $T_1$ and also Fabry-Pérot resonances in the constructive interference. When $\theta$ reaches $\theta_{TIR}$, the line $k_x = cte$ touches the CEC of the barrier region at a single point, that is, the start of TIR. The TIR extends to $\pi/2$ because for $\theta > \theta_{TIR}$ the constant $k_x$ line can no longer intersect the CEC of the barrier region. Now, we consider an amount of $V_0$ at which the size of CEC in the barrier region is close enough to the CEC of incident region, so the maximum of $k_x$ in the barrier region is larger than $k_x$ at $\theta = \pi/2$ in the incident region [see Fig. 4(b)]. In this case, when $\theta$ exceeds $\theta_{TIR}$, the corresponding constant $k_x$ line will start to intersect the CEC of the barrier region one more time at an angle $\theta' > \pi/3$. Therefore, the TIR terminates at $\theta'$ and will not extend to $\pi/2$. This happens at $V_0 < 0.03$ eV and 0.57 eV $< V_0 < 0.6$ eV in Fig. 3(b). At almost normal incident angle, the Klein tunneling happens, regardless of $V_0$ value, similar to Fig. 3(a). However, at a given oblique incident angle $\theta \neq 0$ when $V_0$ varies outside the TIR region, the size of CEC in the barrier region changes and hence the acquired phase $k_y'd$ of the propagating waves inside the barrier region may also change, causing an oscillatory behavior in $T_1$, similar to the behavior of Fig. 3(a). Since $(\partial k_x / \partial k_y)_E$ approaches zero as $\theta$ reaches 62° [see the CEC in the incident region in Fig. 4(b)], $v_y$ and, consequently $T_1$ are very small, independent of $V_0$ values. On the other hand, the transmission probability $T_2$ (not shown here) is zero at $\theta < \theta_c = 54.7°$, while it has the same features as $T_1$ at $\theta > \theta_c$, regardless of $V_0$ values.
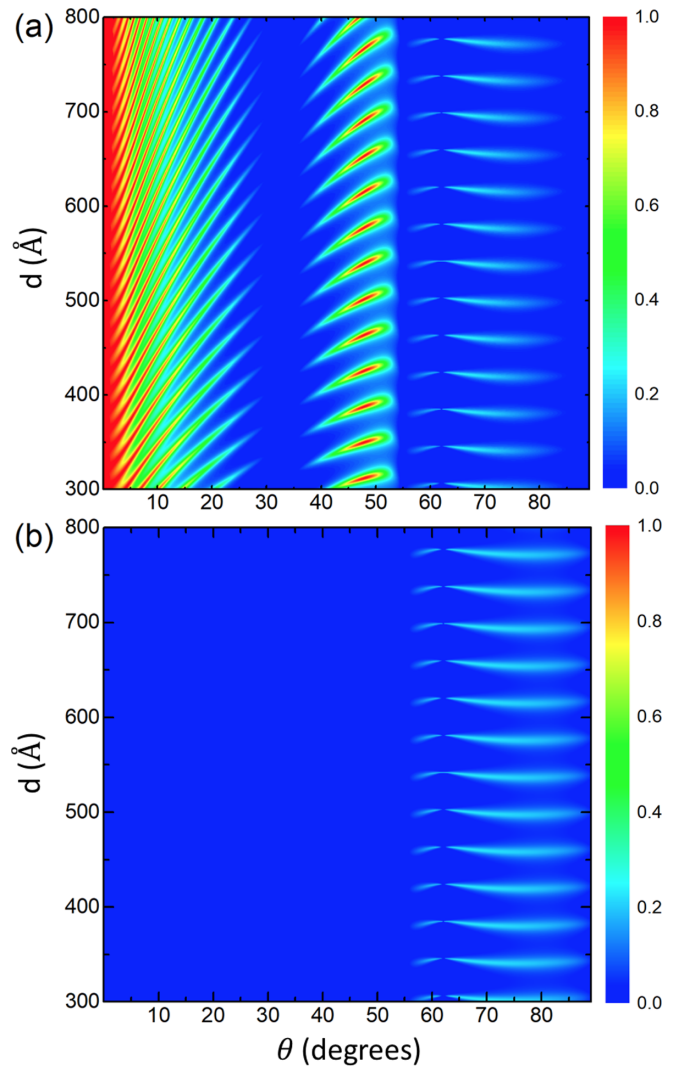


FIG. 5. Calculated transmission probabilities (a) $T_1$ and (b) $T_2$ versus $\theta$ and $d$ for incident electrons with Fermi energy $E_F = 0.3$ eV on a potential barrier of $V_0 = 0.7$ eV.

The transmission probabilities $T_1$ and $T_2$ for the two transmitted waves 1 and 2 as functions of incident angle $\theta$ and the barrier width $d$, at $E_F = 0.3$ eV and $V_0 = 0.7$ eV are shown in Figs. 5(a) and 5(b), respectively. Since $|V_0 - E_F| > E_F$, the TIR does not occur and for a given $d$ value at almost normal incident angle, the Klein tunneling with perfect transmission probability for $T_1$ happens, as shown in Fig. 5(a). As $\theta$ is increased from zero, the acquired phase $k_y'd$ of the propagating waves inside the barrier region changes. This causes oscillations in $T_1$, emerging Fabry-Pérot resonances when constructive interference takes place. If $d$ varies at a fixed oblique incident angle, the acquired phase $k_y'd$ will change so $T_1$ again exhibits oscillations and Fabry-Pérot resonances can emerge. Here, $\theta_c = 54.7°$ is the same as that in Fig. 3(b) because $\theta_c$ depends only on $E_F$. When $\theta$ exceeds $\theta_c$, $T_2$ in Fig. 5(b) takes nonzero values and shows oscillations due to the change of acquired phase $k_y'd$, with varying $\theta$ or $d$. In the vicinity of $\theta = 62°$, both $T_1$ and $T_2$ are very small for the same reason explained in Fig. 3(b).

## III. GOOS-HÄNCHEN SHIFT AND BEAM SPLITTING

The GH shift for a plane wave of electrons is not detectable due to its infinite spatial width. Therefore, to calculate the GH lateral shift, we consider a beam of electrons instead of a plane wave. We model an incident electron beam by using a Gaussian wave packet of surface states as

$$\psi_{in}(\mathbf{r}) = \int_{-\infty}^{+\infty} dk_x f(k_x - k_{x_0}) \chi(k_{y,1}(k_x), E_F) e^{i(k_x x + k_{y,1}(k_x) y)},$$

(5)

where $f(k_x - k_{x_0}) = (\sqrt{2\pi}\Delta_{k_x})^{-1} e^{-(k_x - k_{x_0})^2/2\Delta_{k_x}^2}$ shows the Gaussian angular distribution of width $\Delta_{k_x}$ around central incident angle $\theta_0 = \arctan(|\frac{k_{x_0}}{k_{y,1}}|)$. Analogously, the wave functions of transmitted electron beams can be written as

$$\psi_{tr_1}(\mathbf{r}) = \int_{-\infty}^{+\infty} dk_x f(k_x - k_{x_0}) t_1(k_x)$$
$$\times \chi(k_{y,1}(k_x), E_F) e^{i(k_x x + k_{y,1}(k_x)(y-d))}$$

(6)

and

$$\psi_{tr_2}(\mathbf{r}) = \int_{-\infty}^{+\infty} dk_x f(k_x - k_{x_0}) t_2(k_x)$$
$$\times \chi(-k_{y,2}(k_x), E_F) e^{i(k_x x - k_{y,2}(k_x)(y-d))}.$$

(7)

For the well-collimated electron beams, $f(k_x - k_{x_0})$ is sharp around $k_{x_0}$ such that the spinor components $\chi^\pm = |\chi^\pm| \times e^{i\varphi^\pm}$ can be converted into an exponential form and approximated by keeping the first two terms of the Taylor expansion of its exponent around $k_{x_0}$ as

$$\chi^\pm(k_{y,1}(k_x)) = \exp[\ln \chi^\pm(k_{y,1}(k_x))]$$
$$\simeq \chi^\pm(k_{y,1}(k_{x_0})) \exp\{[\frac{|\dot{\chi}^\pm(k_{y,1}(k_{x_0}))|}{|\chi^\pm(k_{y,1}(k_{x_0}))|}$$
$$+ i\dot{\varphi}^\pm(k_{y,1}(k_{x_0}))](k_x - k_{x_0})\},$$

(8)

where $\dot{\varphi}^\pm(k_{y,1}(k_{x_0}))(|\dot{\chi}^\pm(k_{y,1}(k_{x_0}))|)$ denotes the derivative of $\varphi^\pm(k_{y,1}(k_x))(|\chi^\pm(k_{y,1}(k_x))|)$ with respect to $k_x$, evaluated at $k_x = k_{x_0}$. Substituting Eq. (8) into Eq. (5), using the approximation $k_{y,1}(k_x) \simeq k_{y,1}(k_{x_0}) + \dot{k}_{y,1}(k_{x_0})(k_x - k_{x_0})$ and then evaluating the integral, we obtain the spatial form of the components of the incident beam as

$$\psi_{in}^\pm(\mathbf{r}) = \chi^\pm(k_{y,1}(k_{x_0}))$$
$$\times e^{-[x + \dot{\varphi}^\pm(k_{y,1}(k_{x_0})) + \dot{k}_{y,1}(k_{x_0})y]^2 \Delta_{k_x}^2/2}$$
$$\times e^{\gamma^{\pm 2}/2\Delta_{k_x}^2} e^{i\gamma^\pm \dot{\varphi}^\pm(k_{y,1}(k_{x_0}))}$$
$$\times e^{i[(k_{y,1}(k_{x_0}) + \gamma^\pm \dot{k}_{y,1}(k_{x_0}))y + (k_{x_0} + \gamma^\pm)x]},$$

(9)

where $\gamma^\pm = \frac{\Delta_{k_x}^2 |\dot{\chi}^\pm(k_{y,1}(k_{x_0}))|}{|\chi^\pm(k_{y,1}(k_{x_0}))|}$. As can be seen from the second factor in Eq. (9), the incident beam has a Gaussian shape and the peak location of its upper and lower components at the interface $y = 0$ is given by $x_{in}^\pm = -\dot{\varphi}^\pm(k_{y,1}(k_{x_0}))$. Therefore, the average location of incident beam at the interface $y = 0$ can be expressed as

$$\bar{x}_{in} = -\dot{\varphi}^+(k_{y,1}(k_{x_0}))|\chi^+(k_{y,1}(k_{x_0}))|^2$$
$$- \dot{\varphi}^-(k_{y,1}(k_{x_0}))|\chi^-(k_{y,1}(k_{x_0}))|^2.$$

(10)

It is worth mentioning that the last factor in Eq. (9) shows that the propagation direction of incident-beam components deviates from the central angle $\theta_0$ by the amount of $\delta^\pm \approx \tan \delta^\pm \approx \frac{\gamma^\pm}{k_{y,1}(k_{x_0})}$. This deflection is due to the warping effect, as in the absence of warping, $|\chi^\pm|$ is constant and $\delta^\pm = 0$. Moreover, the third factor in Eq. (9) reveals that the magnitude of the incident beam is adjusted by warping as well.

By comparing Eqs. (6) and (7) with Eq. (5), we can write an expression for the transmitted beam components, similar to Eq. (9), by the substitutions $\chi^\pm \mapsto \chi^\pm t_{1(2)}$, $\varphi^\pm \mapsto \varphi^\pm + \varphi_{t_{1(2)}}$, and $|\chi^\pm| \mapsto |\chi^\pm||t_{1(2)}|$ in Eq. (9) where $\varphi_{t_{1(2)}}$ represents the phase of transmission amplitude $t_{1(2)}$. Therefore, the transmitted beams find also Gaussian shapes just like the incident beam. The average locations of the transmitted beams at the interface $y = d$ read as

$$\bar{x}_{tr_1} = -\dot{\varphi}^+(k_{y,1}(k_{x_0}))|\chi^+(k_{y,1}(k_{x_0}))|^2 - \dot{\varphi}_{t_1}(k_{x_0})$$
$$- \dot{\varphi}^-(k_{y,1}(k_{x_0}))|\chi^-(k_{y,1}(k_{x_0}))|^2$$

(11)

and

$$\bar{x}_{tr_2} = -\dot{\varphi}^+(-k_{y,2}(k_{x_0}))|\chi^+(-k_{y,2}(k_{x_0}))|^2 - \dot{\varphi}_{t_2}(k_{x_0})$$
$$- \dot{\varphi}^-(-k_{y,2}(k_{x_0}))|\chi^-(-k_{y,2}(k_{x_0}))|^2.$$

(12)

The GH lateral shift $\Delta_{tr}$ is defined as the displacement of the peak of transmitted beam at the interface $y = d$ relative to the peak of incident beam at the interface $y = 0$ [23] (see Fig. 2) which is different from classical prediction of electron optics, i.e., Snell's shift $d \tan \theta'$, where $\theta'$ is the refraction angle. Therefore, the GH shift of the transmitted beam 1(2) with the same (different) momentum as (from) that of the incident beam can be obtained from Eqs. (10)–(12) as

$$\Delta_{tr_1} = -\dot{\varphi}_{t_1}(k_{x_0})$$

(13)

and

$$\Delta_{tr_2} = -\dot{\varphi}_{t_2}(k_{x_0}) - \dot{\varphi}^-(-k_{y,2}(k_{x_0}))|\chi^-(-k_{y,2}(k_{x_0}))|^2$$
$$+ \dot{\varphi}^-(k_{y,1}(k_{x_0}))|\chi^-(k_{y,1}(k_{x_0}))|^2.$$

(14)

In deriving Eq. (14), we have used $\dot{\varphi}^+(k_{y,1}(k_{x_0})) = \dot{\varphi}^+(-k_{y,2}(k_{x_0})) = 0$ because in the case of double refraction, upper spinor components in zero potential regions are real. The spatial splitting between the two beams will occur when they have different GH shifts. In this case, the spatial separation between the two beams is given by $\delta\Delta_{tr} = \Delta_{tr_1} - \Delta_{tr_2}$.

The electron spin orientation can be obtained from Eq. (3) as $\mathbf{s} = \langle\sigma\rangle = E_+^{-1}(-v_F k_y, v_F k_x, w(\mathbf{k}))$, indicating the spin-momentum locking of surface electrons in TIs, due to the spin-orbit coupling. Consequently, the spin direction of transmitted beam 2 is rotated relative to the spin direction of both transmitted beam 1 and incident beam by the amount of $\alpha = \arccos(\mathbf{s}_1 \cdot \mathbf{s}_2)$, where $\mathbf{s}_1$ and $\mathbf{s}_2$ are spin orientations of transmitted beams 1 and 2, respectively.

Due to the warping term, $\Delta_{tr_{1(2)}}$ cannot be derived in compact analytical expressions. Therefore, these quantities are calculated numerically using Eqs. (13) and (14). Typical results for GH shifts and the corresponding transmission probabilities are shown in Figs. 6 and 7. The parameters are chosen to avoid TIR and that two transmitted beams propagate.
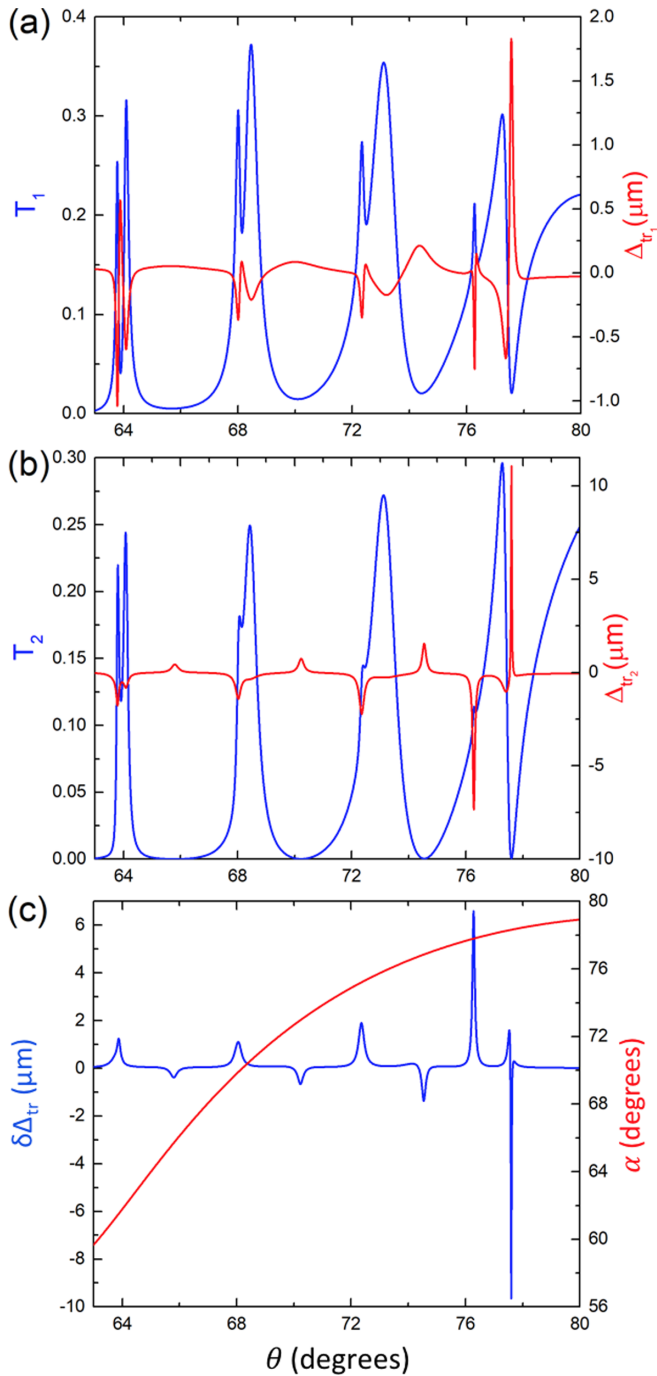
FIG. 6. (a), (b) Calculated transmission probabilities $T_{1,2}$ (blue curves) and the GH shifts $\Delta_{tr_{1,2}}$ (red curves) in transmission as a function of $\theta$. (c) The spatial separation $\delta\Delta_{tr}$ (blue curve) and the angle difference $\alpha$ (red curve) between spin orientation of transmitted beams. The parameters are $E_F = 0.35$ eV, $V_0 = 0.72$ eV, and $d = 490$ Å.
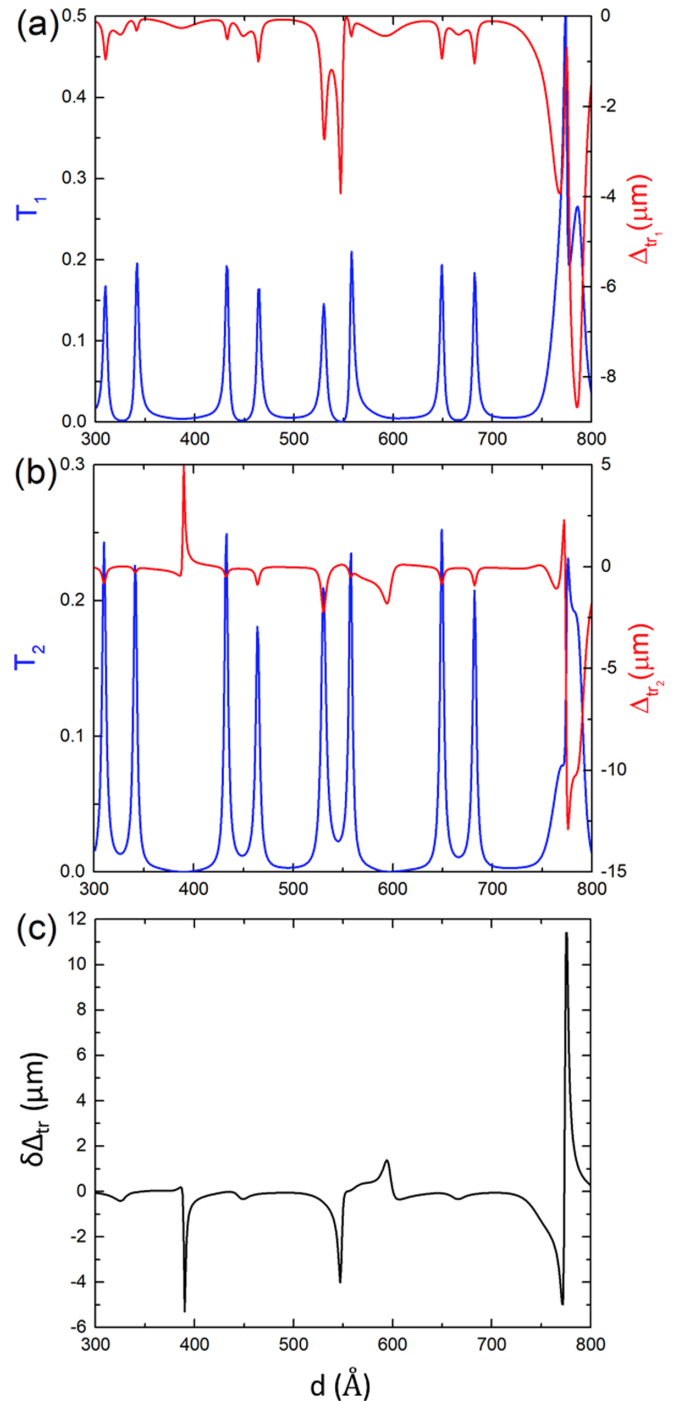
FIG. 7. (a), (b) Calculated transmission probabilities $T_{1,2}$ (blue curves) and the GH shifts $\Delta_{tr_{1,2}}$ (red curves) in transmission as a function of barrier width $d$. (c) The spatial separation $\delta\Delta_{tr}$ between transmitted beams. The parameters are $E_F = 0.3$ eV, $V_0 = 0.65$ eV, and $\theta = 59°$.

Figures 6(a) and 6(b) show the transmission probabilities and the corresponding GH values of the two transmitted beams 1 and 2 in terms of incident angle. Due to the interference effect, $T_{1,2}$ show an oscillatory behavior and some sharp maxima and minima appear for both transmitted beams. In fact, by changing the incident angle, the acquired phase $(k_y' d)$ of every propagating wave along the barrier region varies, which

leads to the oscillation of transmission probabilities. The peak positions of the two beams are almost the same. The corresponding GH shifts (red lines) of the two beams exhibit some strong peaks beside the usual ones with positive and negative values. To explain qualitatively the behavior of lateral shifts and the occurrence of their peaks, we rewrite the formula (13)

and (14) [64] as

$$\Delta_{tr} = \frac{\frac{d}{dk_x}\tan\varphi_t}{1+\tan^2\varphi_t}, \qquad (15)$$

where $\tan\varphi_t = \mathrm{Im}[t(k_x)]/\mathrm{Re}[t(k_x)]$. We approximate $\mathrm{Re}[t(k_x)]$ and $\mathrm{Im}[t(k_x)]$ around a given point $k_{x_0}$ by retaining the first and second terms of their Taylor expansion as $\mathrm{Re}[t(k_x)] \simeq a_R + b_R(k_x - k_{x_0})$ and $\mathrm{Im}[t(k_x)] \simeq a_I + b_I(k_x - k_{x_0})$, where $a_R, b_R, a_I$ and $b_I$ are coefficients of the expansions. By inserting these approximations in Eq. (15), we obtain

$$\Delta_{tr} = \frac{a_R b_I - a_I b_R}{|t(k_x)|^2}. \qquad (16)$$

Moreover, by approximating $\frac{d}{dk_x}|t(k_x)| \simeq \frac{a_R b_R + a_I b_I}{|t(k_x)|}$, Eq. (16) can be written as

$$\Delta_{tr} = \frac{a_R b_I - a_I b_R}{(a_R b_R + a_I b_I)^2}\left(\frac{d}{dk_x}|t(k_x)|\right)^2. \qquad (17)$$

From Eqs. (16) and (17), the local properties of $\Delta_{\mathrm{tr}}$ in the vicinity of a given point $k_{x_0}$, and hence $\theta_0$ can be studied. According to Eq. (17), the absolute value of $\Delta_{\mathrm{tr}}$ at any point depends on the absolute value of the slope of transmission probability at that point. By approaching the sharp maxima and minima points in the blue curves, the slope of $T_{1,2}$ rapidly finds very large values. Therefore, the absolute values of the corresponding GH shifts near these points in the red curves suddenly increase, creating sharp maxima and minima. The sign of the GH shift is determined by the sign of the numerator in Eq. (17). Some deep minima (not exactly zero) for $T_1$ and $T_2$ appear in Figs. 6(a) and 6(b), especially for $T_2$, i.e., $|t(k_x)| \approx 0$. According to Eq. (16), the absolute value of the corresponding lateral shifts at these points can become large and, therefore, local maxima appear at these points, as seen in red curves. Spatial separation between the two beams and the angle between their spin orientations as a function of incident angle are shown in Fig. 6(c) with blue and red curves, respectively. One can see that at the given angle window, $\delta\Delta_{\mathrm{tr}}$ exhibits several pronounced positive peaks, which make the observation of well-separated beams practically more feasible. Note that although at these points the transmission probabilities are far from the perfect splitter case with $T_1 = T_2 = 0.5$, these values are practically considerable. As an example, for the incident angle $\theta = 76.3°$ at which the obtained transmission probabilities for the two beams are 0.2 and 0.11 [see Figs. 6(a) and 6(b)], the spatial separation is about $6.5\mu m$, which is large enough to measure experimentally. At this point, the angle between spin orientations of the two beams is $77.8°$.

The transmission probabilities of the two transmitted beams and the corresponding GH shifts in terms of barrier width $d$ are depicted in Figs. 7(a) and 7(b). By varying the width of the barrier, the acquired phase, $k_y' d$, changes, and hence, $T_1$ and $T_2$ oscillate by revealing several sharp maxima and deep minima. We note that $k_y'$ does not change at a fixed incident angle. The behavior of GH shifts in Figs. 7(a) and 7(b), compared to their corresponding transmission probabilities, are similar to the behavior of GH shifts in Fig. 6(a) and 6(b), compared to their corresponding transmissions. That

means, near the sharp maxima of transmission probabilities where their slope rapidly increases with $d$, the absolute of the corresponding GH shifts increases as well. Also, near the deep minima of transmission probabilities where $|t(d)| \approx 0$, the absolute of the corresponding lateral shifts can be large. Such a similarity is explained as follows: The dependence of transmission coefficients, and hence their phases on $d$, is through the exponential functions $e^{ik_{y,n}'d}$, when the boundary conditions of continuity of the wave function in Eq. (4) and its derivative at the interface $y = d$ are applied. Consequently, the dependence of GH shifts on $d$ should be through exponential functions $e^{ik_{y,n}'d}$ as well. On the other hand, the dependence of transmission coefficients as well as GH shifts on $k_x(\theta)$ comes from $e^{ik_{y,n}'d}$ and also from other terms which vary slowly compared to the exponential functions. Therefore, when $d$ changes, the dependence of GH shift on the transmission probability will be similar to the dependence of GH shift on the transmission probability when $k_x$ or $\theta$ changes.

Spatial separation between the two electron beams as a function of $d$ is shown in Fig. 7(c). One can see that at $d \sim 775$ Å, the separation between the two beams is almost $10\mu m$ and the transmission probabilities of $T_1$ and $T_2$ are $\sim 0.5$ and 0.2, respectively [see Figs. 7(a) and 7(b)]. Also, the absolute value of $\delta\Delta_{\mathrm{tr}}$ peaks between 720 Å and 800 Å exhibits a considerable width, similar to $T_1$ and $T_2$. Moreover, the angle between spin orientations of two beams is $59.1°$, which is independent of $d$ and can be obtained by the values of $E_F$ and $\theta$. Therefore, our findings reveal that TIs with hexagonal warping effects can be utilized to design an electron beam splitter with the ability of spatial separation as large as a few micrometers with a high chance of observation of the well-separated beams.

In this paper, we focused on a barrier with interfaces along the $x$ direction ($\Gamma K$ direction in **k** space) which resulted in double refraction and double GH shifts of electron beams. If the barrier extends along the $y$ direction ($\Gamma M$ direction), due to the highly anisotropic nature of hexagonally warped Fermi surface, triple refraction [10,59] and consequently triple GH shifts can emerge. The occurrence of double and triple GH shifts in different directions can be a signature of a hexagonally warped Fermi surface, while they do not occur in other cases such as trigonally or tetragonally warped Fermi surfaces. Also, observing gaps in the GH shift measurements in terms of electron energy can indicate the existence of an energy gap in the band structure of materials [38]. Nevertheless, determining whether the shape of the Fermi surface can be identified with GH shift measurements requires more research. It is worth mentioning that in Weyl semimetals, it has been shown that the GH and IF shifts of a reflected beam from a gapped medium can provide a probe of the topological Fermi arc at the reflecting surface [46].

## IV. A PROPOSAL FOR GH SHIFT MEASUREMENT

To the best of our knowledge, the GH shift in electronic systems has not been experimentally measured yet due to the smallness of GH shift values and the difficulty in producing a well-collimated electron beam [49,65]. Although the magnitude of the GH shift in total reflection from a single-interface (step potential) is about a Fermi wavelength
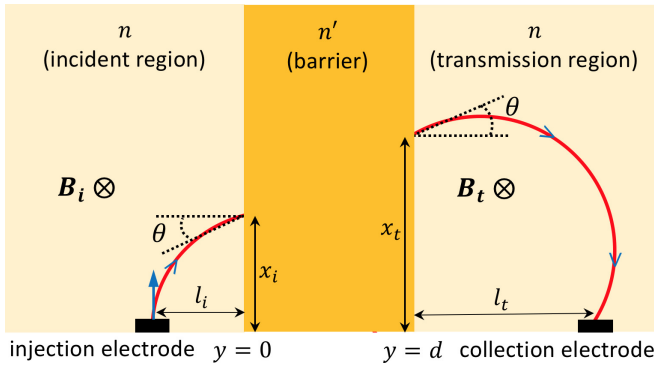
FIG. 8. Schematic view showing the procedure of GH shift measurement. The cyclotron radius $r_i$ in incident region, fixed by both transverse magnetic field $\mathbf{B_i}$ and the Fermi energy, determines the incident angle $\theta$ and also the impact point of electron trajectory at the interface $y = 0$. A varying transverse magnetic field $\mathbf{B_t}$ induces a resonant path with a specific radius $r_t$ which determines the entry point of electron at the interface $y = d$.

of the electron which impedes its direct measurement, it can be enlarged by considering a system acting as a waveguide which causes accumulation of shifts in multiple reflection of electron beams from the waveguide boundaries [29,36,38,65]. Also, in the process of transmitting electrons through potential barrier/well, transmission resonances can occur which enhance the GH shift value considerably [33,34,37,38,43,66]. Note that similar and other mechanisms for amplifying optical GH shifts are considered in literature (see Refs. [67,68]). On the other hand, to directly measure GH shift values, we need a collimator to generate the collimated electron beam and then detect the transmitted/reflected beam from the interface by local gates. Although various proposals for electron collimation in 2D materials [69–73] and surface states of TIs [14,72] are suggested, a decent production of narrow and well-collimated electron beams in such materials has not been attained yet [55,74–76].

Despite the lack of efficient collimation, Chen *et al.* [55] achieved a direct measurement of angle-dependent transmission probability based on TMF measurement scheme. They applied a transverse magnetic field on electrostatically defined $n$-$n'$ ($p$-$n$) junction on graphene and measured the transresistance proportional to the transmission of electrons between an injection electrode (at $n$ side) and a collection electrode (at $n'$ side), while sweeping magnetic field and gate voltage of $n'$ side. In this way, they reached a map in which the first and higher-order resonant peaks appeared. Moreover, using a semiclassical Billiard model, they performed a simulation of electron trajectories whose result was well-matched with that of experimental data. As a result, they reverse-engineered the first-order resonant transport by considering a trajectory for electrons similar to the one that we consider in Fig. 8, which clearly gives the peak positions observed in the experiment as well as in the simulation.

Here, by applying a similar TMF measurement scheme, we propose a procedure for electrons' GH shift measurement on the surface of a TI junction, as schematically shown in Fig. 8. We consider a positive GH shift which mostly occur in the $n$-$n'$-$n$ case. Before explaining the procedure, we give a brief

discussion about survivability of surface states in the presence of a transverse magnetic field. In a TMF phenomenon, it is assumed that the motion of electrons is ballistic, following the classical trajectories [50–52]. This is justified when the electron mean free path $l_e$ is larger than the width of the device in the $x$ direction as well as the separation between the electrodes and interfaces ($l_i$ and $l_t$). The length of $l_e$ is estimated 120 nm for surface electrons in $Bi_2Te_3$ [77]. When the surface classical electrons are subjected to a transverse magnetic field $B$, they follow circular cyclotron orbits with radius $r = E_F/(ev_F B)$ given by Lorentz force, where $e$ is the charge of electron. If the system is treated quantum mechanically, these orbitals get quantized into Landau levels, giving rise to chiral edge states. However, when the magnetic field is not too high, the Landau levels undergo a collapse transition and the edge states can be avoided [17,78]. To match our procedure with the above-mentioned experiment [55], we consider electrons with low Fermi energy in the incident and transmission regions with a circular shape of energy contour similar to that of graphene, which makes an accurate control of electron trajectories in the presence of magnetic field [16]. In the barrier ($n'$) region, there is no magnetic field and the warping effect can be remarkable for large enough $V_0$ values. Under a transverse magnetic field $B_i$, applied on incident electrons injected from a narrow injection electrode, located at the bottom of this region, the electrons undergo a cyclotron motion with radius $r_i = \frac{E_F}{ev_F B_i}$. With some simple calculations, the impact point of electrons and their incident angle at the interface $y = 0$ can be determined as $x_i = \sqrt{l_i(2r_i - l_i)}$ and $\theta = \arctan(\frac{r_i - l_i}{x_i})$, respectively, where $l_i$ is the distance between the injection electrode and the interface (see Fig. 8).

Now we apply an independent transverse magnetic field $B_t$ on the transmission region. $B_t$ bends the transmitted electrons' path down the device into a cyclotron orbit. If the electrons enter the collection electrode placed at the bottom of this region, a peak in the transresistance between the injection and collection electrodes (or corresponding voltage) will occur. Therefore, by tuning $B_t$ in the experiment, it is possible to obtain the amount of $B_t$ and the corresponding radius $r_t = \frac{E_F}{ev_F B_t}$ for which a resonance in magnetic focusing takes place. Having $r_t$ and knowing the angle of incoming electrons (equal to the incident angle) into this region and the distance $l_t$ between the collection electrode and the interface $y = d$, after some straightforward algebraic computations, one can determine the position of the entry point of electrons at the interface $y = d$ as $x_t = r_t \cos\theta + \sqrt{r_t^2 - (l_t - r_t \sin\theta)^2}$. Finally, the GH shift can be obtained by $\Delta_{tr} = x_t - x_i$.

It is important to note that there is a correspondence between the variables used in Chen *et al.*'s experiment [55] and the variables in our proposal. In their experiment, the same magnetic field $B$ is applied to both incident and transmission regions, and the magnetic field $B$ as well as the gate voltage of the transmission region are varied. Also, the angle of entry of electrons into the transmission region ($\theta'$) is different from the incident angle $\theta$. Moreover, $\theta'$ is a function of $\theta$ according to the Snell's law $\sin\theta' = ((E_F - V_i)/(E_F - V_t))\sin\theta$, where $V_i$ and $V_t$ are the gate voltage of incident and transmission regions, respectively. In our proposal, different magnetic fields $B_i$ and $B_t$ are applied to the incident and transmission regions, respectively. The gate voltage of the barrier region is fixed,

while $B_i$ and $B_t$ are variable during the experiment. On the other hand, the transmitted electrons enter the transmission region with the same angle as the incident angle $\theta$ but with a lateral shift $\Delta_{\mathrm{tr}}$ which is a function of the incident angle. Because of such correspondences, we expect that the electrons contributing in the transresistance peak to be those electrons that leave the injection electrode vertically, just like the Chen *et al.* experiment [55].

The above discussion can be presented in a general form as follows. Consider the electrons that leave the injection electrode with arbitrary injection angle $\beta$ (with respect to $y$ axis) at a given $B_i$. The incident angle of these electrons can be calculated as $\theta = \arcsin(\sin \beta - l_i/r_i)$. At a fixed $B_t$, due to the dependence of impact point of electrons on the interface at $y = 0$ and also $\Delta_{\mathrm{tr}}$ to $\theta$, the electrons reach the edge of the device ($y$ axis) in the transmission region at a position that depends on their injection angle $\beta$. Calculating the derivative of $\theta$ with respect to $\beta$, we obtain $d\theta/d\beta = \cos \beta / \sqrt{\cos^2 \beta + 2(l_i/r_i)\sin \beta - l_i^2/r_i^2}$ whose magnitude is zero at $\beta = \pi/2$. This means that the electrons which leave the injection electrode in the small vicinity of angle $\beta = \pi/2$ have the same incident angle $\theta$, and hence the same lateral shift $\Delta_{\mathrm{tr}}$, resulting in the largest density of electrons at a point on the edge of the device in the transmission region. By sweeping the magnetic field $B_t$, $r_t$ is varied, so at a fixed collection electrode position, a peak in transresistance belonging to the electrons that leave the injection electrode vertically, appears. By tuning $B_t$ on larger amounts, the cyclotron radius $r_t$ is reduced, so the electrons can reach the collection electrode after one or more specular reflection from the interface and/or the edge of the device, leading to the formation of the next peaks [52].

Although the present proposal of GH shift measurement was applied to the surface state of Bi$_2$Te$_3$ consisting of a single nondegenerate Dirac cone, this approach can also be utilized in 2D conventional systems such as graphene and other single-layer hexagonal crystals. Nevertheless, in materials consisting of multivalleys, multirefraction can appear, making the observation of the GH shift more complicated than the present paper. Moreover, surface states of TIs are topologically protected against nonmagnetic perturbations compared to the conventional surface states which are sensitively dependent on the geometry of surface structure.

Since in most Dirac materials the GH shift is spin and/or valley dependent, which originates from spin-orbit coupling [36–38,40,66,79], the measurement of the GH shift can provide the possibility of fabrication of spin/valley devices based on EBSs.

## V. CONCLUSION

In summary, we studied theoretically the influence of a hexagonal warping effect on the transport properties and lateral shifts of electrons at the surface of a TI $n$-$n'$-$n$ ($n$-$p$-$n$) junction. It is shown that double refractions occur when the Fermi energy and incident angle of electron beams exceed their critical values. We establish an expression for calculating GH shift values and show that a deflection of the propagation direction of beams from their central propagation directions appears due to the hexagonal warping effect. The dependence of lateral shifts and the corresponding transmissions on system parameters such as incident angle, height, and the width of potential barrier are carefully examined. We show that the system can produce two spatially separated beams with different spin orientations as a result of the GH effect. Therefore, our findings provide an alternative way to construct an electron beam splitter on the basis of TI junctions. Using the physics of TMF phenomenon, we also introduce a procedure for experimentally measuring the GH shift of electron beams in 2D electronic systems which may pave a new route in spin/valleytronics.

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

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

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

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

[5] 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).

[6] K. Kuroda, M. Arita, K. Miyamoto, M. Ye, J. Jiang, A. Kimura, E. E. Krasovskii, E. V. Chulkov, H. Iwasawa, T. Okuda, K. Shimada, Y. Ueda, H. Namatame, and M. Taniguchi, Phys. Rev. Lett. **105**, 076802 (2010).

[7] C. M. Wang and F. J. Yu, Phys. Rev. B **84**, 155440 (2011).

[8] J. Wang, W. Li, P. Cheng, C. Song, T. Zhang, P. Deng, X. Chen, X. Ma, K. He, J. F. Jia, Q.-K. Xue, and B. F. Zhu, Phys. Rev. B **84**, 235447 (2011).

[9] H. K. Pal, V. I. Yudson, and D. L. Maslov, Phys. Rev. B **85**, 085439 (2012).

[10] J. An and C. S. Ting, Phys. Rev. B **86**, 165313 (2012).

[11] P. Rakyta, A. Palyi, and J. Cserti, Phys. Rev. B **86**, 085456 (2012).

[12] Z. B. Siu, M. B. A. Jalil, and S. G. Tan, Sci. Rep. **4**, 5062 (2014).

[13] S. Roy, A. Soori, and S. Das, Phys. Rev. B **91**, 041109(R) (2015).

[14] F. Hassler, A. R. Akhmerov, and C. W. J. Beenakker, Phys. Rev. B **82**, 125423 (2010).

[15] E. V. Repin and I. S. Burmistrov, J. Exp. Theor. Phys. **121**, 509 (2015).

[16] Z.-M. Yu, D.-S. Ma, H. Pan, and Y. Yao, Phys. Rev. B **96**, 125152 (2017).

[17] R. S. Akzyanov and A. L. Rakhmanov, Phys. Rev. B **97**, 075421 (2018).

[18] R. S. Akzyanov and A. L. Rakhmanov, Phys. Rev. B **99**, 045436 (2019).

[19] F. Goos and H. Hänchen, Ann. Phys. (Leipzig) **436**, 333 (1947).

[20] T. Tamir, J. Opt. Soc. Am. A **3**, 558 (1986).

[21] C. W. Hsue and T. Tamir, J. Opt. Soc. Am. A **2**, 978 (1985).

[22] C.-F. Li, Phys. Rev. Lett. **91**, 133903 (2003).

[23] J. Broe and O. Keller, J. Opt. Soc. Am. A **19**, 1212 (2002).

[24] D. W. Wilson, E. N. Glytsis, and T. K. Gaylord, IEEE J. Quantum Electron. **29**, 1364 (1993).

[25] X. Chen, C.-F. Li, and Y. Ban, Phys. Rev. B **77**, 073307 (2008).

[26] X. Chen, X.-J. Lu, Y. Wang, and C.-F. Li, Phys. Rev. B **83**, 195409 (2011).

[27] P. Gruszecki, M. Mailyan, O. Gorobets, and M. Krawczyk, Phys. Rev. B **95**, 014421 (2017).

[28] Z. M. Yu, Y. Liu, Y. Yao, and S. A. Yang, Phys. Rev. Lett. **121**, 176602 (2018)

[29] C. W. J. Beenakker, R. A. Sepkhanov, A. R. Akhmerov, and J. Tworzydlo, Phys. Rev. Lett. **102**, 146804 (2009).

[30] Z. Wu, F. Zhai, F. M. Peeters, H. Q. Xu, and K. Chang, Phys. Rev. Lett. **106**, 176802 (2011).

[31] F. Zhai, Y. Ma, and K. Chang, New J. Phys. **13**, 083029 (2011).

[32] S. Ghosh and M. Sharma, J. Phys.: Condens. Matter **21**, 292204 (2009).

[33] X. Chen, J.-W. Tao, and Y. Ban, Eur. Phys. J. B **79**, 203 (2011).

[34] Y. Song, H.-C. Wu, and Y. Guo, Appl. Phys. Lett. **100**, 253116 (2012).

[35] X. Chen, P. L. Zhao, and X. J. Lu, Eur. Phys. J. B **86**, 223 (2013).

[36] J. F. Sun and F. Cheng, J. Appl. Phys. **115**, 133703 (2014).

[37] H. Ghadiri and A. Saffarzadeh, J. Phys.: Condens. Matter **29**, 115303 (2017).

[38] H. Ghadiri and A. Saffarzadeh, J. Appl. Phys. **123**, 104301 (2018).

[39] X. J. Qiu, Q. Lv, and Z. Z. Cao, J. Phys.: Condens. Matter **31**, 225303 (2019).

[40] E. S. Azarova and G. M. Maksimova, J. Phys. Chem. Solids **100**, 143 (2016).

[41] J. Kuai and H. X. Da, J. Magn. Magn. Mater. **354**, 355 (2014).

[42] Q. D. Jiang, H. Jiang, H. Liu, Q. F. Sun, and X. C. Xie, Phys. Rev. Lett. **115**, 156602 (2015).

[43] R.-F. Zheng, L. Zhou, and W. Zhang, Phys. Lett. A **381**, 3798 (2017).

[44] Y. Liu, Z. M. Yu, H. Jiang, and S. A. Yang, Phys. Rev. B **98**, 075151 (2018)

[45] Y. Liu, Z. M. Yu, J. Liu, H. Jiang, and S. A. Yang, Phys. Rev. B **98**, 195141 (2018)

[46] U. Chattopadhyay, L. K. Shi, B. Zhang, J. C. W. Song, and Y. D. Chong, Phys. Rev. Lett. **122**, 066602 (2019).

[47] W. Wu, S. Chen, C. Mi, W. Zhang, H. Luo, and S. Wen, Phys. Rev. A **96**, 043814 (2017).

[48] G. Ye, W. Zhang, W. Wu, S. Chen, W. Shu, H. Luo, and S. Wen, Phys. Rev. A **99**, 023807 (2019).

[49] X. Chen, X. J. Lu, Y. Ban, and C. F. Li, J. Opt. **15**, 033001 (2013)

[50] T. Taychatanapat, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, Nat. Phys. **9**, 225 (2013).

[51] S. P. Milovanovic, M. Ramezani Masir, and F. M. Peeters, J. Appl. Phys. **115**, 043719 (2014).

[52] V. S. Tsoi, J. Bass, and P. Wyder, Rev. Mod. Phys. **71**, 1641 (1999).

[53] S. Bhandari, G.-H. Lee, K. Watanabe, T. Taniguchi, P. Kim, and R. M. Westervelt, Nano Lett. **20**, 4890 (2020).

[54] L. P. Rokhinson, V. Larkina, Y. B. Lyanda-Geller, L. N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **93**, 146601 (2004).

[55] S. Chen, Z. Han, M. M. Elahi, K. M. M. Habib, L. Wang, B. Wen, Y. Gao, T. Taniguchi, K. Watanabe, J. Hone, A. W. Ghosh, and C. R. Dean, Science **353**, 1522 (2016).

[56] S. Bhandari, G. H. Lee, A. Klales, K. Watanabe, T. Taniguchi, E. Heller, P. Kim, and R. M. Westervelt, Nano Lett. **16**, 1690 (2016).

[57] S. Bhandari, G. H. Lee, P. Kim, and R. M. Westervelt, J. Electron. Mater. **46**, 3837 (2017).

[58] S. Bhandari, G. H. Lee, K. Watanabe, T. Taniguchi. P. Kim, and R. M. Westervelt, 2D Mater. **5**, 021003 (2018).

[59] M.-M. Wu, Z.-M. Yu, and H. Pan, Phys. Lett. A **383**, 237 (2019).

[60] M. Arabikhah and A. Saffarzadeh, J. Phys.: Condens. Matter **31**, 445001 (2019).

[61] At $V_0 = E_F$, the angle in which TIR begins is dependent on $d$ value. By increasing $d$, although the angle decreases, it never reaches zero due to the Klein tunneling effect.

[62] O. Klein, Z. Phys. **53**, 157 (1929).

[63] M. I. Katsnelson, K. S. Novoselov, and A. K. Geim, Nat. Phys. **2**, 620 (2006).

[64] In Eq. (14), the second and third terms can be ignored in the vicinity of the peak positions compared to the first term.

[65] Z.-M. Yu, Y. Liu, and S. A. Yang, Front. Phys. **14**, 33402 (2019).

[66] See the results of this paper in Sec. III.

[67] C.-F. Li and Q. Wang, Phys. Rev. E **69**, 055601(R) (2004).

[68] S. Chen, C. Mi, L. Cai, M. Liu, H. Luo, and S. Wen, Appl. Phys. Lett. **110**, 031105 (2017).

[69] C.-H. Park, Y.-W. Son, L. Yang, M. L. Cohen, and S. G. Louie, Nano Lett. **8**, 2920 (2008).

[70] V. V. Cheianov, V. Falíko, and B. L. Altshuler, Science **315**, 1252 (2007).

[71] A. G. Moghaddam and M. Zareyan, Phys. Rev. Lett. **105**, 146803 (2010).

[72] M.-H. Liu, C. Gorini, and K. Richter, Phys. Rev. Lett. **118**, 066801 (2017).

[73] Y. Betancur-Ocampo, F. Leyvraz, and T. Stegmann, Nano Lett. **19**, 7760 (2019).

[74] F. Libisch, T. Hisch, R. Glattauer, L. A. Chizhova, and J. Burgdorfer, J. Phys.: Condens. Matter **29**, 114002 (2017).

[75] X. Zhou, A. Kerelsky, M. M. Elahi, D. Wang, K. M. M. Habib, R. N. Sajjad, P. Agnihotri, J. U. Lee, A. W. Ghosh, F. M. Ross, and A. N. Pasupathy, ACS Nano **13**, 2558 (2019).

[76] S. W. LaGasse and C. D. Cress, Nano Lett. **20**, 6623 (2020).

[77] Y. Xie, Y. Tan, and A. W. Ghosh, Phys. Rev. B **96**, 205151 (2017).

[78] Y. Li, Q. Wan, Y. Peng, G. Wang, Z. Qian, G. Zhou, and M. B. A. Jalil, Sci. Rep. **5**, 18458 (2015).

[79] In graphene, the electronic GH shift is spin and valley independent. However, with magnetic barrier and/or considering small intrinsic spin-orbit coupling or extrinsic Rashba-type spin-orbit coupling, the GH shift will depend on spin and valley.