Impurity scattering and Friedel oscillations in monolayer black phosphorus

Yong-Lian Zou,[1,2] Juntao Song,[2] Chunxu Bai,[1] and Kai Chang[1,*]

[1]*SKLSM, Institue of Semiconductors, Chinese Academy of Sciences, P.O. Box 912, Beijing 100083, China*

[2]*Department of Physics and Hebei Advanced Thin Film Laboratory, Hebei Normal University, Hebei 050024, China*

(Received 25 March 2016; revised manuscript received 29 June 2016; published 19 July 2016)

We study the impurity scattering effect in black phosphorene (BP) in this work. For a single impurity, we calculate the impurity-induced local density of states (LDOS) in momentum space numerically based on a tight-binding Hamiltonian. In real space, we calculate the LDOS and Friedel oscillation analytically. The LDOS shows strong anisotropy in BP. Many impurities in BP are investigated using the T-matrix approximation when the density is low. Midgap states appear in the band gap with peaks in the DOS. The peaks of midgap states are dependent on the impurity potential. For finite positive potential, the impurity tends to bind negative charge carriers and vice versa. The infinite-impurity-potential problem is related to chiral symmetry in BP.

DOI: 10.1103/PhysRevB.94.035431

I. INTRODUCTION

Graphene has remarkable electronic, optical, and mechanical properties and shows promising applications in electronic devices [1]. Since the successful production of graphene in 2004 [2], there have been many studies on graphene both experimentally and theoretically [1]. However, due to its gapless spectrum, graphene is not a good candidate for on-off devices. It is desirable to find two-dimensional materials with a tunable gap which can be utilized as on-off devices. A promising candidate reported in recent years is monolayer black phosphorene (BP), which has a direct band gap [3].

Monolayer BP has been fabricated in the laboratory using the exfoliation method [3]. Since then, more and more works have focused on single-layer and few-layer BP [4–10]. Bulk BP has a band gap of 0.31–0.35 eV, and single-layer BP has a band gap of about 1.5 eV [3]. Monolayer BP has large mobility as well as a band gap which makes it a promising material in electronic devices. It has different masses along the armchair and zigzag directions. This anisotropy has a great impact on its transport and optical properties. For example, Low *et al.* [6] studied the plasmons in BP and found that the plasmon excitations have different dispersions along different directions.

Impurity is introduced in the process of fabricating materials. Because an impurity can affect the properties of devices remarkably, it is important to understand the impurity effects in monolayer BP. Impurities could be introduced in many ways. Substrates provide a source of impurities, and the adsorbed atoms on BP or missing atoms in BP can induce vacancies and so on [11]. Impurities provide scattering centers to carriers and are the main contribution to the lifetime of carriers in the low-temperature limit. We mainly focus on short-range impurities in this paper [11]. Understanding impurity effects in BP will provide useful information about this new fabricated material.

To understand how an impurity affects carriers in BP, it is necessary to study the density of states (DOS) of carriers. The electronic DOS has many characteristics showing the presence of impurities. There are many physical properties closely related to the DOS of carriers, such as electrical conductivity, optical conductivity, and scanning-tunneling microscope (STM) images. The DOS of carriers is also related to the polarization of the system, which modifies the electronic screening.

For a single impurity, the Fourier transform scanning tunneling spectroscopy (FT-STS), which is Fourier transformation of local density of states (FT-LDOS), can reveal much information about carriers [12–16]. The FT-STS shows interference patterns which originate from interference of incoming waves and outgoing waves scattered by the impurity. The scattering occurs on contours of constant energy, and momentum is transferred from the impurity to carriers, so the amplitudes in FT-STS can reveal what kind of scattering can happen and also the properties of the impurity. For example, the FT-STSs in monolayer graphene and bilayer graphene show different signatures which can be used to distinguish monolayer and bilayer graphene [14]. For surface states of a three-dimensional topological insulator, the backscattering is forbidden. Thus, when the FT-STS amplitude in experiment is weak for the backscattering, it can be viewed as strong evidence for the topological insulator [16]. Therefore, FT-STS is a useful method to investigate material properties.

In real space, the disturbance of the impurity relocates electrons, and the amplitude of the LDOS oscillates and decays away from the position of the impurity. So the electron density at position $\mathbf{r}$ which is obtained by summing the LDOS at $\mathbf{r}$ up to Fermi energy will oscillate and decay in real space. This is called Friedel oscillation (FO). In graphene, FO has been studied in a series of papers [17,18]. FO decays as r^{-3} in graphene, and the oscillation wave vector is $2k_F$, with k_F being the Fermi surface wave vector.

For a finite but small density of impurities, peaks can show up in the DOS. The states associated with these peaks are called midgap states because they appear inside the band gap [7]. In graphene, the midgap states appear at Dirac points where the DOS is zero [19]. These midgap states can be viewed as bound states attracted by the impurity potential.

In this work, we first study FT-STS in BP which is caused by a single impurity; the numerical results are based on a four-band tight-binding Hamiltonian in BP using the T matrix [19–25]. This problem has been investigated in graphene and in topological insulators [13,16]. However, the

*kchang@semi.ac.cn

strong anisotropy in BP has a strong influence on FT-STS, and it is useful to study these anisotropy-related effects in BP. For a single impurity, FO in real space is also investigated based on a two-band model of BP [5]. For the case of many impurities, we focus on short-range and low-density impurities. We present calculations of DOS in BP using both a T-matrix approximation (full Born approximation) and a self-consistent T-matrix approximation (full self-consistent Born approximation) [24,25]. The T matrix is exact for a single impurity and also a good approximation for many impurities as long as the density is low and impurity is short range [25].

The rest of the paper is organized as follows: In Sec. II we introduce the T-matrix method and also the self-consistent T-matrix method. In Sec. III we calculate FT-STS in BP induced by a single impurity. Section IV gives the Friedel oscillation in BP, and Sec. V deals with the low density of impurities using both the T-matrix and self-consistent T-matrix methods. Finally, we conclude in Sec. VI.

II. HAMILTONIAN OF BP AND THE T MATRIX

We start with our calculation using the tight-binding model of monolayer BP. BP has the same structure as graphene but with the atoms puckered. There are four phosphorus atoms in the unit cell of BP, as shown in Fig. 1(a), and the four-band Hamiltonian is

$$H(\mathbf{k}) = \begin{pmatrix} 0 & f_{AB} & f_{AC} & f_{AD} \\ f_{AB}^* & 0 & f_{BC} & f_{BD} \\ f_{AC}^* & f_{BC}^* & 0 & f_{CD} \\ f_{AD}^* & f_{BD}^* & f_{CD}^* & 0 \end{pmatrix}, \tag{1}$$

where

$$\begin{aligned} f_{AB} &= t_1(1+e^{-ik_y a_y}) + t_3(e^{-ik_x a_x} + e^{-ik_x a_x - ik_y a_y}), \\ f_{AC} &= t_4(1+e^{-ik_y a_y} + e^{-ik_x a_x} + e^{-ik_x a_x - ik_y a_y}), \\ f_{AD} &= t_2 e^{-ik_x a_x} + t_5, \\ f_{BC} &= t_2 + t_5 e^{-ik_x a_x}, \\ f_{BD} &= t_4(1+e^{ik_y a_y} + e^{-ik_x a_x} + e^{-ik_x a_x + ik_y a_y}), \\ f_{CD} &= t_1(1+e^{ik_y a_y}) + t_3(e^{-ik_x a_x} + e^{-ik_x a_x + ik_y a_y}), \end{aligned} \tag{2}$$

with a_x, a_y given in Fig. 1(a). A, B, C, and D denote the four atoms in the unit cell. The tight-binding parameters read $t_1 = -1.220$ eV, $t_2 = 3.665$ eV, $t_3 = -0.205$ eV, $t_4 = -0.105$ eV, $t_5 = -0.055$ eV [26].

In this work we assume the impurity is short range and can be modeled as a δ function potential. In the vicinity of the Dirac point in graphene, long-range Coulomb impurity is the main source of scattering [11], while in BP the Coulomb impurity can be treated as short range because of screening. As a result, we mainly focus on short-range impurity. The impurity potential written in $\mathbf{k}$ space is assumed to be

$$U = u \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3}$$

where u is constant in $\mathbf{k}$ space. The only nonzero element of U in the first row and first column means the impurity is near or resides on the single A atom. This kind of impurity can be used to model hydrogen adsorbed on BP or dopants such as carbon, silicon, sulfur, and oxygen [27]. We take $u \to \infty$ for vacancies. In the single-impurity problem, we fix the position of the impurity in the first unit cell and take the impurity site as the origin in real space. In the case of a low-density concentration of impurities, we fix the number of impurities, and their positions are randomly distributed in the system. The final results take the average over all possible configurations of impurities. The Born approximation is often introduced in the calculation if u is small and the impurity is short range, but when a bound state is formed near the impurity, the Born approximation is not justified. Instead, the T matrix should be introduced in the calculation when bound states are formed [23]. The T matrix is exact for a single impurity for any value of u. For a low-density concentration of impurities or vacancies, the T-matrix approximation is also a good approximation [24]. We also calculate the DOS using the self-consistent T-matrix approximation and find that they do not make big differences.

For a single impurity, the LDOS is obtained by taking the imaginary part of the full Green's function. The full Green's function is carried out using the T matrix and Matsubara

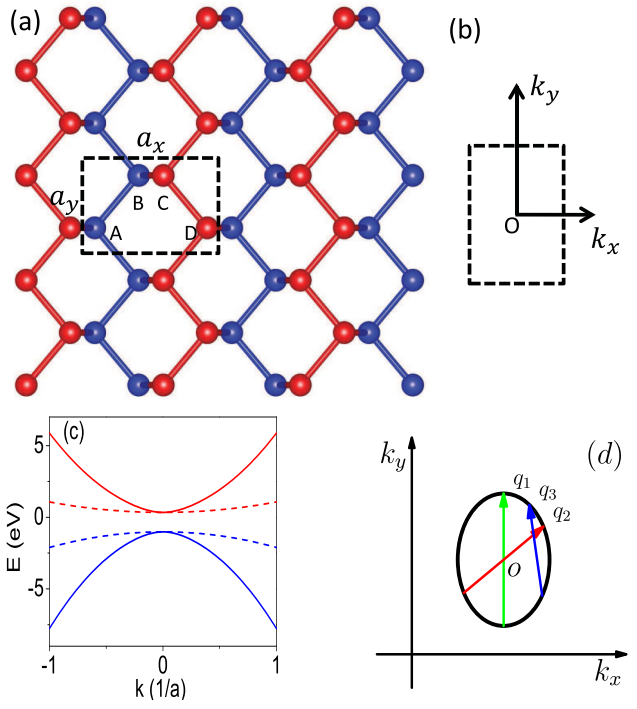

FIG. 1. Lattice structure of monolayer black phosphorene. (a) Red (blue) atoms represent the upper (lower) layer; there are four atoms in a unit cell, which are labeled A, B, C, and D. The unit-cell size along the x direction is $a_x = 4.43$ Å, and that along the y direction is $a_y = 3.27$ Å. (b) The first Brillouin zone (BZ) of BP. (c) Energy dispersion of monolayer BP along the x (solid) and y (dashed) directions. (d) Momentum transferred by impurity scattering at the energy contour.

frequency [24]. Written in momentum space, it is

$$G(\mathbf{k}_1,\mathbf{k}_2,i\omega_n) = G^0(\mathbf{k}_1-\mathbf{k}_2,i\omega_n) + G^0(\mathbf{k}_1,i\omega_n)T_{\rm imp}(\mathbf{k}_1,\mathbf{k}_2,i\omega_n)G^0(\mathbf{k}_2,i\omega_n), \tag{4}$$

where the $T_{\rm imp}$ matrix satisfies the self-consistent equation

$$T_{\rm imp}(\mathbf{k}_1,\mathbf{k}_2,i\omega_n) = U(\mathbf{k}_1,\mathbf{k}_2) + \sum_{\mathbf{k}'} U(\mathbf{k}_1,\mathbf{k}')G^0(\mathbf{k}',i\omega_n)T_{\rm imp}(\mathbf{k}',\mathbf{k}_2,i\omega_n) \tag{5}$$

and the zeroth Green's function in momentum space is

$$G^0(\mathbf{k},i\omega_n) = [i\omega_n - H(\mathbf{k})]^{-1} \overset{i\omega_n\to\omega+i\eta}{=} [\omega+i\eta-H(\mathbf{k})]^{-1}, \tag{6}$$

where η is set to 0.01 eV in our numerical calculation.

Since BP has four atoms in a unit cell, the above equations are a 4×4 matrix equation. Then the T matrix can be obtained as

$$T_{\rm imp}(i\omega_n) = \frac{V}{I - V/N\sum_{\mathbf{k}\in \rm BZ} G^0(\mathbf{k},i\omega_n)}, \tag{7}$$

where the summation is over the first Brillouin zone (BZ), I is a 4×4 identity matrix, and N is the number of unit cells of BP. In the unitary limit ($u\to\infty$) the T matrix reduces to

$$T_{\rm imp}(i\omega) = -\left[\overline{G}^0_{AA}(i\omega)\right]^{-1}\begin{pmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}, \tag{8}$$

where

$$\overline{G}^0_{AA}(i\omega_n) = \frac{1}{N}\sum_{\mathbf{k}\in\rm BZ} G^0_{AA}(\mathbf{k},i\omega_n). \tag{9}$$

Here and throughout the paper, the subscripts A, B, C, and D are used to denote the corresponding matrix element of lattices in the unit cell as shown in Fig. 1(a).

The LDOS induced by a single impurity located at the origin is

$$\delta\rho(\mathbf{q},i\omega_n) = -\frac{1}{N\pi}\sum_{\mathbf{k}\in\rm BZ}[\delta G(\mathbf{k},\mathbf{k}+\mathbf{q}) - \delta G^*(\mathbf{k},\mathbf{k}+\mathbf{q})], \tag{10}$$

where

$$\delta G(\mathbf{k},\mathbf{k}+\mathbf{q}) = G(\mathbf{k},\mathbf{k}+\mathbf{q}) - G^0(\mathbf{k},\mathbf{k}+\mathbf{q}). \tag{11}$$

For many impurities, we take the low density into account, i.e., small $n_i = N_i/N$, where N_i is the number of impurities. We average over all possible positions of impurities and approximate the Green's function of carriers to first order in n_i using the T-matrix approximation. The Green's function is

$$G(\mathbf{k},i\omega_n) = G^0(\mathbf{k},i\omega_n) + G^0(\mathbf{k},i\omega_n)T(i\omega_n)G(\mathbf{k},i\omega_n), \tag{12}$$

which can be solved as

$$G(\mathbf{k},i\omega_n) = \{[G^0(\mathbf{k},i\omega_n)]^{-1} - T(i\omega_n)\}^{-1}. \tag{13}$$

In the above equations, the T matrix is $T(i\omega_n) = Un_i[1 - U\overline{G}^0(i\omega_n)]^{-1}$. If we use the self-consistent T-matrix approach, the T matrix becomes $T(i\omega_n) = Un_i[1-U\overline{G}(i\omega_n)]^{-1}$, where $\overline{G}(i\omega_n) = 1/N\sum_{\mathbf{k}} G(\mathbf{k},i\omega_n)$. The difference between the T-matrix approximation and the self-consistent T-matrix approximation is that T matrix uses the the zeroth-order Green's function, while the latter uses the full Green's function in calculating $T(i\omega_n)$.

III. FT-STS FOR A SINGLE IMPURITY

It is necessary at the beginning to clarify some issues related to experiments and the conditions under which our approach is appropriate. FT-STS is related to the Fourier-transformed LDOS $\rho(\mathbf{r}_i,i\omega_n)$, with $\mathbf{r}_i$ being the position of the unit cell. The experimental situation is complicated. If the tip of the STM has high resolution, then the LDOS related to each of the four atoms can be detected. If the resolution is low, the LDOS detected may be $\sum_{X=A,B,C,D}\rho_{XX}(\mathbf{r}_i)$. In our work we calculate $\rho_{AA}(\mathbf{r}_i)$ for simplicity. In the zero-temperature limit, it is assumed that for all other interactions, inelastic scattering can be incorporated into the broadening for simplicity. The reason is that in the low-temperature limit, inelastic scattering is suppressed, the electron-electron interaction is screened, and Landau's Fermi-liquid picture is valid.

The FT-LDOS is calculated using Eq. (10), which is a convolution integral essentially. The steps in the numerical calculation based on the tight-binding model are as follows: The zeroth-order Green function in k space on the right-hand side of Eq. (4) is Fourier transformed to real space, and we obtain $G^0(E+i\eta,\mathbf{r}_i)$ and $G^0(E+i\eta,-\mathbf{r}_i)$, respectively, and then Fourier transform their products back to $\mathbf{k}$ space including appropriate coefficients.

Before the presentation of numerical results, we discuss how the scattering process influences the interference pattern in FT-STS and do some analytical calculations. First, the energy is conserved because the impurity is time independent; namely, scattering between states occurs on the same energy contour. Second, from Eq. (4), it is easily seen that the dominant contribution to FT-STS is where energy ω is close to the two poles of the two Green's functions simultaneously. The associated quantity is also called joint DOS [12]. Following Pereg-Barnea and MacDonald [15], we use the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian and consider E located in the conduction band, taken into account in the interband coupling as the renormalization of conduction effective masses. The Green's function to first order in u has the form

$$\delta G(\mathbf{q},i\omega) = \frac{1}{N}\sum_{\mathbf{k}}\left[G^0_{AA}(i\omega,\mathbf{k})uG^0_{AA}(i\omega,\mathbf{k}+\mathbf{q})\right]. \tag{14}$$

Transforming the sum into an integral, using Feynman's parameter, changing variables $k_x\to\sqrt{\eta_c'}k_x$, $k_y\to\sqrt{\nu_c}k_y$, and defining $p_x = \sqrt{\eta_c'}q_x$, $p_y = \sqrt{\nu_c}q_y$, it becomes

$$\delta G(\mathbf{q},i\omega) = \frac{A_c}{(2\pi)^2}\frac{u}{\sqrt{\eta_c'\nu_c}}\int d^2\mathbf{k}\frac{1}{i\omega-\mathbf{k}^2}\frac{1}{i\omega-(\mathbf{k}+\mathbf{p})^2} = \frac{A_c}{4\pi^2}\frac{u}{\sqrt{\eta_c'\nu_c}}\int_0^1 dx$$

$$\times \int d^2\mathbf{k} \frac{1}{[i\omega - k^2 - x(1-x)p^2]^2}$$
$$= -\frac{1}{\pi} \frac{uA_c}{\sqrt{\eta_c' v_c} p^2} \frac{1}{\sqrt{4\omega/p^2 - 1}} \arctan \frac{1}{\sqrt{4\omega/p^2 - 1}}, \tag{15}$$

where $A_c = a_x a_y$ is the area of a unit cell. Finally, the FT-LDOS is

$$\delta\rho(\mathbf{q}, E) = \frac{1}{\pi^2} \frac{uA_c}{\sqrt{\eta_c' v_c} p^2} \mathrm{Im}\mathcal{F}\left(\frac{4E}{p^2}\right), \tag{16}$$

with

$$\mathcal{F}(z) = \frac{1}{\sqrt{z-1}} \arctan \frac{1}{\sqrt{z-1}}. \tag{17}$$

The parameters η_c' and v_c related to effective masses are given in the next section.

According to Eq. (16), the FT-STS is zero outside the contour $p^2 = 4E$, i.e., $\eta_c' q_x^2 + v_c q_y^2 = 4E$. This contour becomes a branch cut for $\mathcal{F}(\frac{4E}{p^2})$. In numerical calculations, the broadening of levels causes the FT-LDOS in regions outside this contour to be nonzero. It is clear that the largest part lies on this contour.

The numerical result is shown in Figs. 2(a)–2(d) with $E = 0.3$ eV, $E = 0.7$ eV, $E = 1.5$ eV, and $E = 2.5$ eV, respectively. The scattering occurs on the contours of constant energy, and the results show that the large amplitude of the FT-LDOS comes from scattering exchanging momentum $\mathbf{q} = 2\mathbf{k}_F$ with $\mathbf{k}_F$, corresponding to the Fermi wave vector. There is only intraband scattering in BP, while in graphene scattering can happen between two nonequivalent K points. However, BP shows strong anisotropy in FT-STS. The prominent feature is that the constructive interference occurs on elliptic contours, which is consistent with Eq. (16). The largest part of the FT-LDOS in BP is backscattering, where momentum transfer is two times the Fermi wave vector; that is, the scattering just reversed the wave vector. It is reasonable to consider other kinds of impurities. For example, the bridge-adsorbed, top-adsorbed impurity [27,28]. The FT-STS shows similar behavior for all these impurities. The FT-STS is mainly due to the truncation of the electron sea at the Fermi surface, and the oscillation period is $2k_F$.

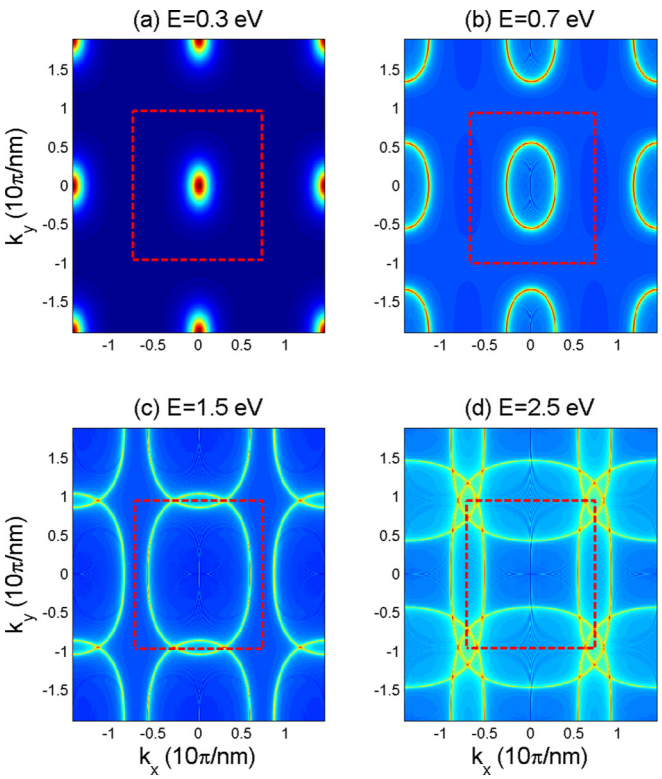

FIG. 2. FT-LDOS in BP with (a) $E = 0.3$ eV, (b) $E = 0.7$ eV, (c) $E = 1.5$ eV, and (d) $E = 2.5$ eV. The red dashed rectangle denotes the first BZ. In the calculation of FT-LDOS, we set $u = 2$ eV. The largest part in the LDOS is shown by the elliptic contours.

Note there are also small interference patterns inside the elliptic contours in Fig. 2. These small parts of the FT-LDOS are due to terms proportional to the product $\mathrm{Re}\frac{1}{E+i\eta-\varepsilon(\mathbf{k})}\mathrm{Im}\frac{1}{E+i\eta-\varepsilon(\mathbf{k}+\mathbf{q})}$, with $\mathbf{k}$ fixed around a point on the energy contour and $\mathbf{q}$ being variable. This small contribution also exists in the FT-LDOS in graphene, but the isotropy in the energy dispersion makes this part the same in all directions, and it will not emerge. In BP, this small part has different values in different directions, and the smallest part emerges. The incoming waves interfere with the outgoing waves after scattering, which is called FO in real space. We study it in detail in the next section. The strong anisotropy in FT-STS of BP makes the pattern easier to identify in STM experiments.

Finally, it is worth comparing FT-STS in BP to that in graphene [13]. In graphene, the FT-STS has one circle around the Γ point and six circles around six corners of the BZ for each one. The circle around the Γ point is due to intravalley scattering, while the other circles are due to intervalley scattering. The intravalley-scattering-induced FT-STS has a long oscillation wavelength and is easily detectable in experiments. The intervalley-scattering-induced FT-STS has a wavelength of oscillation comparable to the lattice. For BP, the situation is different. There is only scattering around the Γ point, and the FT-STS pattern has only one circle around the center of the BZ when the Fermi energy is near the bottom of the conduction band or the top of the valence band. When the Fermi energy increases, the wave vector of the oscillation increases and finally meets the BZ. The large wave-vector oscillation mode appears when the Fermi energy increases while in the graphene large wave-vector oscillation mode appears no matter where the Fermi energy lies. In conclusion, intervalley scattering and intravalley scattering occur in graphene, while only intravalley scattering occurs in BP. So the commensurate effect for FT-STS occurs at any Fermi energy in graphene due to intervalley scattering, while in BP the effect only occurs when the Fermi energy is large enough.

IV. THE k · p HAMILTONIAN AND FO IN BP

To investigate FO in BP, we use the 2×2 $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian of BP. The Hamiltonian can be described as [4,5]

$$H_{kp} = \begin{pmatrix} E_c + \eta_c k_x^2 + v_c k_y^2 & \gamma k_x \\ \gamma k_x & E_v - \eta_v k_x^2 - v_v k_y^2 \end{pmatrix}, \tag{18}$$

where the subscript c (v) labels the conduction (valence) band, $\eta_{c,v}$ and $\nu_{c,v}$ are related to the effective masses by $\eta_{(c,v)} = \hbar^2/2m_{(c,v)x}$ and $\nu_{(c,v)} = \hbar^2/2m_{(c,v)y}$, and the mass parameters are $m_{cx} = 0.151m_e$, $m_{cy} = 1.062m_e$, $m_{vx} = 0.122m_e$, $m_{vy} = 0.708m_e$, with m_e being the free-electron mass. $E_c = 0.34$ eV ($E_v = -1.018$ eV) is the conduction (valence) band edge, and $\gamma = -5.231$ eV Å is the interband coupling coefficient. The corresponding Green's function in this section has matrix components in conduction bands and valence bands, and the subscripts c and v are used to denote the corresponding matrix element. This Hamiltonian can be obtained from the tight-binding Hamiltonian by low-energy expansion to second order of $\mathbf{k}$ near the Γ point based on D_{2h} symmetry of BP [5]. The $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian has strong anisotropic masses along the x and y directions, and we will use it to calculate FO in BP.

Here we briefly describe how to obtain Eq. (18) from tight-binding Hamiltonian (1). The details can be found in Refs. [5,26]. The eigenvector of tight-binding Hamiltonian (1) is given by $[\phi_A\phi_B\phi_C\phi_D]$. Using unitary transformation

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1&0&1&0\\0&1&0&1\\1&0&-1&0\\0&1&0&-1\end{pmatrix}, \tag{19}$$

the tight-binding Hamiltonian can be reduced to a block-diagonal Hamiltonian with each block being a 2×2 Hamiltonian, of which the eigenvector is

$$\Psi = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_A+\phi_C\\ \phi_B+\phi_D\end{pmatrix}. \tag{20}$$

So BP can be described by a 2×2 Hamiltonian. After a rotation of the Pauli matrices $\tau_x \mapsto \tau_z$ followed by $\tau_y \mapsto \tau_x$ [5], this 2×2 Hamiltonian is transformed into Eq. (18). Under this rotation, the impurity potential U is transformed into

$$U = u/\sqrt{2}\begin{pmatrix}1&-i\\ i&1\end{pmatrix}. \tag{21}$$

In this section, we use the Born approximation, which is enough to calculate FO. To first order in U, the modified Green's function in real space is

$$\begin{aligned} G(\mathbf{r},i\omega_n) &= G^0(\mathbf{r},i\omega_n)UG^0(-\mathbf{r},i\omega_n)\\ &= \begin{pmatrix}G^0_{cc}(\mathbf{r},i\omega_n) & G^0_{cv}(\mathbf{r},i\omega_n)\\ G^0_{vc}(\mathbf{r},i\omega_n) & G^0_{vv}(\mathbf{r},i\omega_n)\end{pmatrix}\frac{u}{\sqrt{2}}\begin{pmatrix}1&-i\\ i&1\end{pmatrix}\\ &\times\begin{pmatrix}G^0_{cc}(-\mathbf{r},i\omega_n) & G^0_{cv}(-\mathbf{r},i\omega_n)\\ G^0_{vc}(-\mathbf{r},i\omega_n) & G^0_{vv}(-\mathbf{r},i\omega_n)\end{pmatrix}. \end{aligned} \tag{22}$$

The Green's function in $\mathbf{k}$ space is

$$\begin{aligned} G^0(\mathbf{k},i\omega_n) &= \begin{pmatrix}G^0_{cc}(\mathbf{k},i\omega_n) & G^0_{cv}(\mathbf{k},i\omega_n)\\ G^0_{vc}(\mathbf{k},i\omega_n) & G^0_{vv}(\mathbf{k},i\omega_n)\end{pmatrix}\\ &= \frac{1}{i\omega_n - H_{kp}}. \end{aligned} \tag{23}$$

For convenience, we denote

$$g \equiv \left[i\omega_n - \begin{pmatrix}H_{cc}&0\\0&H_{vv}\end{pmatrix}\right]^{-1} \tag{24}$$

and take the nondiagonal part $\left(\begin{smallmatrix}0&\gamma k_x\\ \gamma k_x&0\end{smallmatrix}\right)$ as a perturbation to get $G^0(\mathbf{k},i\omega_n)$. We will see the Hamiltonian can be reduced to a single-band problem if we study FO with $E > E_c$.

Let the Fermi energy E_F lie in the conduction band such that $k_F \ll \min(1/a_x, 1/a_y)$ and we study the Friedel oscillation in BP. The $\mathbf{k}$-space Green's function $G_{cc} = (g_{cc}^{-1} + H_{cv}g_{vv}H_{vc})^{-1}$, $G_{cv} = g_{cc}H_{cv}G_{vv}$. Since $\eta_v k_{x_F}^2 + \nu_v k_{y_F}^2 \ll (E_c - E_v)$, we expand $H_{cv}g_{vv}H_{vc}$ in $1/(E_F - E_v)$, keeping terms up to quadratic in $\mathbf{k}$, and obtain $G^0_{cc}(i\omega \to E + i\eta) = \frac{1}{E+i\eta-\eta_c k_x^2-\nu_c k_y^2-E_c-\frac{\gamma^2 k_x^2}{E_c-E_v}}$. The nondiagonal part G_{cv} is neglected in our calculation since $\gamma k_{x_F}/(E_F - E_v) \ll 1$. Written in real space,

$$\begin{aligned} G^0_{cc}(\mathbf{r},E) &= \frac{1}{N}\sum_{\mathbf{k}}\frac{\exp i\mathbf{k}\cdot\mathbf{r}}{E-\eta_c' k_x^2 - \nu k_y^2 - E_c}\\ &= -i\frac{A_c}{2\sqrt{\eta_c'\nu_c}}H_0^{(1)}\left(r'\sqrt{\frac{E-E_c}{\sqrt{\eta_c'\nu_c}}}\right), \end{aligned} \tag{25}$$

where $\eta_c' \equiv \eta_c + \gamma^2/(E_c - E_v)$, $r'^2 \equiv (\sqrt{\frac{\nu_c}{\eta_c'}}x^2 + \sqrt{\frac{\eta_c'}{\nu_c}}y^2)$, and $H_0^{(1)}(z)$ is the first-kind Hankel function. In this section, we denote $z = r'\sqrt{\frac{E-E_c}{\sqrt{\eta_c'\nu_c}}}$ for convenience.

The impurity-induced LDOS in real space is obtained from Eq. (22),

$$\begin{aligned} \delta\rho(E,\mathbf{r}) &= -\frac{1}{\pi}\mathrm{Im}[G(E,\mathbf{r})UG(E,-\mathbf{r})]_{cc}\\ &= G_{cc}(E,\mathbf{r})\frac{u}{\sqrt{2}}G_{cc}(E,-\mathbf{r})\\ &= \frac{\sqrt{2}u}{16\pi}A_c^2 J_0(z)Y_0(z). \end{aligned} \tag{26}$$

For a large distance away from the impurity position, i.e., $z \gg 1$, keeping the leading terms in $J_0(z)Y_0(z) = -\cos(2z)/\pi z$ and integrating $\delta\rho(E,\mathbf{k})$ from the edge of the conduction band to the Fermi energy, the leading part in FO induced by impurity over the large distance is

$$\delta n(\mathbf{r}) = \frac{\sqrt{2}u}{16\pi}\frac{A_c^2}{\eta_c'\nu_c}\frac{\sin\left(2\sqrt{\frac{E_F}{\eta_c'}x^2 + \frac{E_F}{\nu_c}y^2}\right)}{\frac{x^2}{\eta_c'} + \frac{y^2}{\nu_c}}, \tag{27}$$

which is our main result in this section.

From Eq. (27), it is seen that FO is anisotropic in BP. The oscillating part in the numerator is due to the discontinuity at the Fermi surface. The anisotropy is related to different masses along different directions. As shown in Fig. 1(d), the curvature of the energy contour is different along the x and y directions. The energy contour in the BZ can be approximately written as $E = \eta_c' k_x^2 + \nu_c k_y^2$ with the same parameters as in Eq. (27) when the Fermi energy lies in the conduction band. We conclude this section by comparing FO in BP to that in graphene. In graphene, the FO decays as r^{-3} due to the cancellation of modes on neighboring sites which decay as r^{-2} [18]. However, there is no such cancellation happening in BP because of the large gap. In BP, the FO shows behavior similar to that of an

ordinary two-dimensional electron gas (2DEG) which shows r^{-2} behavior. Due to different masses of carriers in the x and y directions, FO in BP oscillates anisotropically in the x and y directions.

V. LOW DENSITY OF IMPURITIES

After the investigation of the single-impurity problem, we turn to the many-impurity problem in this section. For low-density impurities which are short range and have finite amplitude, Born approximation is often sufficient. However, to study midgap states in a band gap induced by impurities, Born approximation is not appropriate, and we use the T matrix instead. Impurity is the main source of lifetime for carriers in the low-temperature limit; therefore it is useful to calculate the impurity-induced DOS. We first study the problem of impurities with $u \to \infty$, and then the $t_4 = 0$ problem is considered. Next, the finite-v case is investigated. Finally, we study impurities which have amplitude on all four sites for completeness.

Before going into a detailed analysis of numerical results, it is worth understanding the origin of midgap states. For a single impurity, it is seen from Eq. (4) that impurity-induced bound states comes from poles of $T_{\rm imp}(E)$. For many impurities with a small density, it will be seen that midgap states are related to the poles of $T(i\omega)$ in Eq. (12). Remember that for a low-density impurity, the T-matrix approximation is justified and $T(E) = Un_i[1 - U\overline{G}^0(E + i\eta)]^{-1}$. The DOS is

$$\rho_X(E) = -1/N\pi \mathrm{Im}\left[\sum_{\mathbf{k}} G_{XX}(\mathbf{k}, E + i\eta)\right], \qquad (28)$$

where X denotes A, B, C, or D and $G(\mathbf{k}, E + i\eta)$ is given in Eq. (13). If there is no impurity, BP has crystal symmetry D_{2h}, and DOS on A (B) is equal to that on C (D) [5].

The DOS is related to poles of Green's function. Midgap states come from new poles associated with impurity. Expanding the right-hand side of Eq. (13) to first order in n_i, i.e., $G = G^0 + G^0T(i\omega)G^0$, it is seen that new poles come from $T(i\omega)$. To first order in n_i, the midgap states appear at E, satisfying

$$\det\left[I - V\sum_{\mathbf{k}} G^0(\mathbf{k}, E + i\eta)\right] = 0; \qquad (29)$$

because bound states appear for a single impurity, we denote the solution of this equation as $E_{\rm imp}$. It is worth noting that our analysis is not suitable for long-range or high-density impurities where the interference effect is important.

We compute numerically the DOS in BP in the presence of impurities. The site-dependent DOS is obtained from Eq. (28). We first consider impurities of the form of Eq. (5) and take the limit $u \to \infty$. As shown in Fig. 3, we plot DOS on sites A, B, C, and D in BP separately. It can be seen in Fig. 3 that the DOS on atom A [Fig. 3(a)] shows little evidence of bound states, while the DOS on atom B [Fig. 3(b)] has midgap states. The limit $u \to \infty$ means atom A decouples from the system; the states belonging to the missing A atoms now have zero amplitude on site A, and bound states may emerge. The DOS on atoms C and D shows similar behavior to the DOS on atoms A and B, respectively, because atom A (B) is connected

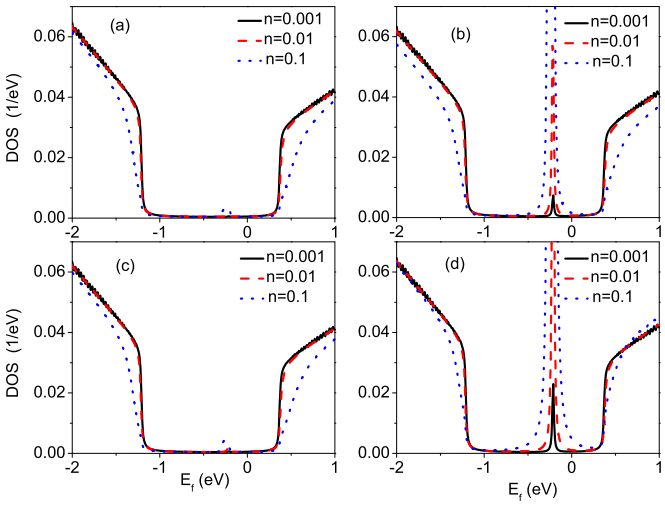

FIG. 3. DOS on lattices (a) A, (b) B, (c) C, and (d) D in the case of $u \to \infty$. The DOS of midgap states on sites A and C is very low. DOS on sites B and D shows peaks in the band gap which indicate midgap states. The impurity density n is proportional to height of peaks.

to atom C (D) through small next-nearest-neighbor t_4, so the midgap state tends to stay on atoms B and D.

To get a deeper understanding of midgap states in BP, we note that the midgap states in BP are similar to midgap sates in graphene in the presence of vacancies. In graphene, the midgap states in the presence of vacancies appear near $E = 0$ eV because of chiral symmetry of the graphene Hamiltonian. Chiral symmetry in graphene is defined as $\sigma_3 H \sigma_3 = -H$, where σ are Pauli matrices acting on pseudospin [20]. Due to chiral symmetry, if there is a state at energy ε, then there is also another state at energy $-\varepsilon$. For every $\mathbf{k}$, there are two eigenstates for the graphene Hamiltonian $H(\mathbf{k})$, with ε and $-\varepsilon$, respectively. If one atom is decoupled from this Hamiltonian, then only one state is allowed to exist, and this state must appear at $\varepsilon = 0$ eV.

Midgap states in BP can also be explained in this way because BP regains chiral symmetry if we set $t_4 = 0$ eV. To see this more clearly, we start from the 2×2 Hamiltonian in Ref. [5], which reads

$$H_{2\times 2} = \begin{pmatrix} f_{AC} & h \\ h^* & f_{AC} \end{pmatrix}, \qquad (30)$$

where h is given in Ref. [5], the detailed expression of which is not relevant here. The Hamiltonian (30) has chiral symmetry if we set $t_4 = 0$ eV. So the midgap states should appear near $E = 0$ eV if $t_4 = 0$ eV, and we verify this by plotting DOS on sites A, B, C, and D, respectively, in BP where we have omitted t_4. The DOS is shown in Fig. 4. As expected, midgap states appear near $E = 0$ eV, which is similar to graphene. So midgap states in BP with nonzero t_4 should also exist, except they are shifted by an energy interval due to the nonzero t_4.

For finite u, the position $E_{\rm imp}$ of midgap states shows different behaviors for positive u and negative u. We plot $u = 10$ eV and -10 eV in Figs. 5 and 6, respectively. As already discussed, impurity density does not affect the position of midgap states, although it may introduce new energy scales

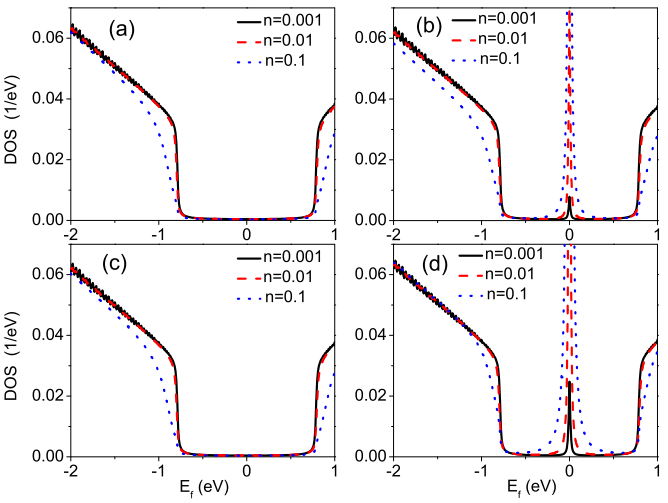

FIG. 4. DOS on sites (a) A, (b) B, (c) C, and (d) D with $t_4 = 0$ and $u \to \infty$. With $t_4 = 0$, BP preserves chiral symmetry, due to which vacancies have bound states exactly at $E = 0$ eV

nu and nW, with W being the width of the band. Impurity density only affects the heights of peaks. Position E_{imp} is determined by u. Positive u tends to bind negative electrons or positive holes, while negative u attracts positive electrons. Note that there is a little difference between DOSs on sites A and on C (also between sites B and D). The reason is that the presence of impurity at site A breaks D_{2h} symmetry and site A is not equivalent to site C in the presence of impurity which resides on site A.

For completeness we study the DOS in the presence of impurities which reside on all four sites, that is to say, impurity of the following kind:

$$U = u \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{31}$$

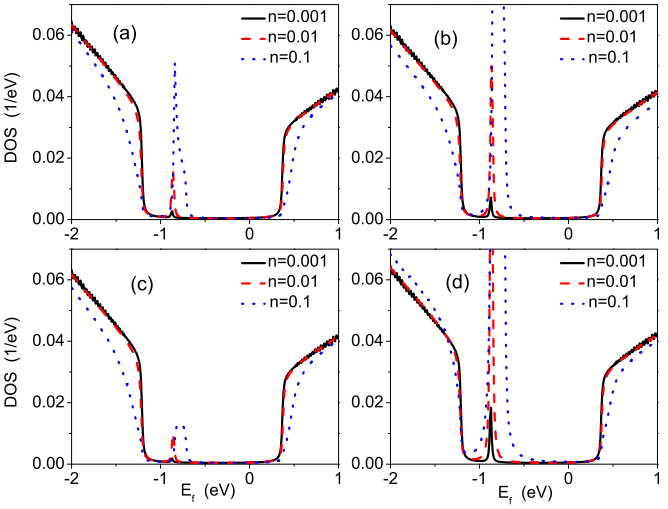

FIG. 5. DOS on sites (a) A, (b) B, (c) C, and (d) D at $u = 10$ eV. For finite u, DOS on sites A and C have finite amplitudes. The positions of the peaks move towards valence bands compared to $u \to \infty$.

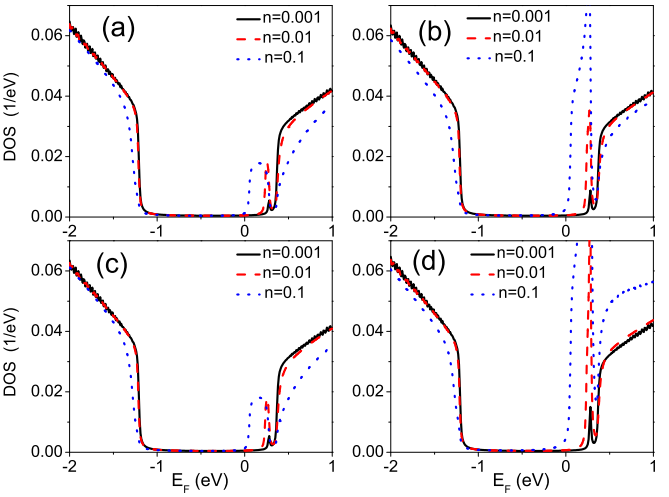

FIG. 6. DOS on sites (a) A, (b) B, (c) C, and (d) D at $u = -10$ eV. For finite u, the DOS on sites A and C has a finite amplitude. The positions of the peaks move towards conduction bands compared to $u \to \infty$.

which has the same amplitude on all four sites. As shown in Fig. 7, there is no midgap states in the gap if we take $u \to \infty$. The DOSs on the four sites are the same, so we present DOS on only site A. The impurity resides on all four sites in a unit cell, so there is no site left to host a bound state with finite energy. For finite u, there is a solution E_{imp} to equation (29). If E_{imp} lies within the gap, midgap states will appear [21]. On the other hand, if the solution E_{imp} lies in the bands, the impurity will modify the DOS in the bands, and no midgap states are induced [7].

Finally, the self-consistent T matrix is used to calculate DOS. The self-consistent T matrix is also called the full self-consistent Born approximation; it replaces the full Green's function with the zeroth-order Green's function in the calculation of the T matrix. The same impurity problem as in Fig. 5 is calculated again using the self-consistent T matrix here. The numerical result is shown in Fig. 8. There is no essential difference from the T-matrix approximation which is shown in Fig. 5; the self-consistent T-matrix method shows the position

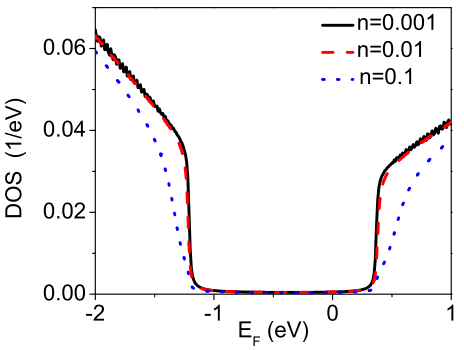

FIG. 7. DOS on site A for impurity potential $u \to \infty$ with the form of Eq. (31), which has amplitudes on all four sites in BP. The DOS on sites B, C, and D are the same as that on site A, which we have not shown here. This kind of impurity does not bind states in the band gap.

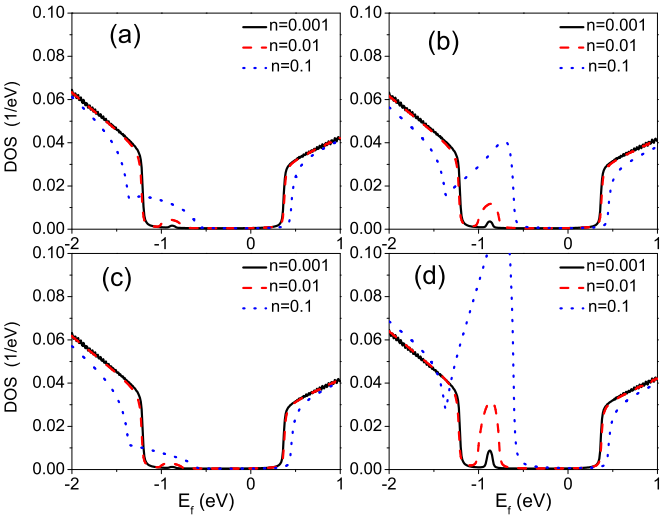

FIG. 8. Self-consistent T-matrix calculation of impurity of the form in Eq. (5). We set $u = 10$ eV as in Fig. 5. Self-consistent calculation takes the finite lifetime of carriers into account, so the width of the peaks is larger than results using the zeroth Green's function.

of E_{imp} is not changed and the amplitude is also proportional to the height of the peaks. But there are some differences indeed. The height and width of the peaks have changed. This can be explained as follows: self-consistent calculation treats the imaginary parts of self-energy more exactly, so the width of the peaks in this approach is more reliable.

VI. DISCUSSION AND CONCLUSIONS

We have discussed the impurity problem in BP. For a single impurity we calculated FT-STS in momentum space and FO in real space. Numerical results of FT-STS are based on a four-band tight-binding Hamiltonian, while FO is analytically calculated based on a 2×2 $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian. The scattering is elastic, so the wave function will get phase shifted after scattering. Due to the phase shift, the interference between incoming and outgoing waves forms patterns in FT-STS. The largest part of the interference amplitude comes from backscattering in FT-STS. Because of the anisotropy in BP, there are two features: one is that contours in FT-STS are elliptic; the other is that small contours appear inside large contours. FO also shows anisotropy in oscillation and decaying directions. FO in BP is different from that in graphene in that it decays as $\frac{1}{x^2/\eta_c' + y^2/v_c}$ in BP but as r^{-3} in graphene.

For the many-impurity problem, we calculated DOS in BP and found that midgap states appear in the band gap. The position E_{imp} of midgap states is related to the amplitude of the impurity potential u. The density determines the height of the peaks. The midgap states appear at $E = 0$ for $u \to \infty$ when t_4 is set to zero, which is due to chiral symmetry of the BP Hamiltonian in the absence of t_4. For finite u, the impurity site tends to bind negative (positive) charge carriers for positive (negative) v. We also calculated DOS using the self-consistent T-matrix approximation and found the width of the peaks is larger than the DOS using the T-matrix approximation.

In this work we have assumed the zero-temperature limit. We have omitted inelastic scattering due to phonons and electron-electron interactions. Indeed, at finite temperature, phonon-mediated scattering is no longer energy conserved, and it may have a large effect on FT-STS. The states over a wide range of energy may take part in the scattering process, and the phase space for scattering is enlarged compared to the zero-temperature limit. So the phonon scattering may not be ignored in FT-STS at finite temperature. For electron-electron interaction, it has been shown that electron-electron interactions have a strong influence on carriers near the Dirac point in graphene [29]. The Fermi velocity will be reshaped, and even the gap is opened at the Dirac point due to exciton condensation [29]. So it is expected that electron-electron interactions may have a great influence on the DOS in BP.

ACKNOWLEDGMENTS

We thank X. Y. Zhou and R. Zhang for many useful discussions; we also acknowledge useful suggestions provided by H. Jiang, J. Liu, and S. G. Cheng. This work was supported by Grant No. 2011CB922204 from the MOST of China and the National Natural Science Foundation (Grants No. 11474085, No. 11174252, No. 11304306, and No. 11504005).

Y.-L.Z. and J.S. contributed equally to this work.

[1] A. H. C. Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).
[2] K. S. Novoselov, A. K. Geim, S. V. Morozv, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306**, 666 (2004).
[3] L. Li, Y. Yu, G. J. Ye, Q. Ge, X. Ou, H. Wu, D. Feng, X. H. Chen, and Y. Zhang, Nat. Nanotechnol. **9**, 372 (2014).
[4] A. S. Rodin, A. Carvalho, and A. H. Castro Neto, Phys. Rev. Lett. **112**, 176801 (2014) .
[5] M. Ezawa, New J. Phys. **16**, 115004 (2014).
[6] T. Low, R. Roldán, H. Wang, F. Xia, P. Avouris, L. M. Moreno, and F. Guinea, Phys. Rev. Lett. **113**, 106802 (2014).
[7] J. M. Pereira and M. I. Katsnelson, Phys. Rev. B **92**, 075437 (2015).
[8] X. Zhou, W.-K. Lou, F. Zhai, and K. Chang, Phys. Rev. B **92**, 165405 (2015).
[9] Y. Jiang, R. Roldán, F. Guinea, and T. Low, Phys. Rev. B **92**, 085408 (2015).
[10] X. Y. Zhou, R. Zhang, J. P. Sun, Y. L. Zou, D. Zhang, W. K. Lou, F. Cheng, G. H. Zhou, F. Zhai, and K. Chang, Sci. Rep. **5**, 12295 (2015).
[11] S. D. Sarma, S. Adam, E. H. Hwang, and E. Rossi, Rev. Mod. Phys. **83**, 407 (2011).
[12] Q.-H. Wang and D.-H. Lee, Phys. Rev. B **67**, 020511 (2003).
[13] C. Bena and S. A. Kivelson, Phys. Rev. B **72**, 125432 (2005).
[14] C. Bena, Phys. Rev. Lett. **100**, 076601 (2008).
[15] T. Pereg-Barnea and A. H. MacDonald, Phys. Rev. B **78**, 014201 (2008).

[16] X. Zhou, C. Fang, W.-F. Tsai, and J. P. Hu, Phys. Rev. B **80**, 245317 (2009).
[17] V. V. Cheianov and V. I. Fal'ko, Phys. Rev. Lett. **97**, 226801 (2006).
[18] A. Bácsi and A. Virosztek, Phys. Rev. B **82**, 193405 (2010).
[19] V. M. Pereira, J. M. B. Lopes dos Santos, and A. H. Castro Neto, Phys. Rev. B **77**, 115109 (2008).
[20] T. O. Wehling, M. I. Katsnelson, and A. I. Lichtenstein, Chem. Phys. Lett. **476**, 125 (2009).
[21] T. O. Wehling, A. V. Balatsky, M. I. Katsnelson, A. I. Lichtenstein, K. Scharnberg, and R. Wiesendanger, Phys. Rev. B **75**, 125425 (2007).
[22] R. R. Biswas and A. V. Balatsky, Phys. Rev. B **81**, 233405 (2010).
[23] A. Balatsky, I. Vekhter, and J.-X. Zhu, Rev. Mod. Phys. **78**, 373 (2006).
[24] B. Y.-K. Hu, E. H. Hwang, and S. Das Sarma, Phys. Rev. B **78**, 165411 (2008).
[25] N. M. R. Peres, F. Guinea, and A. H. Castro Neto, Phys. Rev. B **73**, 125411 (2006).
[26] S. Yuan, A. N. Rudenko, and M. I. Katsnelson, Phys. Rev. B **91**, 115436 (2015).
[27] Y. Guo and J. Robertson, Sci. Rep. **5**, 14165 (2015).
[28] X.-B. Li, P. Guo, T.-F. Cao, H. Liu, W.-M. Lau, and L.-M. Liu, Sci. Rep. **5**, 10848 (2015).
[29] V. N. Kotov, B. Uchoa, V. M. Pereira, F. Guinea, and A. H. Castro Neto, Rev. Mod. Phys. **84**, 1067 (2012).