New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
29 March 2017

REVISED
10 August 2017

ACCEPTED FOR PUBLICATION
25 September 2017

PUBLISHED
24 October 2017

PAPER

The focusing effect of electron flow and negative refraction in three-dimensional topological insulators

Kai-Tong Wang[1], **Yanxia Xing**[2,5], **King Tai Cheung**[3], **Jian Wang**[3], **Hui Pan**[4] and **Hong-Kang Zhao**[1,5]

[1] School of Physics, Beijing Institute of Technology, Beijing 100081, People's Republic of China
[2] Beijing Key Laboratory of Nanophotonics and Ultrafine Optoelectronic Systems, School of Physics, Beijing Institute of Technology, Beijing 100081, People's Republic of China
[3] Department of Physics and the Center of Theoretical and Computational Physics, The University of Hong Kong, Pokfulam Road, Hong Kong, People's Republic of China
[4] College of Physics, Beihang University, Beijing 100191, People's Republic of China
[5] Authors to whom any correspondence should be addressed.

E-mail: xingyanxia@bit.edu.cn **and** zhaohonk@bit.edu.cn

Keywords: nonequilibrium transport, topological insulator, electron focusing, local response, Green's function

Abstract

We numerically study the focusing effect induced by a single p–n junction in three-dimensional topological insulators (3D TIs). It is found that, for either surface states or bulk states of 3D TIs, the corresponding electrons injected from the n/p region can be perfectly focused at the symmetric position in the p/n region. These results suggest that the focusing effect is a general phenomenon in materials which can be described by massless or massive Dirac equations. We also find that the focusing effect is robust against moderate random disorders. In the presence of external magnetic fields, the focusing effect remains good, but the position of the focus point oscillates periodically due to the finite size effect.

1. Introduction

In the 1960s, Veselago theoretically predicted the existence of negative refractive index materials, i.e., left-handed materials [1, 2]. After about 30 years, the first artificial left-handed material was experimentally verified [3]. In general, electromagnetic negative refraction can only be realized in artificial metamaterials with negative ϵ and μ, [4] where ϵ and μ are the electric permittivity and the magnetic permeability, respectively. When concerning the wave vector and the group velocity, the optical rays and the electron flow (the electron's de Broglie wave) are similar. Hence the negative refraction would be achieved in massless Dirac fermion materials, such as graphene. In such a material, the negative refraction is directly related to the perfect Veselago lens [5] and the Klein paradox [6].

The existence of negative refraction in massless Dirac material is natural. The electrons and holes in massless Dirac material are conjugately linked, and the chiralities (or dispersions) in the conduction band and valence band are opposite. Therefore, the potential barrier induced by a p–n junction (PNJ) in massless Dirac materials is highly transparent for the charge carriers [7]. As a result, the electron flow would be negatively refracted and symmetrically focused by the straight interface of the PNJ in the linear dispersion region [8]. Beyond the linear region, the statement on Dirac fermion fall through, but the focusing effect still exists [9]. It means the negative refraction is not limited to two-dimensional (2D) massless Dirac materials. In fact, as shown in figure 1, when electrons with momentum $(k_x,\ k_{\|})$ and velocity $(v_x,\ v_{\|})$ penetrate through the PNJ and become holes with momentum $(-k_x,\ k_{\|})$, due to the opposite dispersion for electrons and holes, the velocity of holes becomes $(v_x,\ -v_{\|})$. Then the negative refraction is formed. As a result, the electron flow is focused by the straight interface induced by the PNJ. Here, '$\|$' denotes the direction along y for a 2D system or y–z plane for three-dimensional systems. So, there are two essential conditions to the focusing effect of the electron flow. One is the opposite dispersions in the conduction band and valence band, the other is the nearly transparent PNJ. In principle, besides massless Dirac fermions, [10, 11] all gapless semi-metals and topological materials [12, 13] described by the quadratic massive Dirac equation in 2D or beyond 2D, such as the 3D topological insulator (TI), are expected

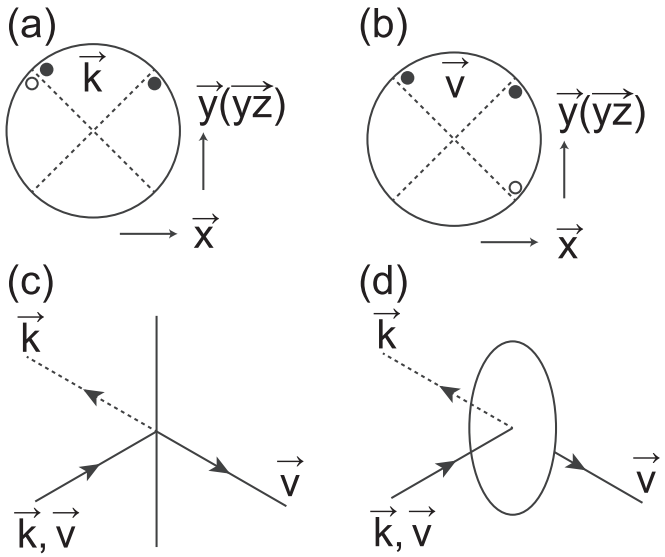

Figure 1. An electron with momentum (k_x, $k_{\|}$) and velocity (v_x, $v_{\|}$) is injected from the left *n* region and scattered by the one-dimensional PNJ (panel (c)) or two-dimensional PNJ (panel (d)) to the right *p* region as a hole (hollow black dots) with momentum ($-k_x$, $k_{\|}$) and velocity (v_x, $-v_{\|}$) due to the opposite dispersions for electrons and holes, which leads to the focusing effect of electrons flow. The corresponding momentum and velocity of electrons (solid circles) and holes (open circles) is shown in the panel (a) and panel (b), respectively.

to have the same effect. Considering the helical resolved characteristics of the TI materials, the focusing effect in TIs can have great potential in the applications of helicity-based electronic optics [12].

In the past years, due to the extraordinary band structure and huge potential in fabricating future devices, TIs have attracted great attentions in condensed matter physics [14–24]. For a 3D TI such as $(Bi,\ Sb)_2Te_3$, [19, 20] electrons on the conducting surface are massless Dirac fermions depicted by a single Dirac cone. On the other hand, when the Fermi energy is beyond the band gap, the TI is similar as conventional semiconductors because of the separated conduction band and valence band. What about the focusing effect of electron flow in 3D TIs considering these different states of matter? Can we find the focusing effect in the deep conduction/valence band?

To answer these questions, we have constructed the PNJ in an infinite 3D TI ribbon as shown in figure 2(a). With the aid of nonequilibrium Green's function (NEGF), we study the local conductance response to the nonequilibrium electron injection in 3D TIs with single PNJ. It is found that when the incident energy is in the bulk gap, the transport processes are dominated by the surface states, and the focusing effect arises only on the surfaces. As shown in figure 2(a), on each side surface, electrons flow injected (the blue points) from the *n* region (blue region) can be focused in the *p* region (red region) in the symmetric position (the red points). It is natural since the surface states of a 3D TI satisfy the 2D massless Dirac equation. When the incident energy is beyond the bulk energy gap, the TI resembles conventional semiconductors. However, because of the conjugated interconnection between the conduction band and valence band, the focusing effect in the bulk is even better. In this case, the electron flow incident from any site with the position of $(-x,\ y,\ z)$ in the *n* region will be focused at position $(x,\ y,\ z)$ in the *p* region. Although supported by both surface and bulk states, the focusing effect can not be observed when these two type of states are mixed near the energy band edges, since surface states and bulk states have different dispersion relations. Furthermore, we have also studied the influence of random scattering and the weak external magnetic field B_z on the focusing effect. It is found that the focusing effect is immune to random disorders. In the presence of weak perpendicular magnetic field B_z, the focus point is deviated by the lateral Lorentz force, but the focusing effect retains well. Owing to the finite size of the scattering region, with the increasing B_z, the position of the focus oscillates periodically with the period of $\Delta B \approx \frac{h/e}{S}$, where S is the area of the central *p* region.

The paper is organized as follows. In section 2, from the low energy effective model, we present the system Hamiltonian in real space using tight-binding technique. Then, both the local partial density and the local conductance describing the local response to the nonequilibrium source, i.e., the incident electron flow, are derived. Section 3 is the numerical results and some discussions. Finally, a summary of our work is presented in section 4.

2. Model and formalism

Through $k \cdot p$ perturbation, the low-energy effective Hamiltonian of 3D TIs can be expanded in the Hilbert space composed of four low-lying states at the Γ point, i.e., $|P1_{z^+,\uparrow}\rangle$, $|P2_{z^-,\uparrow}\rangle$, $|P1_{z^+,\downarrow}\rangle$ and $|P2_{z^-,\downarrow}\rangle$. Correspondingly, the Hamiltonian of infinite 3D TI is written in the following form [15, 20, 25]:

$$H_0(k) = \epsilon_k + M_k\sigma_0\tau_z + A_\perp k_z\sigma_z\tau_x + A_\parallel(k_x\sigma_x + k_y\sigma_y)\tau_x, \tag{1}$$

where σ_α and τ_α represent the real spin (↑ and ↓) and pseudo-spin (signing the orbital $|P1_{z^+}\rangle$ and $|P2_{z^-}\rangle$) with $\alpha = x, y, z$. σ_0 and τ_0 are 2×2 unitary matrices. $\epsilon_k = C_0 + C_\perp k_z^2 + C_\parallel(k_x^2 + k_y^2)$, $M_k = -D_0 + D_\perp k_z^2 + D_\parallel(k_x^2 + k_y^2)$. Here, we set $\epsilon_k = 0$ since it shifts the Dirac point and does not change the topological structure of the Hamiltonian. To investigate the spacial focusing effect, the Hamiltonian expressed in real space is needed. Replacing $k_{x,y,z}$ by $-\mathrm{i}\nabla_{x,y,z}$, we get the 3D effective tight-binding Hamiltonian in a square lattice as follows [26]:

$$H_0 = \sum_{\mathbf{i}} d_{\mathbf{i}}^\dagger H_{\mathbf{i}} d_{\mathbf{i}} + \sum_{\mathbf{i},\alpha} d_{\mathbf{i}}^\dagger H_\alpha d_{\mathbf{i}+\mathbf{a}_\alpha} + \text{h.c.} \tag{2}$$

with

$$H_{\mathbf{i}} = \epsilon_{\mathbf{i}}\sigma_0\tau_0 + \left(D_0 + 2\sum_\alpha \frac{D_\alpha}{a^2}\right)\sigma_0\tau_z$$
$$H_\alpha = \left[-\frac{D_\alpha}{a^2}\sigma_0\tau_z - \mathrm{i}\frac{A_\alpha}{2a}\sigma_\alpha\tau_x\right]\mathrm{e}^{\mathrm{i}\phi_{\mathbf{i},\mathbf{i}+a_\alpha}}$$

where $\alpha = x, y, z$. $d_{\mathbf{i}} = [d_{\mathbf{i},1_z^+,\uparrow}, d_{\mathbf{i},2_z^-,\uparrow}, d_{\mathbf{i},1_z^+,\downarrow}, d_{\mathbf{i},2_z^-,\downarrow}]$ denotes the four low-lying states at the Γ point, and $\epsilon_{\mathbf{i}}$ is the onsite energy at each lattice site. Here, $\mathbf{i} = [\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z]$ is used to indicate the discrete sites of the square lattice with lattice constant a. Considering the perpendicular magnetic field B_z, an extra phase $\phi_{\mathbf{i},\mathbf{i}+a_\alpha} = \frac{e}{\hbar}\int_{\mathbf{i}}^{\mathbf{i}+a_\alpha} \mathbf{A}\cdot\mathbf{dl}$ is induced [27, 28] by the magnetic vector potential $\mathbf{A}$. In the Coulomb gauge, the vector potential is set as $\mathbf{A} = [-By, 0, 0]$ and the magnetic flux at each lattice is then $\Phi_0 = B_z a^2$. For an infinite nanoribbon shown in figure 2(a), y and z dimensions are finite and $x \in [-\infty, \infty]$. The incident electrons are free in the x-direction if the PNJ is absent.

In the presence of the sharp PNJ induced by a step potential $U(x) = E_0[\theta(x) - \theta(-x)]$, the system is composed of a semi-infinite electron-like n region ($x < 0$, the blue region in figure 2(a) where $\epsilon_{\mathbf{i}} = -E_0$) and a semi-infinite hole-like p region ($x > 0$, the red region in figure 2(a) where $\epsilon_{\mathbf{i}} = E_0$). Because the semi-infinite p region and n region are ideally periodic, the incident electrons can only be scattered by the straight interface (the black interface in figure 2(a)), i.e., the sharp PNJ located at $x = 0$. Here, electrons are locally injected through the source terminal H_s and detected through the drain terminal H_d. The total Hamiltonian including the source and drain terminals is then expressed as

$$H = H_0 + H_s + H_d + H_{T,s} + H_{T,d} \tag{3}$$

with

$$H_{s/d} = \sum_k \epsilon_{s/d,k} c_{s/d,k}^\dagger c_{s/d,k}$$
$$H_{T,s/d} = \sum_{k,\mathbf{i}_s} t_{s/d}[d_{\mathbf{i}_{s/d}}^\dagger c_{s/d,k} + \text{h.c.}],$$

where H_0 denotes the infinite nanoribbon with the sharp PNJ. $H_{s/d}$ denotes the source or detecting terminal used to inject or detect electron flow. Phenomenologically, $H_{s/d}$ is expressed in the momentum space. $H_{T,s/d}$ is the coupling between the source or detection electrode and the infinite ribbon. Here, we assume the electron flow is locally injected at the site $\mathbf{i}_s$ in the n region and detected at site $\mathbf{i}_d$ in the p region. For convenience, we define the central scattering region (the solid box in figure 2(a)) enclosing the injecting site and detecting site. Concerning the central scattering Hamiltonian H_c, the total Hamiltonian can also be written in the following form:

$$H = H_c + \sum_{\beta=s,d,l,r} (H_\beta + H_{T,\beta}). \tag{4}$$

Equation (4) describes a typical open system. Here, we can treat H_β as open boundaries, denoting the source electrode, the detection electrode and the left and right semi-infinite lead, respectively. $H_{T,\beta}$ is the coupling between the central scattering region and open boundaries. Obviously, $H_0 = H_c + H_l + H_r + H_T$, and

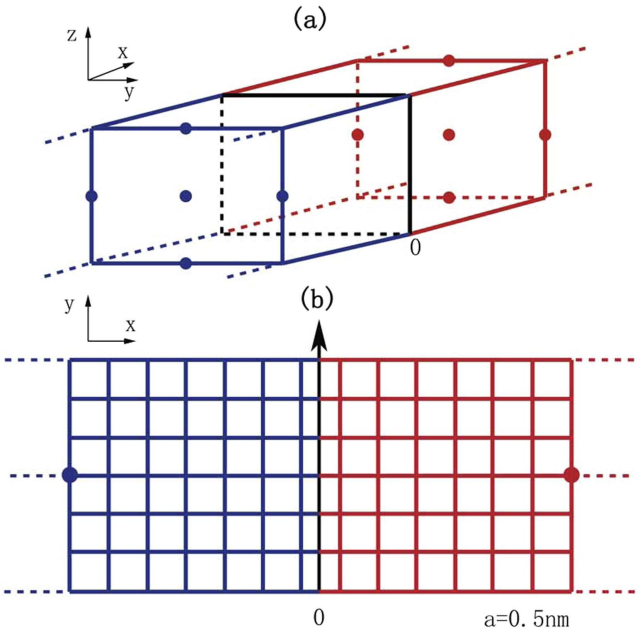

Figure 2. Panel (a): schematic diagram of a two-terminal open system composed of semi-infinite n region (the blue region) and semi-infinite p region (the red region) with a sharp PNJ located at $x = 0$ (the black square). For the bulk states, the incident electron flow is (the central blue dot) scattered by the PNJ and symmetrically focused in the p region (the central red dot). For the surface states, the focusing effect occurs only on the surfaces, the injected electron flow and the focused electron flow are signed by the blue points and red points on the surfaces. Panel (b): the focusing effect in the side surface.

$$\begin{aligned} H_{l/r} &= \sum_{\mathrm{i}\in l/r} d_{\mathrm{i}}^{\dagger} H_{\mathrm{i}} d_{\mathrm{i}} + \sum_{\langle \mathrm{i},\mathrm{j}\rangle \in l/r} d_{\mathrm{i}}^{\dagger} H_{\mathrm{i},\mathrm{j}} d_{\mathrm{j}} \\ H_T &= [H_{l,c} + H_{c,r}] + \mathrm{h.c.} \\ H_{l,c} &= \sum_{\langle \mathrm{i}\in l,\mathrm{j}\in c\rangle} H_x d_{\mathrm{i}}^{\dagger} d_{\mathrm{j}}, \quad H_{c,r} = \sum_{\langle \mathrm{i}\in c,\mathrm{j}\in r\rangle} H_x d_{\mathrm{i}}^{\dagger} d_{\mathrm{j}}. \end{aligned} \tag{5}$$

Next, with the help of the NEGF, the response signals, i.e., the local density ρ_{i} in the scattering region are calculated as follows [29, 30].

$$\rho_{\mathrm{i}} = -\mathrm{i}\int \mathrm{d}E \mathbf{G}^{<}_{\mathbf{ii}}(E), \tag{6}$$

where $\mathbf{G}^{<}_{\mathbf{ii}}$ is the diagonal element of the lesser Green's function. Using Keldysh equation, [31] the lesser Green's function can be written as

$$\mathbf{G}^{<} = \sum_{\beta=l,r,s,d} \mathbf{G}^{r} \mathrm{i}\mathbf{\Gamma}_{\beta} f_{\beta} \mathbf{G}^{a}. \tag{7}$$

Here, $\mathbf{G}^{r}$ and $\mathbf{G}^{a}$ are retarded and advanced Green's functions of the scattering region, respectively. $\mathbf{G}^{r} = \mathbf{G}^{a,\dagger} = [E - H_c - \sum_{\beta}\mathbf{\Sigma}^{r}_{\beta}]^{-1}$, and f_{β} is the Fermi distribution function of the terminal-β. In the nonequilibrium system, the Fermi energy of the terminal-β is shifted by the external bias V_{β}, and $f_{\beta}(E) = f_0(E - eV_{\beta})$, where f_0 is the Fermi distribution function with zero bias. The linewidth function is $\mathbf{\Gamma}_{\beta} = \mathrm{i}(\mathbf{\Sigma}^{r}_{\beta} - \mathbf{\Sigma}^{a}_{\beta})$ with $\mathbf{\Sigma}^{r}_{\beta}$ the retarded self energy induced by the lead-β. For the left and right seimi-infinite leads, $\mathbf{\Sigma}^{r}_{l/r} = \mathbf{H}_{c,l/r}\mathbf{g}^{r}_{l/r}\mathbf{H}_{l/r,c}$, where $\mathbf{H}_{c,l/r}$ is the coupling from the central region to the left or right lead, and $\mathbf{g}^{r}_{l/r}$ is the surface Green's function of the semi-infinite lead, which can be calculated iteratively using transfer matrix [32, 33] or Bloch eigenvector [34, 35]. The source and detection terminals are expressed in the momentum space. In the wide band limit, the self energy of the source or the detecting lead is $\mathbf{\Sigma}^{r}_{s/d} = -\mathrm{i}\pi\rho_0 t^2_{s/d}$. Comparing with injecting terminal, the influence of the detection terminal is much weaker, i.e. $t_d \ll t_s$. In this case, we can neglect $\Sigma^{r/a}_{d}$, then $\mathbf{G}^{r/a} = [E - H_c - \sum_{\alpha}\mathbf{\Sigma}^{r/a}_{\alpha}]^{-1}$ with $\alpha = l, r, s$.

In our calculations, the electron flow is injected from the source terminal, and the left and right semi-infinite leads are all the drain terminals. So, we set $V_s = V$ and $V_l = V_r = 0$. Finally, the lesser Green's function can be divided into equilibrium and nonequlibrium terms, i.e., $\mathbf{G}^{<} = \mathbf{G}^{<}_{0} + \mathbf{G}^{<}_{V}$ with

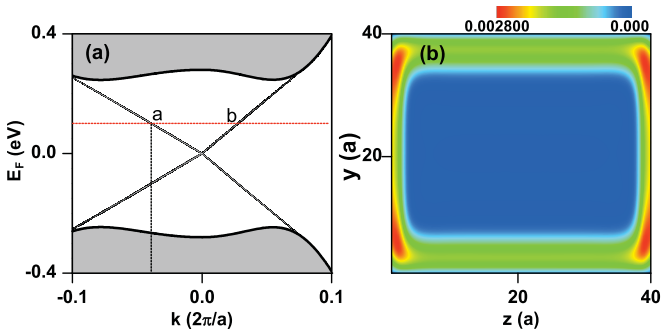

Figure 3. Panel (a): local density of states at the low energy regime on an infinite $x-y$ surface of a semi-infinite 3D TI along the directions of $\overrightarrow{\Gamma K}$ and $\overrightarrow{\Gamma M}$. The black thick lines are the bulk band edges. The red dotted lines denotes the fixed low energy $E_F = 0.1$ eV and corresponding $k_x = 0.04$, with which the local density of states in the finite $(40a \times 40a)$ $y-z$ cross section of the infinite ribbon are plotted in panel (b).

$$\mathbf{G}_0^< = \sum_\beta \mathbf{G}^r i\mathbf{\Gamma}_\beta \mathbf{G}^a f_0$$
$$\mathbf{G}_V^< = \mathbf{G}^r i\mathbf{\Gamma}_s \mathbf{G}^a (f_s - f_0). \tag{8}$$

Here, only the nonequlibrium term contributes to the response signals. It means

$$\rho_\mathbf{i} = \int dE [\mathbf{G}^r \mathbf{\Gamma}_s \mathbf{G}^a]_{\mathbf{ii}} (f_s - f_0).$$

At zero temperature and the linear bias limit, $\rho_\mathbf{i} = \mathbf{G}^r_{\mathbf{i},\mathbf{i}_s} \mathbf{\Gamma}_s \mathbf{G}^a_{\mathbf{i}_s,\mathbf{i}}\, eV_s$. Then, we can define the local partial density

$$\delta\rho_\mathbf{i}/\delta(eV_s) = \mathbf{G}^r_{\mathbf{i},\mathbf{i}_s} \mathbf{\Gamma}_s \mathbf{G}^a_{\mathbf{i}_s,\mathbf{i}}. \tag{9}$$

On the other hand, we can also calculate the local conductance, which is defined as $\sigma_\mathbf{i} = \partial J_\mathbf{i}/\partial V_s$, where $J_\mathbf{i}$ is the current flowing to the detection terminal that is located at site $\mathbf{i}$. According to the Landauer–Büttiker formalism,

$$J_\mathbf{i} = \frac{e}{h} \mathbf{\Gamma}_d \mathbf{G}^r_{\mathbf{i},\mathbf{i}_s} \mathbf{\Gamma}_s \mathbf{G}^a_{\mathbf{i}_s,\mathbf{i}} (eV_d - eV_s).$$

Since we have set $V_d = 0$, the local conductance $\sigma_\mathbf{i}$ is then expressed as

$$\sigma_\mathbf{i} = \frac{e^2}{h} \mathbf{\Gamma}_d \mathbf{G}^r_{\mathbf{i},\mathbf{i}_s} \mathbf{\Gamma}_s \mathbf{G}^a_{\mathbf{i}_s,\mathbf{i}}. \tag{10}$$

Here, $\mathbf{\Gamma}_d = 2\pi\rho_0 t_d^2$ is a constant. Therefore, the local conductance is equivalent to the partial density, i.e., $\sigma_\mathbf{i} \propto \delta\rho_\mathbf{i}/\delta(eV_s)$. In the following numerical calculations, only the local partial density is considered.

3. Numerical results and discussion

In the numerical calculations, parameters of the 3D TI are set as $D_0 = 0.28$ eV, [25] $D_\perp = 10$ eV Å^2, $D_\parallel = 56.6$ eV Å^2, $\epsilon_k = 0$, $A_\perp = 2.2$ eV Å, $A_\parallel = 4.1$ eV Å, [20] the lattice constant $a = 5$ Å. Here we set $E_F = 0$, so the (kinetic) energy (relative to the energy of Γ point) of electrons and holes are $E_n = E_0$ and $E_p = -E_0$, respectively.

3.1. Focusing effect in the linear dispersion regime

For the 3D TI, due to the band inversion near the Γ point, the nontrivial energy gap is induced. When $A_\parallel$ is small, the energy gap is roughly determined by D_0. In our model, the global bulk energy gap is in the interval of $[-0.245, 0.245$ eV$]$. When the incident energy of electrons is within the bulk energy gap, all the bulk states are forbidden, and only the massless Dirac fermions on the surfaces are permitted. In figure 3, we first study the equilibrium density of states ρ_0 in this energy regime. According to fluctuation–dissipation theorem, $\rho_0 = \frac{i}{2\pi}[G_s^r - G_s^{r,\dagger}]$, where G_s^r is the surface Green's function of the semi-infinite 3D TI for fixed Fermi energy. In figure 3(a) we plot the distribution of $\rho_0(k, E_F)$ of the infinite surface along the x–y plane. The infinite surface is denoted by the 2D momentum $k_{x,y}$. In order to show the results intuitively, the momentum is set along the lines of $\overrightarrow{\Gamma M}$ $(k_x = k_y = k)$ and $\overrightarrow{\Gamma K}$ $(k_x = k, k_y = 0)$, where $\Gamma = (0, 0)$, $M = (-0.5, 0)$ and $K = (0.5, 0.5)$ are all the high symmetry points in the momentum space. From figure 3(a), we can clearly see the linear dispersion of the massless Dirac cone within the bulk energy gap that is bounded by the thick black lines in figure 3(a). Fixing $E_F = 0.1$ eV, the corresponding momentum k_x along the path of $\overrightarrow{\Gamma M}$ is determined (the

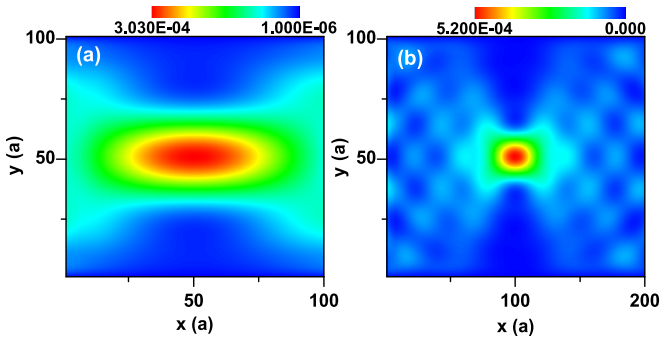

Figure 4. The focusing effect dominated by surface states for $E_0 = 0.1$ eV and 0.2 eV. Panel (a): distribution of the local partial density at the bottom layers in the p region while the electron flow is injected from the position $(-49.5a,\ 49.5a,\ 0.5a)$ in the n region. The potential height of PNJ $E_0 = 0.1$ eV. Panel (b): $E_0 = 0.2$ eV, the electron flow is injected from $(-99.5a,\ 49.5a,\ 0.5a)$ and focused at $(99.5a,\ 49.5a,\ 0.5a)$. The width and height of the ribbon are $W_y = 100a$ and $W_z = 6a$, respectively.

black dotted line in figure 3(a)). With these E_F and k_x, the local density of states in the finite cross section $(40a \times 40a)$ of the infinite ribbon is plotted in figure 3(b). From figure 3(b), we can clearly see the boundary states (the red region along the boundary) in the $y-z$ section, depicting the surface states in the infinite ribbon.

From figure 3, we have confirmed the surface states in both momentum space and real space when the incident energy is within the bulk energy gap. In the following, we will study the focusing effect induced by these surface states. In figure 4, we focus on the focusing effect in surfaces of the 3D TI nanoribbon with a straight PNJ potential within bulk energy gap. In figures 4(a) and (b), we plot the distribution of local partial density $\delta\rho/\delta eV$ at the bottom layer in the p region $(x > 0)$ for $E_0 = 0.1$ eV and 0.2 eV, respectively. The results in other surfaces are similar (not shown). The width and height of the ribbon are set as $W_y = 100a$, $W_z = 6a$, respectively. The local partial density on the bottom surface is the sum of the two layers at the lowest bottom, i.e., the layers at $z = 0.5a$ and $z = 1.5a$. In figure 4(a), the electron flow is injected from the position $(-49.5a,\ 49.5a,\ 0.5a)$ in the n region, and focused at the position $(49.5a,\ 49.5a,\ 0.5a)$ in the p region. While in the figure 4(b), the electron flow is injected from $(-99.5a,\ 49.5a,\ 0.5a)$ and focused at $(99.5a,\ 49.5a,\ 0.5a)$. It can be seen as long as the energy is within the energy gap, the electron flow can be perfectly focused at the symmetric position in the p region, no matter where it is injected in the n region, as discussed in figure 1. It is similar to graphene [8, 9] systems. Besides, we also find that the focusing effect for the higher energy $(E_0 = 0.2$ eV) is better, which is totally different from the focusing effect in graphene systems, for which the focusing effect is worse for the higher energy [9]. In addition, due to the extra scattering induced by the boundaries of the nanoribbon, there are regular interference patterns when the scattering region is long, as shown in figure 4(b).

In the linear dispersion regime, the focusing effect is dominated by the surface states. Therefore, the focusing effect cannot happen in the deep positions of the ribbon. In figure 5, considering a nanoribbon with size $W_y = 100a$ and $W_z = 13a$, the distribution of local partial density $\delta\rho/\delta eV$ along the diagonal line of the p region, i.e., $(x = y) \in [0,\ 100a]$, is plotted. Assuming the electron flow is injected from the middle layer $(z = 6.5a$, red lines) and bottom layer $(z = 0.5a$, black and blue lines), we plot $\delta\rho/\delta eV$ at the bottom, middle and top layers in the p region. Here, $\delta\rho/\delta eV$ of the middle layers is the sum of the five middle layers, i.e., the layers located at $z = 4.5$–$8.5a$. $\delta\rho/\delta eV$ of the bottom (top) layers is the sum of the four lowest (highest) layers, i.e., the layers located at $z = 0.5$–$3.5a$ (9.5–12.5a). It is found when injecting electron flow from middle layer, the local response in the p region is uniformly small in all layers, which means no focusing effect happens when the electron source is located deep inside the bulk. On the other hand, when injecting electron flow from the bottom layer, $\delta\rho/\delta eV$ increases abruptly in the center of the bottom layers (the black line in figure 5), and becomes very small in the middle and top layers (the blue lines). In a word, in the low energy linear dispersion regime, the focusing effect of electron flows is dominated by the surface states and arises only in the surfaces of 3D TI ribbons.

3.2. Focusing effect in the high energy regime

When the Fermi energy is beyond the bulk energy gap, incoming carriers are no longer described by linear Dirac equations. How about the focusing effect in this case? In figure 6(a), we show the distribution of equilibrium density of states $\rho_0(k,\ E_F)$ in the whole energy regime for an infinite surface of a semi-infinite 3D TI. Both discrete surface states (gray lines) and continuous bulk states (gray region bounded by black band edges) appear in figure 6(a). Two cases are considered: Fermi energy is near the band edge $(E_F = 0.33$ eV) and deep in the conduction band $(E_F = 0.9$ eV) (see the red dotted lines in figure 6(a)). With these Fermi energies, the local

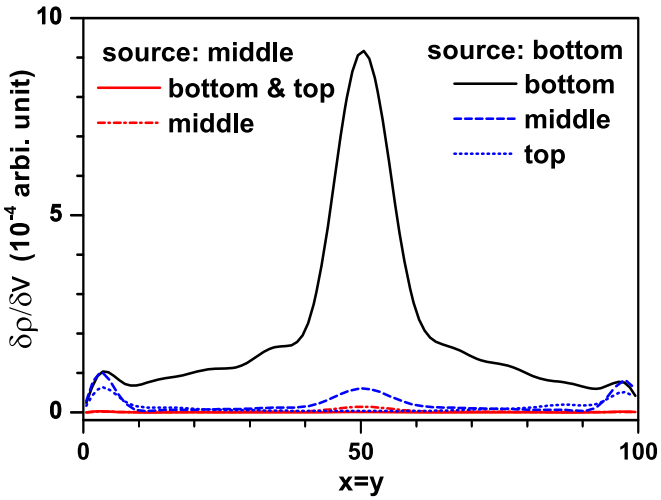

Figure 5. The distribution of local partial density along the diagonal line of the p region, i.e., $(x = y) \in [0,\ 100a]$ for injecting energy $E_0 = 0.2$ eV (surface states dominated focusing). The electron flow is injected from the middle layer ($z = 6.5a$) or bottom layer ($z = 0.5a$). The width and height of the ribbon are $W_y = 100a$ and $W_z = 13a$, respectively.

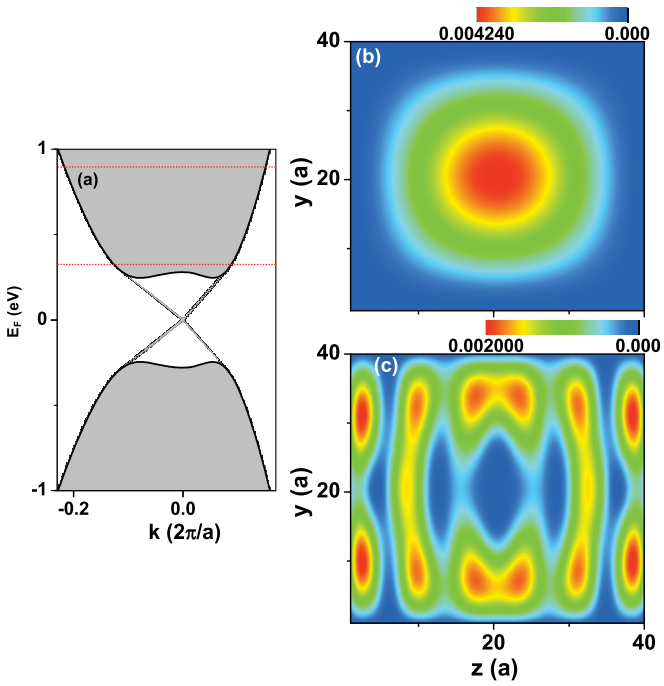

Figure 6. Local density of states at high energy regime for an infinite $x-y$ surface of a semi-infinite 3D TI. The momentum k is along the directions of Γ–M and Γ–K. The red dotted lines denote the fixed Fermi energies $E_F = 0.33$ eV and 0.9 eV, which correspond to the puddle regime of surface and bulk states, and the bulk states dominated regime, respectively. Corresponding to $E_F = 0.9$ eV and 0.33 eV, the local density of states in the finite $(40a \times 40a)$ $y-z$ cross section of the infinite ribbon are plotted in panels (b) and (c), respectively.

density of states in the finite cross section $(40a \times 40a)$ of the infinite ribbon plotted in figures 6(b) and (c). We can see near the band edge ($E_F = 0.33$ eV, panel (c)) the surface states of the infinite ribbon are disturbed. Meanwhile the bulk states have not yet predominated. When E_F is deep in the conduction band ($E_F = 0.9$ eV, panel (b)), the surface states disappear completely. Then, the system is dominated by the bulk states that do not obey the massless Dirac equation. In the following we will study the focusing effect induced by the bulk states.

We first set the PNJ potential $E_0 \gg D_0$. In this case, the bulk states dominate the transport processes. As a result, the focusing effect can occur in bulk as well as the surfaces of the ribbon, which is different from the focusing effect in the linear energy regime. Injecting electrons flow from the bottom layer located at $z = 0.5a$, we plot the local partial density $\delta\rho/\delta eV$ in the center of the p region at every layer signed by its z coordinate in figure 7(a). The height of the nanoribbion is $W_z = 6a$. It can be seen $\delta\rho/\delta eV$ becomes maximum at the bottom

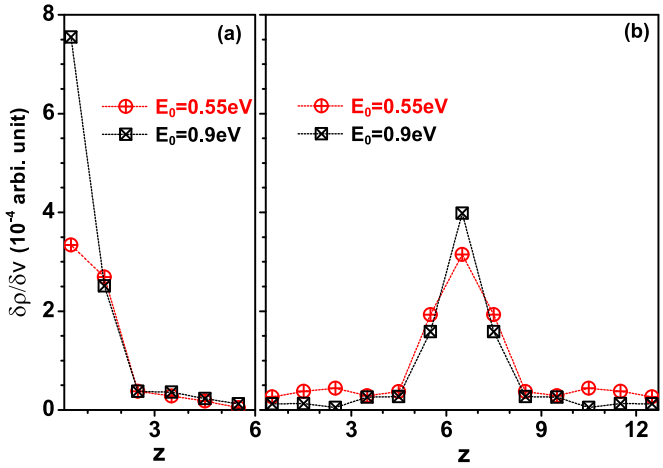

Figure 7. In the high energy regime, $E_0 = 0.55$ eV and 0.9 eV, the $\delta\rho/\delta V$ at the center ($x = 0.5W_x$, $y = 0.5W_y$) of the p region versus z. Panel (a): $W_z = 6a$, the electron flow is injected from the bottom layers ($z = 0.5a$). Panel (b): $W_z = 13a$, the electron flow is injected from the middle layer ($z = 6.5a$).

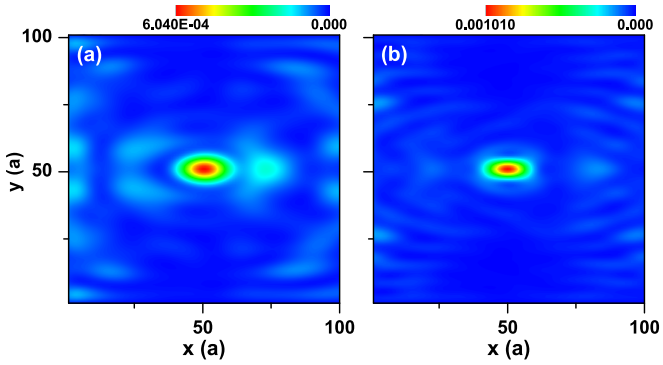

Figure 8. The focusing effect dominated by bulk states. Panel (a): the distribution of local partial density in the p region of the infinite ribbon with a sharp PNJ for $E_0 = 0.55$ eV. Panel (b): $E_0 = 0.9$ eV. The other parameters: the width $W_y = 100a$, the height $W_z = 6a$.

x–y plane with $z = 0.5a$ since the electron source also locates at the bottom layer. When deviating from the bottom layer, $\delta\rho/\delta eV$ reduces abruptly to nearly zero. Next, we set $W_z = 13a$ and inject the electron flow from the middle layer ($z = 6.5a$). $\delta\rho/\delta eV$ in the center of the p region at every layer is plotted in figure 7(b). We can see $\delta\rho/\delta eV$ becomes maximum at the middle bulk layer ($z = 6.5a$) and decreases abruptly at other layers. In other words, wherever injected from the n region, the electron flow can always be perfectly focused to the symmetric site in the p region. It means that besides 2D massless Dirac fermions, conventional semi-metals can also produce the focusing effect.

In general, it is difficult to embed the source lead deep into the system. So, in the following, the source terminal is assumed to be located in the bottom surface.Indeed the focusing effect occurs on the bottom surface as well. However, we must keep in mind that this focusing is dominated by bulk states, not by surface states. In figure 8, we plot the distribution of local partial density $\delta\rho/\delta eV$ in the p region for $E_0 = 0.55$ eV and 0.9 eV that are all deep inside the bulk energy band. Here, the local partial density in the bottom surface is the sum of the two lowest layers ($z = 0.5a$ and $1.5a$). From figure 8, we can see the perfect focusing effect in the high energy regime. The higher the PNJ potential, the better the focusing effect. Different from figure 4, this focusing effect is induced by the conventional bulk states with quasi-quadratic dispersion relations.

Up to now, we have shown that both the surface states (with linear dispersion) and the bulk states (with quadratic dispersion) can produce focusing effect. We wonder if the focusing effect is going to happen when the surface states and the bulk states are mixed. In figure 9, we plot the distribution of local partial density $\delta\rho/\delta eV$ in the p region for $E_0 = 0.33$ eV. For $E_0 = 0.33$ eV, the Fermi energy is near the band edge, and surface sates and bulk states coexist as shown in figure 6(b). From figure 9, we can see the focusing effect induced by the mixed states is much worse compared to figures 4 or 8 in which pure surface states or bulk states are dominant. It is not

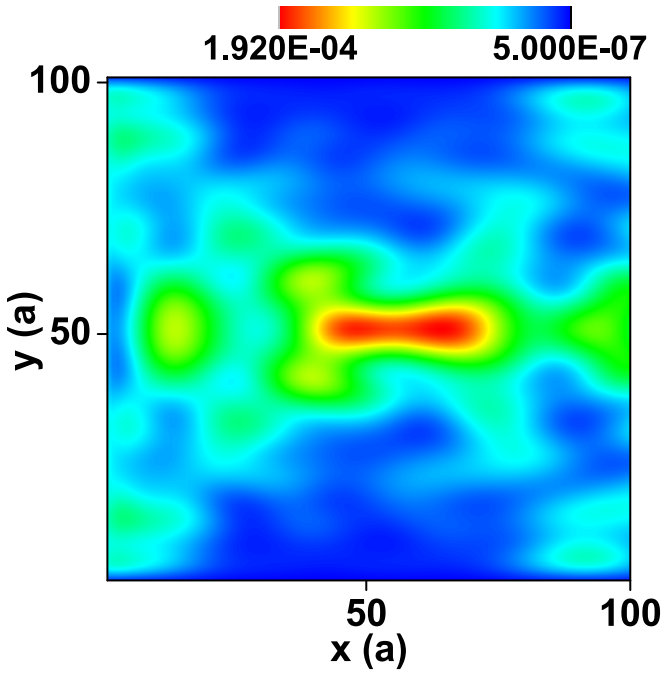

Figure 9. The distribution of the local partial density of the p region in the puddle regime of surface and bulk states for $E_0 = 0.33$ eV. The other parameters are the same as figure 8.

strange because the dispersions of the surface states and bulk states are different. The mixed states can't synchronously penetrate the PNJ through the Klein tunneling. [8] The poor focusing in figure 9 just reveals the different behaviors between the surface states and bulk states.

3.3. Influence of disorders on the focusing effect

In a real device, disorders are inevitable. In this subsection, we will study the influence of disorders. In general, the impurities may appear near the interface due to the preparation of the PNJ. Disorders induce random scattering which is simulated by the random on-site potential [36, 37] $\delta\epsilon_i$ uniformly distributed in the interval $[-w/2,\ w/2]$, where w is the disorder strength. Due to the disorder, the on-site energy now becomes $\epsilon_i + \delta\epsilon_i$. In the numerical calculations, disorders are distributed near the PNJ in the region from $x = -5.5a$ to $x = 5.5a$. The numerical results are averaged over 200 random configurations.

In figures 10 and 11, we study the focusing effect in the presence of random disorders. Both the surface states dominated regime ($E_0 = 0.2$ eV) and bulk states dominated regime ($E_0 = 0.9$ eV) are studied. In figures 10(a) and (b), we plot the distributions of local partial density for $E_0 = 0.2$ eV and $E_0 = 0.9$ eV, respectively. The disorder strength is set as $w \geqslant 2$ eV that is very strong comparing to the PNJ potential E_0. From figure 10, we can see in the presence of strong disorders, focusing effects dominated by either surface states or the bulk states survive successfully. Comparing figure 10(b) with figure 8(b), it can be found in the presence of strong disorders the distribution patterns of $\delta\rho/\delta eV$ are hardly affected. Therefore, the focusing effects of both surface states and bulk states are all immune to random disorders. It means the robust focusing effect is a general phenomenon, and not limited to the massless Dirac fermions, which is promising for related device designs.

Comparing figures 8(b) and 10(b), we can see that although the focusing patterns are kept well, the focusing intensity, i.e., the maximum $\delta\rho/\delta V$, is reduced severely in the presence of strong disorders. In figure 11, we plot the maximum values of the local partial density versus the disorder strength w. Here the maximum value is the sum of the local partial density of the nine sites around the central focus (located at $x = 49.5a$, $y = 49.5a$). From figure 11, we can find for the bulk states dominated focusing ($E_0 = 0.9$ eV), the intensity is maintained when $w < E_0$. When the disorder is strong enough to destroy the PNJ, the maximum $\delta\rho/\delta eV$ decreases rapidly, and the focusing patterns are finally smeared. Correspondingly, the focusing effect induced by the PNJ is out of work. For the surface states dominated focusing ($E_0 = 0.2$ eV), the maximum $\delta\rho/\delta eV$ remains unchanged even at very strong disorder ($w \gg E_0$) because of the topological nature of the surface states. The focusing effect is kept until the disorder is strong enough to destroy the topological surface states. In summary, the surface states dominated focusing effect is coarse but more robust than the bulk states dominated one. In another words, the bulk states dominated focusing effect is finer but frangible compared to the surface states dominated case.

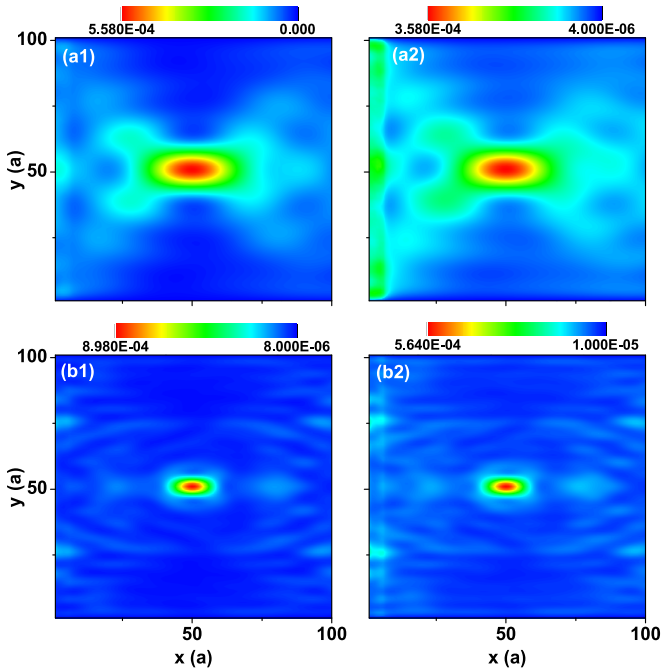

Figure 10. Distributions of the local partial density in the p region of the disordered PNJ. The top and bottom panels are corresponding to the surface states dominated regime, $E_0 = 0.2$ eV for $w = 2$ eV and 4 eV (panels (a1) and (a2)), and bulk states dominated regime, $E_0 = 0.9$ eV for $w = 2$ eV and 3 eV (panels (b1) and (b2)), respectively. The other parameters are the same as figure 8.

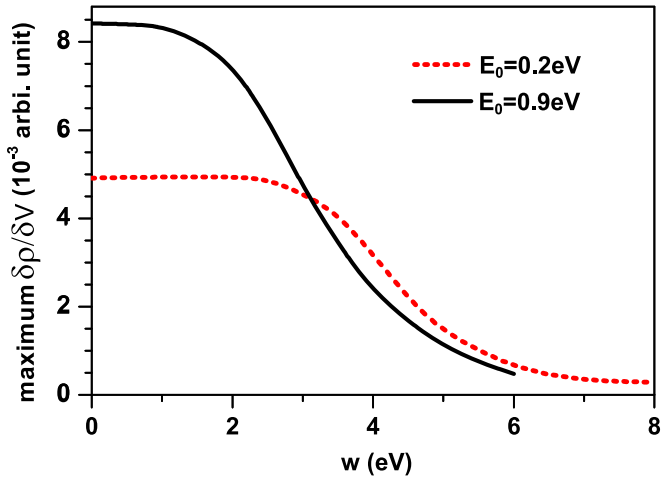

Figure 11. The maximum values of the local partial density in the p region versus the disorder strength w in surface sates dominated regime, $E_0 = 0.2$ eV and bulk states dominated regime, $E_0 = 0.9$ eV.

3.4. Effect of magnetic fields on the focusing effect

Besides random disorders, the focusing effect is disturbed by external magnetic fields. In the following, we will study the influence of the perpendicular magnetic field B_z. In figures 12(a) and (b), we plot the distributions of the local partial density for $E_0 = 0.2$ eV and 0.9 eV, respectively. The magnetic field is set as $B_z = 4T$. It can be seen the focusing pattern is disturbed and the focus point deviates from the central position of the p region. Furthermore, comparing figures 8(b) and 12(b), we can see the focusing intensity is also reduced by B_z. This is because the magnetic vector potential produces additional transverse velocity that breaks the conservation of momentum k_y. As a result, the symmetric focusing process as analyzed in figure 1 falls down and the focusing intensity is then weakened. Although weakened by the magnetic field, the focusing effect still exists in the p region. From figure 12, we can clearly see the deflected focus point.

Let us analyze how the magnetic field affects the focusing effect. In the presence of a magnetic field B_z, the concomitant vector potential induces the anomalous transverse velocity δv_y [38]. In our model, δv_y is positive

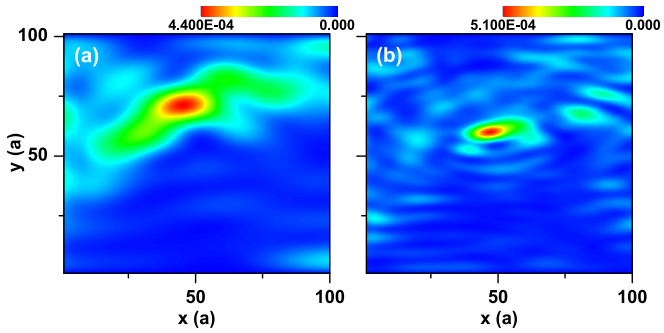

Figure 12. Distributions of the local partial density of the p region in the presence of external magnetic fields. Panel (a): surface states dominated regime, $E_0 = 0.2$ eV. Panel (b): bulk states dominated regime, $E_0 = 0.9$ eV. The magnetic filed $B_z = 4T$.

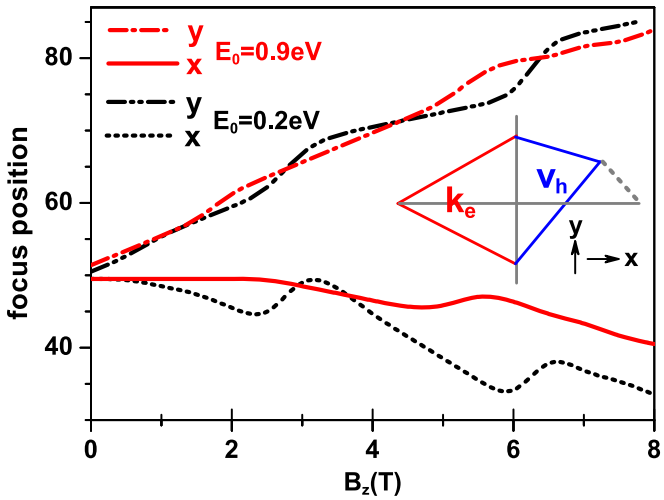

Figure 13. Position coordinates of the focused electron flow versus the magnetic field B_z for $E_0 = 0.2$ eV (the black lines) and 0.9 eV (the red lines), respectively.

(negative) for electrons (holes). Assuming the electrons are injected in the middle of the ribbon and the momentum k_e is symmetrically distributed, the velocity of holes v_h is shifted by the positive δv_y as shown in the inset of figure 13. Then we can expect that the focus point is deflected. It deviates from the center of the p region with negative δx and positive δy (gray dotted line in the inset of figure 13). To track the deflection of the focus point in the presence of magnetic fields, we plot x and y coordinates of the focus position versus B_z in the main panel of figure 13. It can be seen with the increasing of B_z the focus point deviates from the central point ($x = 49.5a$, $y = 49.5a$) of the p region. This conclusion is valid for both surface states ($E_0 = 0.2$ eV) and bulk states ($E_0 = 0.9$ eV). The deflections of the focus point are roughly the same for $E = 0.2$ eV and $E = 0.9$ eV. Besides, due to the finite size of the central scattering region, the track of the focus point oscillates periodically with the oscillating period of $\Delta B \approx \frac{h/e}{W_x W_y}$. The reason is that when the magnetic flux is quantized by h/e (h is the Planck constant and e is the electron charge), the anomalous transverse velocity δv_y becomes maximum and δx or δy is correspondingly maximum. Furthermore, the oscillation is more violent for lower E_0 because the subbands are more discrete.

4. Conclusion

In summary, based on the tight-binding Hamiltonian and NEGF technique, the focusing effect of electron flows in 3D TIs with a single PNJ is systematically studied. It is found that, for either surface states or bulk states of 3D TIs, the electron flow injected from the n/p region can be perfectly focused at the symmetric position in the p/n region of the PNJ. These facts suggest that the focusing effect is a general phenomenon in TIs described by massless or massive Dirac equations. However, focusing patterns are destroyed when the Fermi energy of electron flows is at the band edges, where surface and bulk states are mixed. This exception is attributed to the

incompatible dispersion relations of the surface and bulk states of TIs. We also found the focusing pattern is robust against moderate random disorders, but the focusing intensity are severely reduced at strong disorders. In the presence of a weak magnetic field, the focusing effect remains well, but the position of the focus point oscillates periodically due to the finite size effect. These numerical findings are beneficial to the application of topological materials.

Acknowledgments

This work is supported by grants from National Natural Science Foundation of China (No. 11674024, No. 11374246 and No. 11574019). Jian Wang acknowledges the Research Grant Council (Grant No. 17311116) and the University Grant Council (Contract No. AoE/P-04/08) of the Government of HKSAR.

ORCID iDs

King Tai Cheung https://orcid.org/0000-0002-3706-3828

References

[1] Veselago V G 1968 *Sov. Phys.—Usp.* **10** 509
[2] Ward D W, Nelson K A and Webb K J 2005 *New J. Phys.* **7** 213
[3] Shelby R A, Smith D R and Schultz S 2001 *Science* **292** 77
[4] Smith D R, Pendry J B and Wiltshire M C K 2004 *Science* **305** 788
[5] Pendry J B 2000 *Phys. Rev. Lett.* **85** 3966
[6] Güney D Ö and Meyer D A 2009 *Phys. Rev.* A **79** 063834
[7] Katsnelson M I, Novoselov K S and Geim A K 2006 *Nat. Phys.* **2** 620
[8] Cheianov V V, Fal'ko V and Altshuler B L 2007 *Science* **315** 1252
[9] Xing Y, Wang J and Sun Q-F 2010 *Phys. Rev.* B **81** 165425
[10] Fleury R and Alù A 2014 *Phys. Rev.* B **90** 035138
[11] Péterfalvi C G, Oroszlány L, Lambert C J and Cserti J 2012 *New J. Phys.* **14** 063028
[12] Zhao L, Wang J, Liu J, Xu Y, Gu B-L, Xue Q-K and Duan W 2015 *Phys. Rev.* B **92** 041408
[13] Sessi P *et al* 2016 *Phys. Rev.* B **94** 075137
[14] Hasan M Z and Kane C L 2010 *Rev. Mod. Phys.* **82** 3045
[15] Qi X-L and Zhang S-C 2011 *Rev. Mod. Phys.* **83** 1057
[16] Murakami S, Nagaosa N and Zhang S-C 2003 *Science* **301** 1348
[17] Bernevig B A, Hughes T L and Zhang S-C 2006 *Science* **314** 1757
[18] König M, Wiedmann S, Brüne C, Roth A, Buhmann H, Molenkamp L W, Qi X-L and Zhang S-C 2007 *Science* **318** 766
[19] Chen Y L *et al* 2009 *Science* **325** 178
[20] Zhang H, Liu C-X, Qi X-L, Dai X, Fang Z and Zhang S-C 2009 *Nat. Phys.* **5** 438
[21] Yu R, Zhang W, Zhang H-J, Zhang S-C, Dai X and Fang Z 2010 *Science* **329** 61
[22] Chang C-Z *et al* 2013 *Science* **340** 167
[23] Zhang W, Yu R, Zhang H-J, Dai X and Fang Z 2010 *New J. Phys.* **12** 065013
[24] Shan W-Y, Lu H-Z and Shen S-Q 2010 *New J. Phys.* **12** 043048
[25] Liu C-X, Qi X-L, Zhang H-J, Dai X, Fang Z and Zhang S-C 2010 *Phys. Rev.* B **82** 045122
[26] Zhang L, Zhuang J, Xing Y, Li J, Wang J and Guo H 2014 *Phys. Rev.* B **89** 245107
[27] Chen J-C, Wang J and Sun Q-F 2012 *Phys. Rev.* B **85** 125401
[28] Zhang S-F, Jiang H, Xie X C and Sun Q-F 2014 *Phys. Rev.* B **89** 155419
[29] Datta S 1997 *Electronic Transport in Mesoscopic Systems* (Cambridge: Cambridge University Press)
[30] Haug H and Jauho A-P 2008 *Quantum Kinetics in Transport and Optics of Semiconductors* 2nd edn (Berlin: Springer)
[31] Jauho A-P, Wingreen N S and Meir Y 1994 *Phys. Rev.* B **50** 5528
[32] Sancho M P L, Sancho J M L and Rubio J 1984 *J. Phys. F: Met. Phys.* **14** 1205
[33] Sancho M P L, Sancho J M L and Rubio J 1985 *J. Phys. F: Met. Phys.* **15** 851
[34] Lee D H and Joannopoulos J D 1981 *Phys. Rev.* B **23** 4988
[35] Lee D H and Joannopoulos J D 1981 *Phys. Rev.* B **23** 4997
[36] Xing Y, Zhang L and Wang J 2011 *Phys. Rev.* B **84** 035110
[37] Chen L, Liu Q, Lin X, Zhang X and Jiang X 2012 *New J. Phys.* **14** 043028
[38] Xiao D, Chang M-C and Niu Q 2010 *Rev. Mod. Phys.* **82** 1959