Spintronic signatures of Klein tunneling in topological insulators

Yunkun Xie, Yaohua Tan, and Avik W. Ghosh

Phys. Rev. B **96**, 205151 — Published 28 November 2017

DOI: 10.1103/PhysRevB.96.205151

Spintronic signatures of Klein tunneling in topological insulators

Yunkun Xie,* Yaohua Tan, and Avik W. Ghosh
Charles L. Brown Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, 22904 USA.

Klein tunneling, the perfect transmission of normally incident Dirac electrons across a potential barrier, has been widely studied in graphene and explored to design switches, albeit indirectly. We show an alternative way to directly measure Klein tunneling for spin-momentum locked electrons crossing a PN junction along a three dimensional topological insulator surface. In these topological insulator PN junctions (TIPNJs), the spin texture and momentum distribution of transmitted electrons can be measured electrically using a ferromagnetic probe for varying gate voltages and angles of current injection. Based on transport models across a TIPNJ, we show that the asymmetry in the potentiometric signal between PP and PN junctions and its overall angular dependence serve as a direct signature of Klein tunneling.

I. INTRODUCTION

Klein tunneling - a consequence of quantum electrodynamics where relativistic particles pass through a high potential barrier unimpeded[1] - is an intriguing phenomenon that has yet to be directly observed in experiments. Researches closest to testing the KT phenomenon are mostly conducted in graphene, with recent progress in the demonstration of anomalous broadened quantized states in a graphene quantum dot[2] and negative index[3] in graphene. It has also been invoked to engineer a gate tunable pseudogap in graphene at high mobility, making it potentially useful for both low power digital and high speed analog switches[4–6]. Exciting as it is, a direct measurement of Klein tunneling in graphene is very hard because electron flow in graphene sums over all momenta equally and current measurements cannot differentiate those mixed electron momenta. To overcome this difficulty, recent progress in graphene focused on either collimating electrons through a particular gate geometry[7] or through a specially designed electron source[8]. Both methods narrow the electron momenta distribution but can potentially suffer from gate edge roughness or reduced signal intensity due to electron absorption in the source structure. Here we propose an alternative experimental setup to measure KT in a different system that doesn't need complicated gate/source structure - 3D topological insulator (TI) surface. The TI surface, such as Bi_2Se_3, has a simple Dirac cone band structure[9] reminiscent of graphene, except its branches are labeled by spins rather than pseudospins. This unique band structure makes TI a potential candidate for spintronics applications: Carriers along the surface have their spins locked with their linear momentum[10], which can generate polarized spins with charge injection and apply a sizeable spin torque on a magnet[11–13]. Recently we suggested that a TIPNJ can be used as a gate tunable spin filter to amplify charge to spin conversion at a magnetic source and increase spin polarization at the drain[14]. Such a tunable torque can have potential applications in all spin logic[15]. Beyond applications, the TI surface state offers opportunities to study the fundamental physics of Dirac electrons such as Veselago focusing and Klein tunneling. In this paper, we propose a new way to measure KT in 3D TI surface. The core of our proposed idea relies on the measurement of electron spin potential on the TI surface through a spin selective ferromagnetic probe. Since momenta couple with spin on TI surface, the spin selective probe can also be momentum selective. The method of potentiometric measurement with a ferromagnetic probe to detect the spin structure on TI surface has been well-established both theoretically[16,17] and experimentally[18–21]. Here we model a potentiometric measurement on a TIPNJ and demonstrate from detailed calculations that the angle and voltage dependent potentials measured at the probe bear direct signatures of Klein tunneling across the PN junction.

Our paper is organized as follows: In section II, we first define the proposed experimental setup as well as theoretical model for TIPNJ and ferromagnetic probe. Then we derive the analytical equations followed by a brief summary of the numerical techniques. In section III we describe the simulation results for two different experimental setups. In section IV we discuss possible solutions to some realistic issues expected in the proposed experiments.

II. MODELING METHODS

Fig. 1(a) shows a schematic structure of the TI pn junction in a potentiometric measurement setup. The TI surface can be chemically doped into P or N-type, as demonstrated in multiple experiments[22,23]. The figure shows a P-doped TI surface with a top gate on the source side that can swing it electrostatically to N-type. Recent experiment has already shown an innovative way to put atomically abrupt gate on TI to create in-plane pn Junction[24]. The rest of the P-type TI surface is exposed and a ferromagnetic probe is placed on top of the exposed surface to monitor the voltage at different gate bias and angular orientations (the orientation can be altered by using multiple contacts at relative angles, as we discuss later).

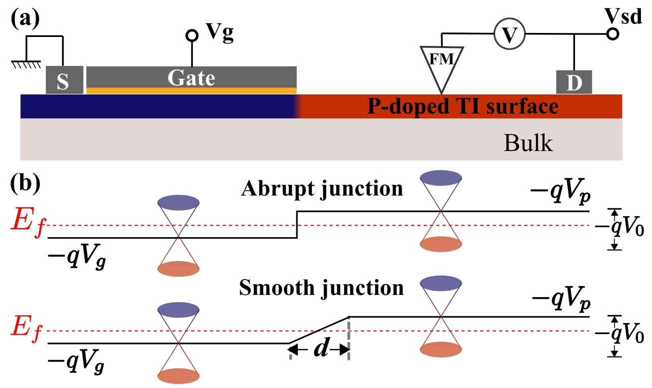

FIG. 1. (a) A basic setup for potentiometric measurement on a topological insulator PN junction. (b) The electrostatic potential profiles (abrupt and smooth) across TI PN junction.

The TI surface states can be described by the $k \cdot p$ Hamiltonian when the electron energy under consideration is close to the Dirac point[10]:

$$H = v_F \hat{\mathbf{z}} \cdot (\boldsymbol{\sigma} \times \mathbf{p}) \tag{1}$$

where $\hat{\mathbf{z}}$ is the normal vector of the surface and v_F is the speed of electrons near the Dirac point. $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. It should be emphasized that this parameterized surface Hamiltonian ignores any bulk leakage current that could control the strength of the measured voltage. In binary TI compounds such as BiSb, Bi_2Se_3, Bi_2Te_3, it can be challenging to separate the surface contribution from the dominant bulk contribution[25–27]. One possible solution is to use ternary compounds like Bi_2Te_2Se with low carrier density in the bulk[28]. Minimizing the leakage current into the bulk of TI is still an active research topic that is outside the scope of this paper. Here we only discuss the pure surface states of 3D TI.

The electrostatic potential across the TI PN junction is given by:

$$\begin{aligned} V(x) &= -qV_p, \quad \text{exposed P side} \\ &= -qV_g', \quad \text{gate side} \end{aligned} \tag{2}$$

where $E_p = -qV_p$ is the energy difference between the local electron chemical potential and the Dirac point ($E = 0$). V_g' is the effective potential on the source side of the TI surface under the gate voltage V_g as shown in Fig.1(a). For the simplicity of the discussion, we assume good electrostatic control of the gate on the TI surface (gate capacitance much larger than other capacitors in the system) that gives $V_g' \approx V_g$. Two potential profiles are depicted in Fig. 1(b), one with an abrupt potential change at the junction interface while the other assumes a smooth transition. Later we will first derive the analytical results for electron transmission in abrupt junctions and then extend it to smooth junctions, which is closer to a realistic profile[2]. For smooth junctions, the transition region between N and P is set to 50 nm wide and the FM probe is placed 80 nm from the junction interface.

In the ballistic limit, the electrons only scatter near the PN junction interface. A weakly coupled ferromagnetic voltage probe can detect the local chemical potential of the non-equilibrium electrons with different spin orientations. To calculate the voltage measured by the FM probe, we treat it as a third contact (Büttiker probe) besides source and drain. From Landauer theory[29,30], the exchange of electrons between the voltage probe and the TI surface follows the following equations:

$$\begin{aligned} I_{in} &= \mathrm{Tr}\left[\Gamma_{\mathrm{FM}} G^n\right] = \mathrm{Tr}\left[\Gamma_{\mathrm{FM}}\left(f_s A_s + f_d A_d\right)\right] \\ I_{out} &= f_p \mathrm{Tr}\left[\Gamma_{\mathrm{FM}} A\right] = f_p \mathrm{Tr}\left[\Gamma_{\mathrm{FM}}\left(A_s + A_d\right)\right] \end{aligned} \tag{3}$$

where $I_{in}\,(I_{out})$ is the incoming (outgoing) currents through the probe. Γ_{FM} is the coupling between the FM probe and the TI surface. G^n is the correlation matrix while $A_s(A_d)$ are the partial spectral functions populated by the source (drain). $A = A_s + A_d$ is the total spectral function. f_s, f_d, f_p are the Fermi-Dirac distribution functions of the source, drain and the floating probe respectively.

The coupling between the FM probe and the TI surface depends on the magnetization of the FM probe $\mathbf{m} = (m_x, m_y, m_z)$ and electron spin $\boldsymbol{\sigma}$ of the TI surface:

$$\Gamma_{\mathrm{FM}}(\mathbf{m}) = \gamma_0 \left(1 + P_{\mathrm{FM}} \mathbf{m} \cdot \boldsymbol{\sigma}\right) \tag{4}$$

where $\gamma_0 = \frac{\gamma_p + \gamma_{ap}}{2}$ is the average coupling between the FM probe and the TI surface when the magnetization of the probe is in parallel or anti-parallel alignment with the surface electron spin. $P_{\mathrm{FM}} = (\gamma_p - \gamma_{ap})/(\gamma_p + \gamma_{ap})$ is the 'polarization' of the FM probe, representing the sensitivity of the FM probe to the electron spins.

The voltage signal measured by the FM probe is determined by its distribution function f_p, which can be solved based on the condition that a voltage probe draws zero net current $I_{in} = I_{out}$:

$$\begin{aligned} f_p(\mathbf{m}) &= \frac{(f_s - f_d)\,\mathrm{Tr}\left[\Gamma_{\mathrm{FM}} A_s\right]}{\mathrm{Tr}\left[\Gamma_{\mathrm{FM}} A\right]} + f_d \\ &= \lambda(\mathbf{m})(f_s - f_d) + f_d \end{aligned} \tag{5}$$

f_p varies when the magnetization $\mathbf{m}$ points to different directions. We use the dimensionless parameter $\lambda(\mathbf{m})$ to characterize the dependence of the voltage signal on the direction of the magnetization. At low-temperature and small bias, the Fermi-Dirac distribution reduces to a step function and chemical potential of the probe can be expressed as:

$$\mu_p(\mathbf{m}) = \lambda(\mathbf{m})(\mu_s - \mu_d) + \mu_d \tag{6}$$

Experimentally instead of switching the magnetization of the FM probe we can drive current along two opposite directions (source to drain and vice-versa), then relate the measured voltage difference $\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m})$ to $\Delta\lambda(\mathbf{m}) = \lambda(\mathbf{m}) - \lambda(-\mathbf{m})$ through the charge current and the ballistic resistance of the junction:

$$\Delta\lambda(\mathbf{m}) = \frac{\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m})}{qIR_B}$$
$$R_B = \frac{h}{q^2 T(E_f)} \tag{7}$$

where R_B is the gate voltage dependent ballistic resistance of the junction, calculated using the average transmission at the Fermi energy.

We can further define a quantity $p(\mathbf{m})$ for the measured 'polarization' of the TI surface electrons along the magnetization direction $\mathbf{m}$:

$$p(\mathbf{m}) = \frac{\lambda(\mathbf{m}) - \lambda(-\mathbf{m})}{\lambda(\mathbf{m}) + \lambda(-\mathbf{m})} = \frac{\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m})}{\mu_p(\mathbf{m}) + \mu_p(-\mathbf{m}) - 2\mu_d} \tag{8}$$

The physical interpretation of Eq. 8 becomes obvious when we substitute Eq. 4 into Eq. 8 and see that $\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}(\mathbf{m})A\right] = \mathrm{Tr}\left[\Gamma_{\mathrm{FM}}(-\mathbf{m})A\right]$ due to the time reversal symmetry of TI surface states. Eq. 8 reduces to:

$$p(\mathbf{m}) = P_{\mathrm{FM}} \frac{\mathrm{Tr}[(\mathbf{m}\cdot\boldsymbol{\sigma})\gamma_0 A_s]}{\mathrm{Tr}[\gamma_0 A_s]} \tag{9}$$

when Eq. 9 is evaluated in the bias window, it indicates the spin polarization of the non-equilibrium electrons along direction $\mathbf{m}$. Notice that P_{FM} also appears in the equation to account for the sensitivity of FM probe. Our definition is compatible with the polarization defined in[16] for homogeneous TI surface.

A. Analytical formalisms

For infinitely large TI surface with an abrupt PN junction potential profile, the eigen-functions to Eq. 1 are given by:

$$|\psi\rangle_\sigma = \frac{1}{\sqrt{2S}} \begin{pmatrix} 1 \\ -sie^{i\theta} \end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{10}$$
$$s = \mathrm{sgn}(E_{\mathbf{k}})$$

where $\mathrm{sgn}(E_k) = 1$ is for the N type and $\mathrm{sgn}(E_k) = -1$ for the P type. S is the surface area. At the PN junction, the transmitted and scattered electrons are connected by:

$$\begin{aligned} |\psi\rangle_\sigma &= |\psi_i\rangle_\sigma + r|\psi_r\rangle_\sigma \\ |\psi\rangle_\sigma &= t|\psi_t\rangle_\sigma \end{aligned} \tag{11}$$

where $|\psi_i\rangle_\sigma, |\psi_r\rangle_\sigma, |\psi_t\rangle_\sigma$ are the incoming, reflected and transmitted electron wave functions respectively (see Fig. 2) and r/t is the reflection/transmission coefficient. Solving Eq. 11 with wave function continuity condition at the junction $\mathbf{r} = 0$, we get the transmission coefficient:

$$\text{for NP} \qquad t = \frac{e^{i\theta_i} - e^{-i\theta_t}}{e^{-i\theta_i} - e^{i\theta_t}} \tag{12}$$

$$\text{for PP} \qquad t = \frac{e^{i\theta_i} - e^{i\theta_t}}{e^{-i\theta_i} - e^{-i\theta_t}} \tag{13}$$

It is convenient to replace θ_t with $\theta_t + \pi$ in the NP case so that the expressions for t are the same in both PP and NP cases. The incident and transmitted angles are connected through the conservation of k_y across the junction: $(E - qV_n)\sin\theta_i = (E - qV_p)\sin\theta_t$, we can calculate the transimission probability $|t|^2$:

$$|t|^2 = \frac{\cos^2\theta_i}{\cos^2\left(\frac{\theta_i+\theta_t}{2}\right)} \tag{14}$$

Smooth PN junction. The effect of a smooth PN junction (as shown in Fig. 1(b)) is an additional exponential factor from the abrupt junction case (Eq.14). Here we borrow the result for the transmission coefficient $|t|^2_{\mathrm{smooth}}$ from the smooth graphene PN junction (see details in[6]):

$$\begin{aligned} |t|^2_{\mathrm{smooth}} &= \frac{\cos^2\theta_i}{\cos^2\left(\frac{\theta_i+\theta_t}{2}\right)} e^{-\pi\frac{k_i k_t}{k_i+k_t}\sin\theta_i\sin\theta_t d} \\ &= \frac{\cos^2\theta_i}{\cos^2\left(\frac{\theta_i+\theta_t}{2}\right)} e^{-\pi\frac{\hbar v_F}{qV_0}k_t^2\sin^2\theta_t d} \end{aligned} \tag{15}$$

where d is the transition length between N region and P region. $V_0 = |V_p - V_n|$ is the potential difference from N region to P region. Notice that the exponential factor should only be added in cases with different types (such as PN, NP) across the junction . In other cases (such as PP', NN'), the difference between abrupt and smooth junctions is negligible, which can be seen in our comparison between analytical and numerical results later.

In the small bias window near E_f, the charge current is given by:

$$I = \frac{q}{h}T(E_f)(\mu_s - \mu_d)$$
$$T(E_f) = \frac{qV_pW}{hv_F}\int_{-\pi/2}^{\pi/2}|t|^2\cos\theta_t d\theta_t \tag{16}$$

where $T(E_f)$ is the electron transmission across the junction and W is the width of the TI surface. Knowing the transmission coefficient allows us to calculate $\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}A_s\right]$

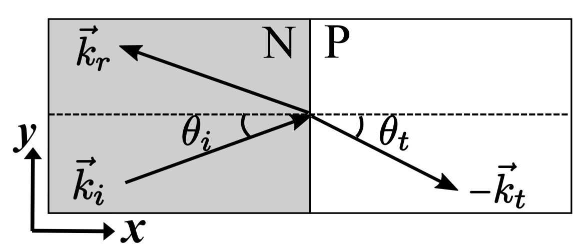

FIG. 2. Incident, reflected and transmitted electrons waves in a TI pn junction.

(Eq. 3-6):

$$\begin{aligned}&\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}A_s\right]\\ &=W\sum_{v_x(\mathbf{k}_t)>0}[1+P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{s}(\mathbf{k}_t)t(\mathbf{k}_t)]\delta(E_f-E(\mathbf{k}_t))\end{aligned}\tag{17}$$

$\mathbf{s}(\mathbf{k}_t)$ is the spin orientation of the transmitted electron with wave vector $\mathbf{k}_t$. $t(\mathbf{k}_t)$ is the transmission coeffecient given by Eq. 13. $\lambda(\mathbf{m})$ in Eq. 6 can then be calculated:

$$\begin{aligned}\lambda(\mathbf{m})&=\frac{\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}A_s\right]}{\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}A\right]}\\ &=\frac{\sum_{v_x(\mathbf{k}_t)>0}[1+P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{s}(\mathbf{k}_t)t(\mathbf{k}_t)]\delta(E_f-E(\mathbf{k}_t))}{\sum_{\mathbf{k}_t}[1+P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{s}(\mathbf{k}_t)]\delta(E_f-E(\mathbf{k}_t))}\\ &=\frac{\sum_{v_x(\mathbf{k}_t)>0}[1+P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{s}(\mathbf{k}_t)t(\mathbf{k}_t)]\delta(E_f-E(\mathbf{k}_t))}{\sum_{\mathbf{k}_t}\delta(E_f-E(\mathbf{k}_t))}\end{aligned}\tag{18}$$

The last step in Eq. 18 holds because each pair of states $\mathbf{k}_t, -\mathbf{k}_t$ cancel each other due to the time reversal symmetry of TI surface Hamiltonian $\mathbf{s}(\mathbf{k}_t)=-\mathbf{s}(-\mathbf{k}_t)$. Assume the ferromagnetic voltage probe has an in-plane magnetization (m_x, m_y). Substitute the transmission coefficient into Eq. 18 and replace $\sum$ with $\frac{S}{4\pi^2}\int d^2k$. For the denominator, notice that it is just the density of states on the P side. Therefore $\lambda(\mathbf{m})$ can be calculated:

For PP:

$$\begin{aligned}\lambda(\mathbf{m})=&\left|\frac{E_f-qV_g}{E_f-qV_p}\right|\times\\ &\int_{-\pi/2}^{\pi/2}\frac{\cos^2\theta_i\left(1+P_{\mathrm{FM}}m_x\sin\theta_t-P_{\mathrm{FM}}m_y\cos\theta_t\right)}{2\pi\cos^2\left(\frac{\theta_i+\theta_t}{2}\right)}d\theta_t\end{aligned}\tag{19}$$

For NP:

$$\begin{aligned}\lambda(\mathbf{m})=&\left|\frac{E_f-qV_g}{E_f-qV_p}\right|\times\\ &\int_{-\pi/2}^{\pi/2}\left\{\frac{\cos^2\theta_i\left(1+P_{\mathrm{FM}}m_x\sin\theta_t-P_{\mathrm{FM}}m_y\cos\theta_t\right)}{2\pi\cos^2\left(\frac{\theta_i+\theta_t}{2}\right)}\right.\\ &\left.\cdot\exp\left[\frac{(E_f-qV_p)^2\pi d}{|V_g-V_p|q\hbar v_F}\right]\right\}d\theta_t\end{aligned}\tag{20}$$

where $\theta_t=\sin^{-1}[(E_f-qV_g)/(E_f-qV_p)\sin\theta_i]$.

B. Numerical approach

In general cases, the ballistic electron/spin transport on the TI surface can be numerically modeled with the Non-Equilibrium Green's Function (NEGF) method. An artificial term $\sigma^z=\gamma\hbar v_F\sigma^z(k_x^2+k_y^2)$ is added to the surface Hamiltonian Eq.1 to avoid the fermion doubling problem as have been done in the previous studies[14,16]. The modified TI surface Hamiltonian is discretized on a square lattice by the finite difference method[14]:

$$\begin{aligned}H=&\sum_i\epsilon c_i^\dagger c_i+\sum_i\left(t_x c_{i,i}^\dagger c_{i,i+1}+\mathrm{H.C.}\right)\\ &+\sum_j(t_y c_{j,j}^\dagger c_{j,j+1}+\mathrm{H.C.})\end{aligned}\tag{21}$$

$$\epsilon=-4\hbar v_F\frac{\alpha}{a}\sigma^z\quad t_x=\hbar v_F\left[\frac{i}{2a}\sigma^y+\frac{\alpha}{a}\sigma^z\right]\tag{22}$$

$$t_y=\hbar v_F\left[-\frac{i}{2a}\sigma^x+\frac{\alpha}{a}\sigma^z\right]\tag{23}$$

where a is the square mesh size ($a=5\,\mathrm{nm}$ is chose for the simulations). $\alpha=\gamma/a$ is a fitting parameter and $\alpha=1$ describes the correct bandstructure near the Dirac cone[14]. Periodic boundary condition is assumed in the transverse direction to simulate infinitely wide TI surface. The retarded green's function is given by:

$$G^R(E,\mathbf{k}_\perp)=(E+\delta-H(\mathbf{k}_\perp)-\Sigma_s(E,\mathbf{k}_\perp)-\Sigma_d(E,\mathbf{k}_\perp))^{-1}\tag{24}$$

where E is the energy and $\mathbf{k}_\perp$ is the transverse wavevector. $\Sigma_{s,d}$ are self-energies from the source and drain. The FM probe is assumed to be weakly coupled to the TI surface so the effect of Σ_p (assign a very small value) on electron transport is neglected when calculating $G^R(E,\mathbf{k}_\perp)$. Then the spectral functions can be calculated numerically through the NEGF formalism:

$$\begin{aligned}A_s&=G^R\Gamma_s G^{R\dagger},\quad\Gamma_s=i(\Sigma_s-\Sigma_s^\dagger)\\ A_d&=G^R\Gamma_d G^{R\dagger},\quad\Gamma_d=i(\Sigma_d-\Sigma_d^\dagger)\end{aligned}\tag{25}$$

$\lambda(\mathbf{m})$ is then calculated from the matrix forms of Γ_{FM}, A_s, A.

III. RESULTS

A. Varying gate voltage: from PP to NP junction.

The impact of a TIPNJ on surface electron transport is summarized schematically in Fig. 3(a). Consider a small source-drain bias near the Fermi energy, as shown in Fig. 1(b). As the gate voltage varies from $V_g=V_p$ to $V_g=-V_p$, the TI switches from a homogeneous P-doped surface to an NP junction. Electrons see a potential barrier from the N region to the P region. In a normal semiconductor, such a barrier creates decaying electron waves in the P region and results in a vanishing current. For Dirac type TI surface, however, the junction acts like a collimator for electrons, filtering out electrons with large incident angles but preserving the normally incident modes that cannot back-scatter due to spin conservation. The resulting electron transmission for various gate voltages is plotted in Fig. 3(a). This behavior can translate to the gate voltage dependence of $\Delta\lambda(\mathbf{m})$ defined in

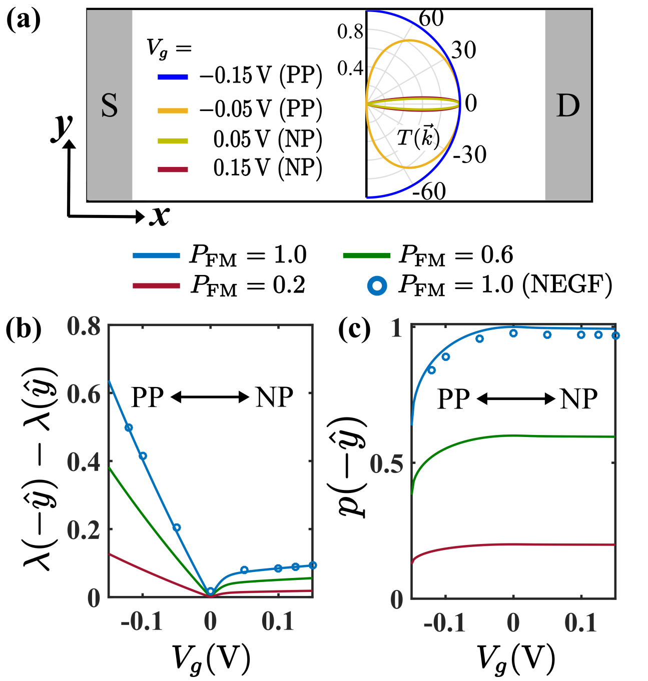

FIG. 3. (a) Schematic plot of the electron transmission through the junction at different gate voltages. (b) Gate voltage dependence of $\Delta\lambda(-\hat{\mathbf{y}}) = \lambda(-\hat{\mathbf{y}}) - \lambda(\hat{\mathbf{y}})$ for various probe sensitivities. (c) The measurable polarization of TI surface electrons along $\hat{y}$ direction. The circles are benchmark results from NEGF simulations.

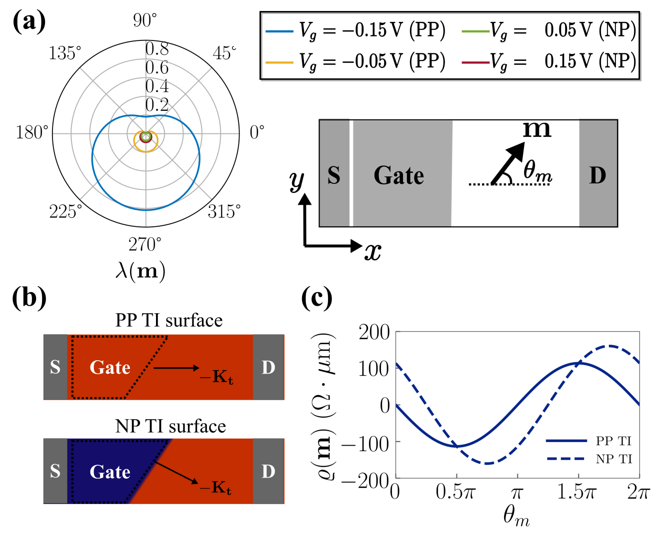

FIG. 4. (a) Angular dependence of $\lambda(\hat{\mathbf{m}})$ for different gate voltages. (b) Schematics of a tilted gate on TI surface. (c) Compare the angular dependence of $\varrho(\mathbf{m})$ in PP and NP cases (See Appendix A for the analysis of $\varrho(\mathbf{m})$).

Eq. 7. Fig. 3(b) shows the gate voltage dependence of $\Delta\lambda = \lambda(-\hat{y}) - \lambda(\hat{y})$. $\Delta\lambda$ first goes down as we move from PP to PI (I: intrinsic), then goes up a bit and saturates in the NP region. The decrease of $\Delta\lambda(\mathbf{m})$ in the PP region is due to a mismatch of modes between the gate side and the probe side as the Fermi energy approaches the Dirac point (intrinsic doping) on the gate side. When $V_g = 0$ V the Fermi level on the gate side lies exactly on the Dirac point with zero density of states and thus $\Delta\lambda(\mathbf{m}) = 0$. It is worth mentioning that the 'zero' is an idealized simplification. A rigorous calculation involves integration over the bias window which would result in a small but non-zero value.

When the gate side is switched to the N region, the angular filtering effect shows up and results in a smaller value of $\Delta\lambda(\mathbf{m})$ compared to its symmetric point (with the same $|V_g|$) in the PP region. Since the normal incident mode is not affected by the potential barrier, a small but near constant $\Delta\lambda(\mathbf{m})$ shows up in the NP region as V_g increases. This asymmetry between PP and NP region and the non-vanishing $\Delta\lambda(\mathbf{m})$ in the NP region separates the TI surface from other 2D systems such as graphene or Rashba systems where there is either $\Delta\lambda(\mathbf{m}) = 0$ in all regions due to spin degeneracy (graphene) or $\Delta\lambda(\mathbf{m}) = 0$ in the transmitted N region due to decaying waves in a potential barrier for massive tunneling electrons (Rashba).

We can further demonstrate collimation in TIPNJ by plotting polarization $p(-\hat{\mathbf{y}})$ as a function of the gate voltage, as shown in Fig. 3(c). Electrons moving along the $\hat{\mathbf{x}}$ direction carry $-\hat{\mathbf{y}}$ spin. Right across the NP junction, filtered electrons have a narrower $\mathbf{k}$ distribution compared to the homogeneous PP case, and thus higher (close to 100%) spin polarization. In reality, this kind of measurement is limited by the sensitivity of the FM probe, but a clear and significant increase of polarization should be observable as we proceed from homogeneous PP case to NP doping with reasonable P_{FM} values.

B. Angular dependence of $\lambda(\mathbf{m})$.

Our discussion so far focused on measurement along two opposite directions ($\pm\hat{\mathbf{y}}$), assumed to be orthogonal to the electron transport direction. For an arbitrary orientation of the magnetization $\mathbf{m}$, $\lambda(\mathbf{m})$ is a cosine function of the relative angle between the magnetization $\mathbf{m}$ and the spin orientation of the non-equilibrium electrons. Fig. 4(a) shows the angular dependence of $\lambda(\mathbf{m})$ with different gate voltages. From homogeneous PP to NP junction, apart from the change in the magnitude, $\lambda(\mathbf{m})$ remains the same cosine function. This is because the FM probe cannot isolate individual modes but measures the sum over all transport modes. In our basic setup, the PN junction filters electrons with large incident angles but the transmitted modes are still symmetrically distributed with respect to $\hat{\mathbf{x}}$. Therefore the average momenta in the PP and NP junction only differ from each other by their magnitude. To experimentally observe the normal tunneling mode, we can put a tilted gate that is not orthogonal to the transport direction (see Fig. 4(b)). A tilted gate will not affect the results from the homogeneous case but will collimate the electrons to a different

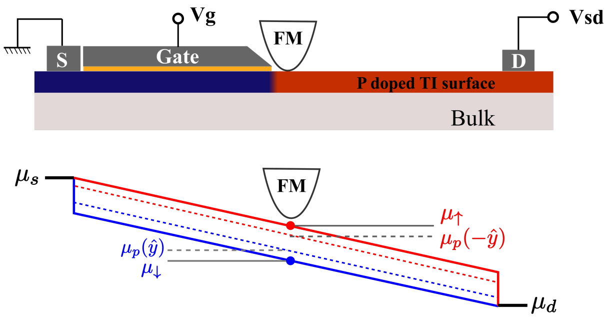

FIG. 5. Top. A possible experimental setup for diffusive system. Bottom. A schematic chemical potential profile in a diffusive system.

angle for NP, thereby creating a phase shift in the angular dependence of $\lambda(\mathbf{m})$. Since we only care about the phase of $\lambda(\mathbf{m})$, we can define an angular function as:

$$\varrho(\mathbf{m}) = \frac{\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m})}{qJP_{\mathrm{FM}}} \tag{26}$$

which will scale $\Delta\mu_p(\mathbf{m})$ by the charge current density J and make the PP and NP cases easier to compare, as shown in Fig. 4(c).

IV. DISCUSSIONS

A. Ballistic versus diffusive limit.

Note that we formulated our equations Eq.3-8 assuming a ballistic channel where $\mu_p(\mathbf{m})$ can be directly related to the chemical potentials from the source and drain. However, our analysis can be easily adopted to a diffusive system with a different interpretation. μ_s and μ_d in the previous discussions should be replaced by the local chemical potential $\mu_\uparrow$ and $\mu_\downarrow$ for spin up and spin down channels, as indicated in Fig. 5. All of our previous discussions are still valid given the following conditions: in a diffusive system, a momentum scattering event can disrupt the collimation effect of the NP junction. To be able to detect the Klein tunneling physics of the junction, the probe needs to be placed very close to the junction, preferably within the mean free path of the TI surface electrons ($\sim$ 120 nm estimated in Bi_2Te_3[31]). To place the probe in such short distance from the gate edge, it possibly requires either a very thin gate ($<$ 100 nm) or specially etched shape (as shown in Fig. 5) to avoid crashing with the probe. From the discussion of $p(\mathbf{m})$ earlier, we need information on μ_d (replace by $\mu_\downarrow$) at the junction. One way to do this is to use a normal voltage probe to map out the resistance from junction to the drain to extract the slope shown in Fig. 5, and then estimate the local electrochemical potential from the applied drain bias.

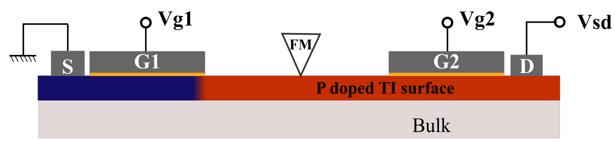

FIG. 6. One possible experimental measurement set-up. Intrinsically P-doped topological insulator under a N type gate near the source. The FM probe is placed on the exposed P side.

B. Possible experimental set-up.

Ideally we would like to rotate the magnetization of the ferromagnetic probe to map out the angle-dependent voltage signals. To our knowledge such a reorientation of an FM probe is challenging. Even fixing the magnetization of the FM probe orthogonal to the transport direction is not straightforward. Instead, we propose placing two separate gates near the source and drain (Fig. 6), creating a symmetric system. Only one of the gates is used at a time to create an N region on one side. When the current direction is switched, we flip the gate polarities on both sides and the entire system is mirrored. Another possibility is to put two probes (one FM, one normal) close to each other and measure the voltage difference between them. It is not difficult to show that $\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m}) = 2(\mu_p(\mathbf{m}) - \mu_{nm})$ where μ_{nm} is the voltage measured at the non-magnetic probe.

To summarize, we propose a straightforward potentiometric measurement on a TIPNJ with a FM voltage probe. We worked out quasi-analytical results for the voltage measurements which is also benchmarked with the numerical NEGF simulations. Our analysis predicts gate voltage dependent asymmetrical features - linear dependence of $\Delta\lambda$ in the PP regime and saturation in the NP regime. In a slightly different setup, the angular phase of the signal directly bear out signatures of Klein tunneling in the TI. We have also discussed non-idealities (probe polarization, momentum scattering) that may influence quantitative details seen in the experiment.

ACKNOWLEDGMENTS

We wish to acknowledge the generous support from NSF Grant No. CCF1514219 and NRI. We are also thankful for the discussions with Prof. Supriyo Datta and his student Shehrin Sayed from Purdue University, Dr. An Ping Lee and Saban M Hus from Oak Ridge National Laboratory (ORNL), and Prof. Nitin Samarth from Penn State University. This work used Rivanna high performance computing system at the University of Virginia.

Appendix A: Angular dependence for tilted junction

Here we show that ϱ in Eq. 26 has a phase shift in tilted NP junction compared to the homogeneous PP case. From Eq. 20 we have:

$$\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m}) = (\lambda(\mathbf{m}) - \lambda(-\mathbf{m}))(\mu_s - \mu_d)$$

where $\Delta\lambda(\mathbf{m}) = \lambda(\mathbf{m}) - \lambda(-\mathbf{m})$ can be calculated from Eq. 18:

$$\begin{aligned}\Delta\lambda(\mathbf{m}) &= \frac{\sum_{v_x(\mathbf{k}_t)>0} 2P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{s}(\mathbf{k}_t)t(\mathbf{k}_t)\delta(E_f - E(\mathbf{k}_t))}{\sum_{\mathbf{k}_t}\delta(E_f - E(\mathbf{k}_t))} \\ &= \frac{P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{S}}{\pi} \\ \mathbf{S} &= \sum_{v_x(\mathbf{k}_t)>0}\mathbf{s}(\mathbf{k}_t)t(\mathbf{k}_t)\end{aligned} \tag{A1}$$

Instead of calculating the electron transmission in Eq. 16 explicitly, we rewrite it as the summation of transmission over all forward propagating modes:

$$\begin{aligned}T(E_f) &= \frac{qV_pW}{hv_F}\sum_{v_x(\mathbf{k}_t)>0}\hat{\mathbf{x}}\cdot\hat{\mathbf{v}}_t t(\mathbf{k}_t) = \frac{qV_pW}{hv_F}\hat{\mathbf{x}}\cdot\mathbf{K} \\ \mathbf{K} &= \sum_{v_x(\mathbf{k}_t)>0}\hat{\mathbf{v}}_t t(\mathbf{k}_t)\end{aligned} \tag{A2}$$

where $\hat{\mathbf{v}}_\mathbf{t}$ is the unit vector along the velocity of mode $\mathbf{k}_t$. It is easy to see $\mathbf{S} = \mathbf{K}\times\hat{\mathbf{z}}$ due to the spin-momentum locking. Since $J = \frac{q}{Wh}T(E_f)(\mu_s - \mu_d)$, ϱ in Eq.26 can be expressed as:

$$\begin{aligned}\varrho(\mathbf{m}) &= \frac{\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m})}{qJP_{\mathrm{FM}}} = \frac{h^2v_F}{\pi q^3V_p}\frac{\mathbf{m}\cdot\mathbf{S}}{\hat{\mathbf{x}}\cdot\mathbf{K}} \\ &= \frac{h^2v_F}{\pi q^3V_p}\frac{(\mathbf{z}\times\hat{\mathbf{m}})\cdot\mathbf{K}}{\hat{\mathbf{x}}\cdot\mathbf{K}}\end{aligned} \tag{A3}$$

For a homogeneous PP junction, $\mathbf{K}\propto\hat{\mathbf{x}}$ and $\mathbf{S}\propto-\hat{\mathbf{y}}$. Therefore $\varrho(\mathbf{m})\propto-\sin\theta_m$. For the NP case, only the normal mode can pass through the junction, which means $\mathbf{K}$ is normal to the junction. Therefore $\varrho(\mathbf{m})\propto\sin(\theta_m - \delta_g)$ where δ_g is the angle of the tilted gate.

* yx3ga@virginia.edu; Charles L. Brown Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, 22904 USA.

[1] O. Klein, "Die reflexion von elektronen an einem potentialsprung nach der relativistischen dynamik von dirac," Zeitschrift für Physik **53**, 157–165 (1929).
[2] C. Gutiérrez, L. Brown, C.-J. Kim, J. Park, and A. N. Pasupathy, "Klein tunnelling and electron trapping in nanometre-scale graphene quantum dots," Nature Physics **12**, 1069–1075 (2016).
[3] S. Chen, Z. Han, M. M. Elahi, K. M. Masum. Habib, L. Wang, B. Wen, Y. Gao, T. Taniguchi, K. Watanabe, J. Hone, *et al.*, "Electron optics with pn junctions in ballistic graphene," Science **353**, 1522–1525 (2016).
[4] C. Beenakker, "Colloquium: Andreev reflection and klein tunneling in graphene," Reviews of Modern Physics **80**, 1337 (2008).
[5] N. Stander, B. Huard, and D. Goldhaber-Gordon, "Evidence for klein tunneling in graphene p- n junctions," Physical Review Letters **102**, 026807 (2009).
[6] R. N. Sajjad and A. W. Ghosh, "Manipulating chiral transmission by gate geometry: switching in graphene with transmission gaps," ACS nano **7**, 9808–9813 (2013).
[7] M.-H. Liu, C. Gorini, K. Richter, "Creating and steering highly directional electron beams in graphene," Physical Review Letters **118**, 066801 (2017).
[8] A. W. Barnard, A. Hughes, A. L. Sharpe, K. Watanabe, T. Taniguchi, and D. Goldhaber-Gordon, "Absorptive pinhole collimators for ballistic dirac fermions in graphene," Nature Communications **8** (2017).
[9] Y. Chen, J. Analytis, J.-H. Chu, Z. Liu, S.-K. Mo, X.-L. Qi, H. Zhang, D. Lu, X. Dai, Z. Fang, *et al.*, "Experimental realization of a three-dimensional topological insulator, bi2te3," Science **325**, 178–181 (2009).
[10] X.-L. Qi and S.-C. Zhang, "Topological insulators and superconductors," Reviews of Modern Physics **83**, 1057 (2011).
[11] A. Mellnik, J. Lee, A. Richardella, J. Grab, P. Mintun, M. Fischer, A. Vaezi, A. Manchon, E. Kim, N. Samarth, *et al.*, "Spin-transfer torque generated by a topological insulator," Nature **511**, 449–449 (2014).
[12] J. Han, A. Richardella, S. Siddiqui, J. Finley, N. Samarth, and L. Liu, "Room temperature spin-orbit torque switching induced by a topological insulator," arXiv preprint arXiv:1703.07470 (2017).
[13] M. Jamali, J.-Y. Chen, D. R. Hickey, D. Zhang, Z. Zhao, H. Li, P. Quarterman, Y. Lv, M. Li, K. A. Mkhoyan, *et al.*, "Room-temperature perpendicular magnetization switching through giant spin-orbit torque from sputtered bixse (1-x) topological insulator material," arXiv preprint arXiv:1703.03822 (2017).
[14] K. M. Masum. Habib, R. N. Sajjad, and A. W. Ghosh, "Chiral tunneling of topological states: Towards the efficient generation of spin current using spin-momentum locking," Physical review letters **114**, 176801 (2015).
[15] B. Behin-Aein, D. Datta, S. Salahuddin, and S. Datta, "Proposal for an all-spin logic device with built-in memory," Nature nanotechnology **5**, 266–270 (2010).
[16] S. Hong, V. Diep, S. Datta, and Y. P. Chen, "Modeling potentiometric measurements in topological insulators including parallel channels," Physical Review B **86**, 085131 (2012).
[17] S. Sayed, S. Hong, and S. Datta, "Multi-terminal spin valve on channels with spin-momentum locking," Scientific reports **6** (2016).
[18] C. Li, O. vant Erve, J. Robinson, Y. Liu, L. Li, and B. Jonker, "Electrical detection of charge-current-induced spin polarization due to spin-momentum locking in bi2se3," Nature nanotechnology **9**, 218–224 (2014).

[19] J. S. Lee, A. Richardella, D. R. Hickey, K. A. Mkhoyan, and N. Samarth, "Mapping the chemical potential dependence of current-induced spin polarization in a topological insulator," Physical Review B **92**, 155312 (2015).

[20] J. Tian, I. Miotkowski, S. Hong, and Y. P. Chen, "Electrical injection and detection of spin-polarized currents in topological insulator bi2te2se," Scientific reports **5** (2015).

[21] C. Li, O. vant Erve, Y. Li, L. Li, and B. Jonker, "Electrical detection of the helical spin texture in a p-type topological insulator sb2te3," Scientific Reports **6** (2016).

[22] B. Zhou, Z. Liu, J. Analytis, K. Igarashi, S. Mo, D. Lu, R. Moore, I. Fisher, T. Sasagawa, Z. Shen, *et al.*, "Controlling the carriers of topological insulators by bulk and surface doping," Semiconductor Science and Technology **27**, 124002 (2012).

[23] N. H. Tu, Y. Tanabe, Y. Satake, K. K. Huynh, and K. Tanigaki, "In-plane topological pn junction in the three-dimensional topological insulator bi2- xsbxte3- ysey," Nature Communications **7** (2016).

[24] S. H. Kim, K.-H. Jin, B. W. Kho, B.-G. Park, F. Liu, J. S. Kim, and H. W. Yeom, "Atomically abrupt topological pn junction," ACS nano (2017).

[25] J.G. Checkelsky, Y.S. Hor, M.-H. Liu, D.-X. Qu, R.J. Cava, and N.P. Ong, "Quantum interference in macroscopic crystals of nonmetallic Bi_2Se_3," Physical Review Letters **103**, 246601 (2009).

[26] A.A. Taskin and Y. Ando, "Quantum oscillations in a topological insulator $Bi_{1-x}Sb_x$," Physical Review B **80**, 085303 (2009).

[27] N. P. Butch, K. Kirshenbaum, P. Syers, A. B. Sushkov, G. S. Jenkins, H. D. Drew, and J. Paglione, "Strong surface scattering in ultrahigh-mobility Bi_2Se_3 topological insulator crystals," Physical Review B **81**, 241301 (2010).

[28] S. Jia, H. Ji, E. Climent-Pascual, M.K. Fuccillo, M.E. Charles, J. Xiong, N.P. Ong, and R.J. Cava, "Low-carrier-concentration crystals of the topological insulator Bi_2Te_2Se," Physical Review B **84**, 235206 (2011).

[29] S. Datta, *Electronic transport in mesoscopic systems* (Cambridge university press, 1997).

[30] A. Ghosh, *Nanoelectronics: A Molecular View* (World Scientific, 2016).

[31] F. Xiu, L. He, Y. Wang, L. Cheng, L.-T. Chang, M. Lang, G. Huang, X. Kou, Y. Zhou, X. Jiang, *et al.*, "Manipulating surface states in topological insulator nanoribbons," Nature nanotechnology **6**, 216–221 (2011).

[32] V. V. Cheianov, V. Fal'ko, and B. Altshuler, "The focusing of electron flow and a veselago lens in graphene pn junctions," Science **315**, 1252–1255 (2007).