This is the accepted manuscript made available via CHORUS. The article has been published as:

Electron transport in graphene/graphene side-contact junction by plane-wave multiple-scattering method

Xiang-Guo Li, Iek-Heng Chu, X.-G. Zhang, and Hai-Ping Cheng

Phys. Rev. B **91**, 195442 — Published 28 May 2015

DOI: 10.1103/PhysRevB.91.195442

Electron transport in graphene/graphene side-contact junction by plane-wave multiple scattering method

Xiang-Guo Li[1], Iek-Heng Chu[1], X.-G. Zhang[1,2] and Hai-Ping Cheng[1*]
[1] *Department of Physics and Quantum Theory Project, University of Florida, Gainesville, FL 32611*
[2] *Center for Nanophase Materials Sciences and Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831*

Electron transport in graphene is along the sheet but junction devices are often made by stacking different sheets together in a "side-contact" geometry which causes the current to flow perpendicular to the sheets within the device. Such geometry presents a challenge to first-principles transport methods. We solve this problem by implementing a plane-wave based multiple scattering theory for electron transport. This implementation improves the computational efficiency over the existing plane-wave transport code, scales better for parallelization over large number of nodes, and does not require the current direction to be along a lattice axis. As a first application, we calculate the tunneling current through a side-contact graphene junction formed by two separate graphene sheets with the edges overlapping each other. We find that transport properties of this junction depend strongly on the AA or AB stacking within the overlapping region as well as the vacuum gap between two graphene sheets. Such transport behaviors are explained in terms of carbon orbital orientation, hybridization, and delocalization as the geometry is varied.

I. INTRODUCTION

Graphene has some of the most fascinating electrical, thermal and mechanical properties[1,2] that promise to make it an important material for broad applications. However, to realize such a promise we need to learn more than its bulk properties. For example, most large-area graphene films are produced as polycrystalline sheets[3–6] containing multiple small domains, usually connected by one of two types of boundaries, end contacts where two domains connect within the same sheet with direct atomic bonding and usually referred to as grain boundaries[7–11], and side contacts formed by stacking edge regions of the two graphene domains side by side with van der Waals force holding them together, which are also observed in recent experiment[12]. There is a great potential for useful devices[13,14] using side-contact junctions, in which the overlapping region is the device region while the rest of the two graphene domains act as electrodes. First-principles transport study of either types of boundaries in graphene sheets can be challenging because of poor screening due to low dimensionality. Moreover, the side-contact junctions present a particularly difficult problem because of its unaccommodating geometry for computational methods designed to deal with layer-structured systems.

There are two basic approaches to apply the mesoscopic theory of Laudauer and Buttiker[15,16] to study quantum transport of electrons within the first-principles method. The first approach is to use localized basis sets[17–20] which allow the calculation of the Green's function of an electrode-device-electrode assembly and a straightforward transport calculation based on the Green's function method. However, localized basis sets do not work well for tunneling through large vacuum gaps that require a faithful description of vacuum electron wave functions. The second approach is based on the scattering theory of plane waves. It has been shown to be completely equivalent to the nonequilibrium Green's function method in the case of noninteracting electrons[21]. A rigorous first-principles method[22,23] (Choi-Ihm method) based on scattering theory and pseudopotentials is implemented within the PWCOND part of the QUANTUM ESPRESSO package[24]. This code has become one of the standard tools for quantum transport studies[25–29]. Within this second approach there is also a somewhat different implementation based on the Korringa-Kohn-Rostoker (KKR) band theory, alternatively called the multiple scattering theory, that is adapted to layer-structured systems and thus named the layer-KKR method[21,30]. This method has been particularly successful in the study of spintronics[31–33]. While localized basis is inadequate for the study of tunneling current between two graphene sheets in the side-contact junction, neither the Choi-Ihm method nor the layer-KKR method can be effectively applied to this problem as well, as we will discuss below. This dissatisfaction compels us to search for a third implementation of the plane wave scattering method.

Both the layer-KKR and Choi-Ihm methods use a two-dimensional plane-wave basis and divide the system under study into a stack of sufficiently thin slices along the transport direction. A generalized complex band structure[34] or transmission matrices are then computed by stitching these slices together with appropriate boundary conditions. Each method has its own advantages and drawbacks. On the one hand, the layer-KKR method first solves the Green's function of individual atomic layers and then stitches these layers together using a layer doubling technique based on multiple scattering theory. This approach is more efficient and yields the scattering matrix for any part of the system of interest thus providing more information about the transport properties of the system. However, it has a serious drawback. It is implemented within the muffin-tin or the atomic sphere approximations (ASA), requiring that the space be divided into spheres around each atom within which the Kohn-Sham potential is spherically symmetric. On the other hand, the Choi-Ihm method does not require spherical approximations. It uses much thinner slices and stitches the slices together by matching boundary conditions of the wave functions between the slices. While this method is in principle rigorous, it is computationally expensive. Moreover, it only yields transport information for

the whole system at the end without providing any scattering matrix of the individual parts of the system.

Here we present a plane-wave multiple scattering algorithm that combines the advantages of the layer-KKR and Choi-Ihm methods, and show that it can be an efficient first-principles quantum transport method for tunnel junctions represented by the side-contact junction of graphene sheet discussed above and other devices characterized by low symmetry. We first completely reformulate the multiple scattering theory within a plane-wave basis, then implement an algorithm similar to the layer-KKR method but without the muffin-tin or ASA approximations. Additional computational speedup is realized by calculating the complex bands of each electrode using only its two-dimensional primitive unit cell and then folding the complex bands into the smaller Brillouin zone corresponding to the larger two-dimensional supercell of the scattering region (described in Section II), and by incorporating the method developed by Srivastava *et al.*[35] for low-symmetry nonorthogonal lattices based on the three-dimensional Bloch theorem. While the impementation of both improvements is straightforward, neither is available in the PWCOND code. As a first application, we use the new method to study conduction through a graphene/graphene side-contact junction. Our calculations show that there are significant differences in transmission coefficients and their scaling with the overlapping area between AA and AB stacked graphene sheets. This difference is explained in terms of the orbital delocalization and barrier variation for the two geometries.

The paper is organized as follows. An overview of the theory and the computational approach are given in Section II. In Section III, we discuss the calculation of transport properties of graphene/graphene side-contact junctions and present the results. Conclusions are in Section IV. A detailed derivation of the method is provided in Appendices.

II. MODEL AND METHODS

A. Overview of the method

We consider a system as sketched in FIG. 1, consisting of a central scattering region connected to left and right semi-infinite electrodes, which in general can be different from each other. Mapping this to the graphene side-contact junction, the two electrodes are two semi-infinite graphene sheets with bulk potentials, and the scattering region contains the overlapping region plus a few extra layers outside the overlapping region on both sides to ensure convergence. The plane-wave multiple scattering theory method produces either complex band structure for each of the bulk electrodes or the transmission coefficients for the electrode-device-electrode assembly. The calculation is divided into two stages. In the first stage, the complex band structures of both electrodes are calculated. Even though the basic formalism for this stage is equivalent between the new method and the Choi-Ihm method, the use of the multiple scattering theory speeds up the new method by about 17% on a single processor for the same supercell size and allows it to be more efficiently parallelized (see Appendix D). A further speedup is achieved by recognizing that the complex band structures of the electrodes can be computed using primitive cells only, and be folded into a larger supercell that matches the transverse dimension (perpendicular to the current) of the scattering region at the end of this stage (Appendix B). In the second stage, a set of linear equations are solved to obtain the transmission coefficients. While the final step of solving the linear system is identical to the Choi-Ihm method, most of the computational time is spent in the steps needed to set up the final equations. Again implementing these steps using scattering matrices makes the new method significantly more efficient for parallel computation.

To set up the equations, we first apply in each region the two-dimensional periodic boundary conditions in the transverse (x, y) directions, which are perpendicular to the direction of transport. Systems that are not periodic along the transverse directions can be approximated by sufficiently large two-dimensional supercells with periodic boundary conditions. The Kohn-Sham equation[36] within the framework of the plane-wave pseudopotential method is,

$$E\psi(\mathbf{r}) = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) + V_{loc}(\mathbf{r})\psi(\mathbf{r}) + \sum_{\alpha lm} C_{\alpha lm} \sum_{\mathbf{R}_\perp} e^{i\mathbf{k}_\perp\cdot\mathbf{R}_\perp} W^\alpha_{lm}(\mathbf{r} - \boldsymbol{\tau}^\alpha - \mathbf{R}_\perp) \quad (1)$$

$$C_{\alpha lm} = \sum_{uv} D_{lm,uv} \int d^3\mathbf{r}' [W^\alpha_{uv}(\mathbf{r}' - \boldsymbol{\tau}^\alpha)]^* \psi(\mathbf{r}'), \quad (2)$$

where $V_{loc}(\mathbf{r})$ is the total screened local potential that includes the local part of ionic pseudopotential, electrostatic and exchange-correlation potential due to valence electrons; $W^\alpha_{lm}(\mathbf{r})$ are a set of projector functions associated with atom α at position $\boldsymbol{\tau}^\alpha$, and form the nonlocal part of the pseudopotential with the coefficients $D_{lm,uv}$[37,38]; We use subscript $\perp$ to indicate vectors in the xy plane. $\mathbf{k}_\perp$ and $\mathbf{R}_\perp$ are wave vectors and lattice vectors in the xy plane, respectively.

Following Choi and Ihm[22], we divide each scattering or electrode region into slices perpendicular to the transport direction (z). The slices are sufficiently thin so that within each slice the local potential can be treated as independent of z. Below Eqs. (1)-(7) are reproduced from Ref. 22 for the sake of completeness. The steps after Eq. (7) are different from the Choi-Ihm method. If one such region starting from $z_0 = 0$ and ending at $z_N = d$ is divided into N slices, then Eq. (1) can be rewritten for each slice (labeled by superscript p) as,

$$E\psi^p(\mathbf{r}) = -\frac{\hbar^2}{2m}\nabla^2\psi^p(\mathbf{r}) + V^p_{loc}(\mathbf{r}_\perp)\psi^p(\mathbf{r}) + \sum_{\alpha lm} C_{\alpha lm} \sum_{\mathbf{R}_\perp} e^{i\mathbf{k}_\perp\cdot\mathbf{R}_\perp} W^\alpha_{lm}(\mathbf{r} - \boldsymbol{\tau}^\alpha - \mathbf{R}_\perp). \quad (3)$$

The solutions of the above inhomogeneous differential equation for each slice are to be connected together using the boundary conditions that the wave functions and their derivatives are continuous to yield the total wave function $\psi(\mathbf{r})$ of the entire region in Eq. (1). The first step is to find a set of basis functions,

$$\psi^p_n(\mathbf{r}) = \phi^p_n(\mathbf{r}_\perp) e^{\pm i k^p_n (z - z_p)} \quad (4)$$

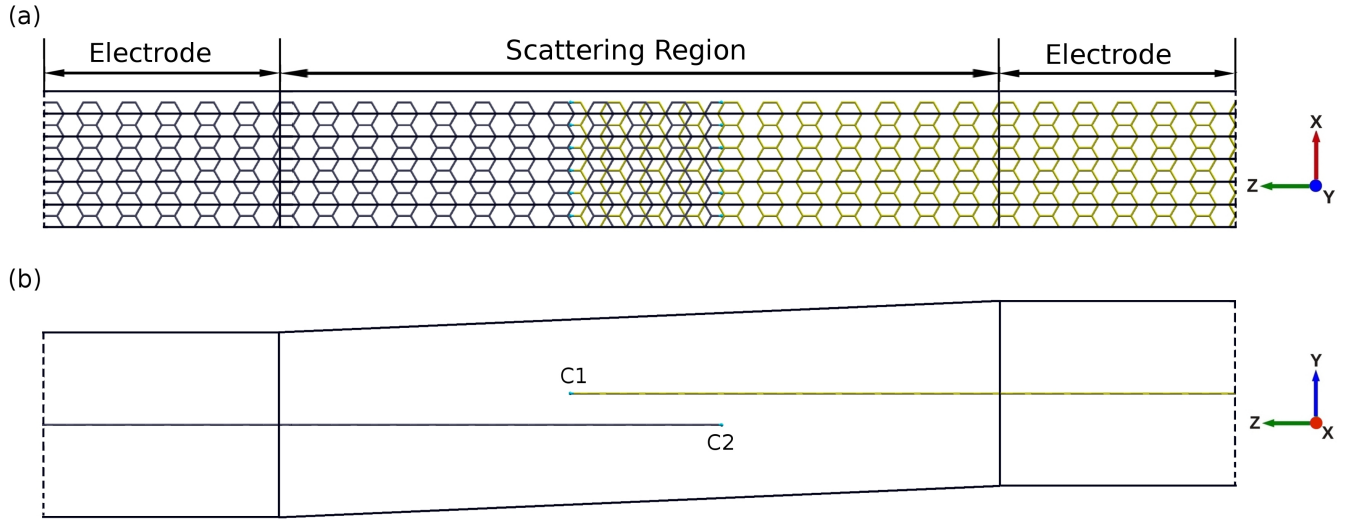

FIG. 1. (Color online) Schematic **(a)** top and **(b)** side views of a graphene side-contact junction. The two graphene layers are connected by one sheet overlapping another to form a bilayer boundary region. The system is periodic along the x direction (six supercells are drawn in the top view) and has a large vacuum gap (more than 15Å) along the y direction. Along the z direction, the system contains the overlapping region and several electrode (graphene) buffer "layers" to ensure the convergence of the potential at the boundary to that in bulk electrodes. The unit cell contains one edge carbon atom on each graphene sheet, denoted as C1 (top layer) and C2 (bottom layer), respectively.

satisfying a homogeneous equation obtained by setting all C_{alm}'s in Eq. (3) to zero, where n labels different homogeneous solutions and goes up to N_{2D}, the cutoff number of the reciprocal lattice vectors $\mathbf{G}_\perp$ in the xy plane, $k_n^p = \sqrt{2m(E - E_n^p)/\hbar^2}$ is the n^{th} wave vector along z direction in the p^{th} slice, E is the total incident energy and E_n^p is the energy eigenvalue of the following equation,

$$E_n^p \phi_n^p(\mathbf{G}_\perp) = \frac{\hbar^2}{2m}|\mathbf{k}_\perp + \mathbf{G}_\perp|^2 \phi_n^p(\mathbf{G}_\perp) + \sum_{\mathbf{G}'_\perp} V_{loc}^p(\mathbf{G}_\perp - \mathbf{G}'_\perp)\phi_n^p(\mathbf{G}'_\perp), \quad (5)$$

where $\phi_n^p(\mathbf{G}_\perp)$ is the Fourier transform of $\phi_n^p(\mathbf{r}_\perp)$. In addition to the homogeneous solution basis set, we also need a particular solution basis set as described below. Starting from Eq. (3) we set only one of C_{alm}'s to one and all others to zero,

$$E\psi_{\alpha lm}^p(\mathbf{r}) = -\frac{\hbar^2}{2m}\nabla^2\psi_{\alpha lm}^p(\mathbf{r}) + V_{loc}^p(\mathbf{r}_\perp)\psi_{\alpha lm}^p(\mathbf{r}) + \sum_{\mathbf{R}_\perp} e^{i\mathbf{k}_\perp\cdot\mathbf{R}_\perp} W_{lm}^\alpha(\mathbf{r} - \tau^\alpha - \mathbf{R}_\perp), \quad (6)$$

which yields a particular solution, $\psi_{\alpha lm}^p(\mathbf{r})$, for each (α, l, m). Both the homogeneous and the inhomogeneous solution basis sets satisfy the Bloch boundary condition within the xy plane, $\psi_{n(\alpha lm)}(\mathbf{r} + \mathbf{R}_\perp) = e^{i\mathbf{k}_\perp\cdot\mathbf{R}_\perp}\psi_{n(\alpha lm)}(\mathbf{r})$. The general solution of Eq. (3) is written as a linear combination of ψ_n^p and $\psi_{\alpha lm}^p$,

$$\psi_{k_i}^p(\mathbf{r}) = \sum_n A_{n,k_i}^p \phi_n^p(\mathbf{r}_\perp) e^{ik_n^p\cdot(z-z_p)} + \sum_n B_{n,k_i}^p \phi_n^p(\mathbf{r}_\perp) e^{-ik_n^p\cdot(z-z_p)} + \sum_{\alpha lm} C_{\alpha lm,k_i}\psi_{\alpha lm}^p(\mathbf{r}), \quad (7)$$

where k_i labels the wavevector of the Bloch and evanescent states when computing complex band structure, and labels the wavevector of the incident waves of the entire scattering region when computing the transmission coefficients.

At this point we take a different approach than Choi and Ihm[22]. The coefficients A_{n,k_i}^p and B_{n,k_i}^p, which depend on k_i, are the only unknowns in the wave function and are determined by matching the wave functions between adjacent slices, usually in a transfer matrix[39] formulation. However, a transfer matrix approach is usually numerically unstable because of the appearance of the exponential factors in Eq. (7). For k_n^p containing nonzero imaginary parts (corresponding to evanescent states), these terms are exponentially decaying and growing waves. Therefore, some transfer matrix elements, as given by Eq. (A14), will grow exponentially while some others will decay exponentially, creating a numerically unstable system. As iteration proceeds, information for the decaying modes will be lost, causing the numerical solution to diverge. To avoid this numerical instability, we separate the waves into forward and backward waves. The forward waves are those that propagate or decay in the positive z direction, the backward waves are those that propagate or decay in the negative z direction. By iterating both types of waves along their exponentially decaying directions we can avoid the numerical instability, a practice already adopted in a number of previous studies[40,41]. This is accomplished by rearranging the boundary conditions between the slices in terms of incident waves of slice $p+1$ with coefficients A^p and B^{p+1}, the outgoing waves of slice $p+1$ with coefficients A^{p+1} and B^p using the scattering matrix[42,43]

$S(p, p+1)$ [see Eq. (A15)] that couples them,

$$\begin{bmatrix} A^{p+1} \\ B^{p} \end{bmatrix} = \begin{bmatrix} S_{11}(p,p+1) & S_{12}(p,p+1) \\ S_{21}(p,p+1) & S_{22}(p,p+1) \end{bmatrix} \begin{bmatrix} A^{p} \\ B^{p+1} \end{bmatrix} + \begin{bmatrix} h^a(p,p+1)C \\ h^b(p,p+1)C \end{bmatrix}. \tag{8}$$

The coefficient $h(p, p+1)$ is also defined in Eq. (A15). Clearly, $S(p,p) = I$, $h^a(p,p) = 0$ and $h^b(p,p) = 0$.

Calculating the scattering matrix for each slice is the first step to obtain the total scattering matrix $S(1,N)$ of the entire region. The second step is to stitch the slices together by applying a doubling technique similar to that employed in the layer-KKR method. In general, one can obtain the scattering matrix for a collection of slices m through k by combining the scattering matrices for m through n and for n through k ($m < n < k$) using the following multiple scattering equations[44],

$$\begin{aligned}
S_{11}(m,k) &= S_{11}(n,k)[I - S_{12}(m,n)S_{21}(n,k)]^{-1}S_{11}(m,n),\\
S_{12}(m,k) &= S_{12}(n,k) + S_{11}(n,k)[I - S_{12}(m,n)S_{21}(n,k)]^{-1}\times \\
&\quad S_{12}(m,n)S_{22}(n,k),\\
S_{21}(m,k) &= S_{21}(m,n) + S_{22}(m,n)S_{21}(n,k)\times \\
&\quad [I - S_{12}(m,n)S_{21}(n,k)]^{-1}S_{11}(m,n),\\
S_{22}(m,k) &= S_{22}(m,n)S_{22}(n,k) + S_{22}(m,n)S_{21}(n,k)\times \\
&\quad [I - S_{12}(m,n)S_{21}(n,k)]^{-1}S_{12}(m,n)S_{22}(n,k),\\
h^a(m,k) &= h^a(n,k) + S_{11}(n,k)[I - S_{12}(m,n)S_{21}(n,k)]^{-1}\times \\
&\quad [S_{12}(m,n)h^b(n,k) + h^a(m,n)],\\
h^b(m,k) &= h^b(m,n) + S_{22}(m,n)h^b(n,k) + S_{22}(m,n)\times \\
&\quad S_{21}(n,k)[I - S_{12}(m,n)S_{21}(n,k)]^{-1}\times \\
&\quad [S_{12}(m,n)h^b(n,k) + h^a(m,n)].
\end{aligned} \tag{9}$$

Each time these equations are applied, the number of slices represented by the scattering matrix is doubled, thus the name "layer-doubling". The final scattering matrix $S(1,N)$ is obtained by applying layer-doubling repeatedly until all slices are contained within the scattering matrix.

While the scattering matrix formalism is numerically more stable than the transfer matrix formalism, the latter is computationally faster than the former. To achieve the optimal balance between speed and numerical stability, we iterate two of the transfer matrices I_{11} and I_{12} along with the scattering matrices S_{21} and S_{22} (see Eq. (A16)) for the first few doubling steps, when the condition number of the transfer matrix I_{11} is reasonable. As soon as the condition number of I_{II} turns bad, we switch over to iterate entirely with the scattering matrices S_{11}, S_{12}, S_{21} and S_{22}.

The final inhomogeneous equation is identical in form to Eq. (8),

$$\begin{bmatrix} A^{N} \\ B^{1} \end{bmatrix} = \begin{bmatrix} S_{11}(1,N) & S_{12}(1,N) \\ S_{21}(1,N) & S_{22}(1,N) \end{bmatrix} \begin{bmatrix} A^{1} \\ B^{N} \end{bmatrix} + \begin{bmatrix} h^a(1,N)C \\ h^b(1,N)C \end{bmatrix}. \tag{10}$$

Eq. (10) provides a direct connection between the total incident waves of the entire region with coefficients A^1 and B^N and the total outgoing waves of the entire region with coefficients A^N and B^1. Eqs. (8)-(10) are the essence of the layer multiple scattering method.

There are $4N_{2D} + N_{orb}$ unknown coefficients counting A ($2N_{2D}$), B ($2N_{2D}$) and C (N_{orb}), with N_{orb} being the total number of nonlocal spheres characterized by W^{α}_{lm}. As mentioned above, the total wave function of the entire region in Eq. (1) can be obtained by connecting the wave function in each slice [see Eq. (7)], which means that the total wave function can be expressed as a function of the unknown coefficients A, B and C. Therefore, the definition of C in Eq. (2) provides N_{orb} equations about the unknown coefficients. In addition, Eq. (10) gives us another $2N_{2D}$ equations. So a total of $2N_{2D}+N_{orb}$ linear equations are provided by Eqs. (2) and (10). Additional $2N_{2D}$ equations are needed, which are provided by the boundary conditions below. Two different types of boundary conditions are needed depending on the physical problem under study. For complex band structures, the generalized Bloch boundary conditions[45] are employed along the transport direction (see Appendix B). For the transmission coefficients of the scattering region, continuity conditions for the wave function and its derivative are applied at the interfaces between the scattering region and each side of the electrodes (see Appendix C).

To compute the complex band structure of an electrode region, the generalized Bloch conditions along the z-direction [see Eqs. (B1), (B2) in Appendix B] need to be applied as the boundary conditions. With Eq. (10), we can express A^N and B^1 as a function of A^1, B^N and C. In addition, we can denote the coefficients C with nonlocal spheres completely fitting the electrode region as $C_{(\alpha lm)'}$, and those with nonlocal spheres crossing the boundaries of the electrode region as $C_{(\alpha lm)}$. From the definition of $C_{(\alpha lm)'}$ in Eq. (2), we can express $C_{(\alpha lm)'}$ as a function of A^N, B^1, A^1, B^N and $C_{(\alpha lm)}$. Therefore, Eq. (10) and the definition of $C_{(\alpha lm)'}$ in Eq. (2) give us an expression of A^N, B^1, $C_{(\alpha lm)'}$ as a function of A^1, B^N and $C_{(\alpha lm)}$. So in the definition of $C_{(\alpha lm)}$ [Eqs. (B3) and (B4)] and Bloch condition Eqs. (B1) and (B2), we can substitute A^N, B^1, $C_{(\alpha lm)'}$ with A^1, B^N and $C_{(\alpha lm)}$. As a result, the only unknowns in Eqs. (B3) and (B4), and (B1) and (B2) are A^1, B^N and $C_{(\alpha lm)}$. Rearranging these equations by setting the unknowns A^1, B^N and $C_{(\alpha lm)}$ as X, we can obtain the generalized eigenvalue problem that takes the form

$$P \cdot X = e^{ikd} e^{i\mathbf{k}_\perp \cdot \mathbf{a}_{3\perp}} Q \cdot X, \tag{11}$$

where $\mathbf{a_3}$ is the third lattice vector which is not in the xy plane; $\mathbf{k} = (\mathbf{k}_\perp, k)$ is the wave vector. Solving this equation yields a set of forward waves $\{\psi^+_k(\mathbf{r})\}$, backward waves $\{\psi^-_k(\mathbf{r})\}$ and also complex band structure $k(E)$.

For transmission coefficient calculations, we need to match the boundary conditions at $z_0 = 0$ and $z_N = d$ for the wavefunction and its derivative [see Eqs. (C14)-(C16) in Appendix C]. Through this process, A^1 and B^1 can be expressed as a function of A^0, B^0 and C [see Eq. (C5)]; A^N and B^N can be expressed as a function of A^{N+1}, B^{N+1} and C [see Eq. (C6)], where A^0 and B^0, A^{N+1} and B^{N+1} are the wave coefficients in the 0^{th} and $(N+1)^{th}$ slices, as defined in Appendix C. Therefore, the unknowns in the definition of C of Eq. (2), and Eqs.

(C14)-(C16) are A^{N+1}, B^0 and C (noting that A^0 and B^{N+1} specify the incident waves, which are known). Rearranging these equations by setting the unknowns A^{N+1}, B^0 and C as X, we can write a set of linear equations in the matrix form

$$M \cdot X = D. \tag{12}$$

The reflection and transmission matrices can be obtained from the solution of X in Eq. (12). The details about the dimension of matrices P, Q, M, D as well as the elements of X are discussed in Appendices B and C. Eqs. (11) and (12) have the same form and same dimensions as the corresponding equations in the Choi-Ihm method[22] but the contents of these equations obtained from the layer multiple scattering theory and thus are entirely different from those in the Choi-Ihm method which are obtained from matching boundary conditions for the layer wave functions.

Next we compare the efficiency of this method against that of the Choi-Ihm method. In the first stage, in which the complex band structures of the electrodes are calculated, there are two speedups. First, by using the primitive cells of the electrodes, the dimension of Eq. (11) for complex band calculations can be reduced by up to several fold for a large calculation. This strategy can also be implemented in the original Choi-Ihm method, although it is not yet available in PWCOND. Second, on a single processor, the new method speeds up the complex band calculations by about 17% according to benchmark runs. There is an additional speedup in a parallel calculation. Since this parallel speedup is the same for both stages, we describe it specifically for the second stage. The most time consuming parts for either stages are the steps to set up Eq. (11) for the complex band calculation or Eq. (12) for the transmission coefficient calculation, not the final step of solving either equation (which is parallelized efficiently but does not consume much computational time). Both the new method and the Choi-Ihm method have identical Eqs. (5) and (6) used at the beginning of the calculation, for which nearly perfect parallelization can be achieved. The major difference between the two methods is in the steps immediately following this calculation. In the multiple scattering method, the scattering matrix for each slice [Eq. (8)] is computed, then the doubling technique [Eq. (9)] is applied to obtain the scattering matrix [Eq. (10)] for the entire system. In the Choi-Ihm method, the wave functions in individual slices are matched across the boundaries in a layer-doubling process until the wave function of the entire system is obtained. In both methods the bottleneck for parallelization is the layer-doubling step, which can scale at best as $\log_2 L$ where L is the number of processors used. However, while the Choi-Ihm method relies entirely on the layer-doubling technique for the wave functions, in the multiple scattering method much of the computation load is shifted to the extraction of the scattering matrices for individual layers, which can be parallelized with nearly 100% efficiency. Once the scattering matrices of all slices are calculated, evaluating Eq. (9) is twice faster than matching boundary conditions for the wave functions across layer boundaries. Therefore the prefactor of the $\log_2 L$ term for the new method is half of that for the Choi-Ihm method.

B. Models and computational details

We apply the plane-wave-multiple-scattering transport method to study the transport properties of a graphene/graphene side contact junction, as illustrated in FIG. 1. The edges of the graphene sheets are zigzag H-terminated in the supercell, the unit cell contains one edge carbon atom for each sheet with a dangling bond saturated by the H atom. We denote these edge atoms as C1 (top graphene layer) and C2 (bottom graphene layer), respectively. Transmission and reflection coefficients of the incident electrode Bloch states through the scattering region are calculated by applying the scattering boundary conditions, Eqs. (C14-C16). The electronic structures of the electrodes and the scattering region are calculated separately using the QUANTUM ESPRESSO package to obtain the self-consistent potentials. Fourteen unit cells of graphene outside the overlapping region are added to the scattering region on each side to ensure the convergence of the potential at the boundaries of the scattering region (see FIG. 1). The electronic structure calculation for the scattering region is obtained by repeating the region in the z direction. To achieve translational invariance under an unit lattice vector along the z direction, the graphene sheet on the left must coincide with the graphene sheet on the right after translation. The best way to accomplish this is to tilt the supercell such that the lattice vector $\mathbf{a_3}$ of the scattering region is no longer perpendicular to the xy plane (see FIG. 1b). Such a configuration cannot be handled by the existing PWCOND code as we discussed earlier. The PWCOND code can only treat this system by doubling the size of the supercell to make the repeating lattice vector perpendicular to the xy plane, thus incurring an 8-fold increase in computation time and 4-fold increase in memory requirement.

We use the Rappe-Rabe-Kaxiras-Joannopoulos ultrasoft pseudopotentials[46] for the C, H atoms with the Perdew-Burke-Ernzerhof exchange-correlation functional[47]. Because of the zigzag H-terminated edges which have non-zero magnetization[48], the calculation is spin-polarized. The energy cutoffs are 60 Ry and 360 Ry for the wave function and charge density, respectively. Gaussian smearing with a width of 0.05 eV is used for the energy levels. The scattering region is periodic along the x direction with a thick vacuum layer of more than 15 Å in the y direction. A 2.46 Å × 20 Å rectangular area is used for the supercell's xy plane. For self-consistent calculations, the z dimension of the supercell is much larger because of the added unit cells of electrodes, and also depends on the width of the overlapping region. A $20 \times 1 \times 1$ k-space mesh is sufficient to sample the Brillouin zone (BZ) for the scattering region.

Graphene has a zero density of states (DOS) at the Dirac point. Using it as electrodes requires the transport calculations to be performed at an energy away from the Dirac point to ensure a finite transmission. In our calculation the incident electron energy is chosen to be 0.1 eV above the Dirac point. At this energy, electrode conduction channels only exist within a small volume of the reciprocal space. The reciprocal space is discretized as $k_\perp = (m^x \Delta k_x, m^y \Delta k_y)$, where $\Delta k_{x(y)}$ is the mesh

interval along $x(y)$ direction, and $m^{x(y)}$ runs from $N_i^{x(y)}$ to $N_f^{x(y)}$. The number of $k_\perp$ points along the $x(y)$ direction in the reciprocal space is $N^{x(y)} = N_f^{x(y)} - N_i^{x(y)} + 1$, which defines the number of conduction channels in the electrode. The transmission coefficient $\mathrm{T}(m^x\Delta k_x, m^y\Delta k_y)$ for each mesh point is computed using the method described in Section II. Because the supercell along the y direction is the vacuum layer of y direction in the electrode in our model system is sufficiently large that the total energy is independent of k_y, the transmission coefficients are also independent of k_y.

Because of the small number of conduction channels in graphene, even if the junction does not contain any scattering the total transmission is still small, a situation we refer to as "electrode-limited" conduction. In order to distinguish the effect of junction scattering from that of the electrode-limited conduction, we compute a transmission probability per electrode conduction channel as follows:

$$\bar{T} = \frac{1}{N^x N^y} \sum_{m^x=N_i^x}^{N_f^x} \sum_{m^y=N_i^y}^{N_f^y} T(m^x\Delta k_x, m^y\Delta k_y). \tag{13}$$

Eq. (13) can be reduced to the following form since the transmission coefficients are independent of k_y:

$$\bar{T} = \frac{1}{N^x} \sum_{m^x=N_i^x}^{N_f^x} T(m^x\Delta k_x, 0). \tag{14}$$

We find that $\Delta k_x = 4.167 \times 10^{-4}(2\pi/a_1)$ is sufficient to converge $\bar{T}$ to 1%. Here a_1 is the lattice constant of electrode in x direction. In particular, if the k mesh in the reciprocal space is infinitesimal or continuous, Eqs. (13) and (14) can be rewritten in an integral form:

$$\bar{T} = \frac{1}{\Omega_{\mathrm{eff}}} \int_{\Omega_{\mathrm{eff}}} T(k_x, k_y) dk_x dk_y, \tag{15}$$

$$\bar{T} = \frac{1}{\lambda_{\mathrm{eff}}^x} \int_{\lambda_{\mathrm{eff}}^x} T(k_x, 0) dk_x. \tag{16}$$

where the effective projected Fermi area of the electrode, denoted as Ω_{eff}, is the area projected to xy plane in k space from the effective Fermi volume of the electrode; λ_{eff}^x is the effective projected Fermi length in x direction of k space for the electrode graphene.

The conductance of the junction is readily calculated from the transmission probability using the Laudauer formula[15]. This is the usual approach taken by almost all quantum transport studies. However, if the transmission is close to unity, the Landauer formula becomes less accurate[49]. In this case, accurate calculation of the junction conductance can be obtained by combining the Landauer approach with the Boltzmann transport equation[49,50]. Boltzmann equation is also needed for non-ballistic transport properties as well as inelastic electron scattering[51,52], which can be important in graphene.

III. RESULTS AND DISCUSSION

A side contact formed between two graphene sheets has already been observed experimentally and is found to be more resistive than grain boundaries formed between graphene domains within the same sheet[10]. The conduction through the side contact can be greatly improved by increasing the overlapping area in the experimental samples. Therefore, the effect on the transport properties of a side contact junction due to the overlapping area as a function of the interlayer distance between the two graphene sheets is the focus of this first-principles study. We begin with the limiting case that the two graphene domains have no overlapping with each other. In this case the edges play an important role, where the dominant factor is the orbital hybridization between the states from the edges of the two graphene layers, which varies with the interlayer distance. This hybridization affects the degree of localization of the electron orbitals near the edges, thus have a deciding role on the transport properties.

A. Tunneling without overlapping between two graphene layers

We first examine how the interlayer distance ($d_\perp$) and horizontal distance ($d_\parallel$) affect the tunneling properties when there is no overlap between the two sheets (see FIG. 2a). The spin-polarized calculations show that the edge of each graphene layer has a small amount of magnetization and the antiferromagnetic (AFM) bonding state between the two edges is the ground state. The carbon atoms on equivalent positions from the two edges in each layer, e.g., C1 and C2, have the same magnitude of magnetization with opposite signs. Due to this spatially antisymmetric spin configuration, the transmission coefficient for the spin-down channel is identical to that for the spin-up channel[53,54]. In the following, all the transmission coefficients calculated are for the spin-up electrons.

With fixed $d_\perp$, the calculated transmission as a function of $d_\parallel$ shows an exponential decay, as plotted in FIG. 2b for $d_\perp = 0$ and $d_\perp = 2.1$ Å. Such a result is expected for simple tunneling through a vacuum barrier with a thickness equal to the separation $d_\parallel$.

The dependence on the interlayer distance ($d_\perp$) is more complex. We set $d_\parallel = 0$ which yields the maximum transmission for non-overlapping sheets. In this case, the transmission coefficient first increases with the distance, reaching its maximum at $d_\perp = 2.1$ Å, followed by an exponential decrease for large distances (see FIG. 2c). The distance of 2.1 Å coincides with the minimum of the total absolute magnetization as shown in FIG. 2c, which also plots the total energy as a function of $d_\perp$ with and without van der Waals (vdW) correction. The energy difference between ferromagnetic (FM) and AFM states is 5 meV when $d_\perp = 1.5$ Å. For other interlayer distances, we cannot find the FM states in the calculation. With vdW functional, the interlayer distance at equilibrium position is about 3.4 Å and the binding energy between the two layer graphene increases from 22 meV (without vdW) to 39 meV (with vdW) per unit cell length along x direction (see FIG.1). The fact that the maximum transmission coincides with the minimum of absolute magnetization suggests that there is a "cancellation" of the magnetic moments between the two opposing edges due to the majority spin electrons from each

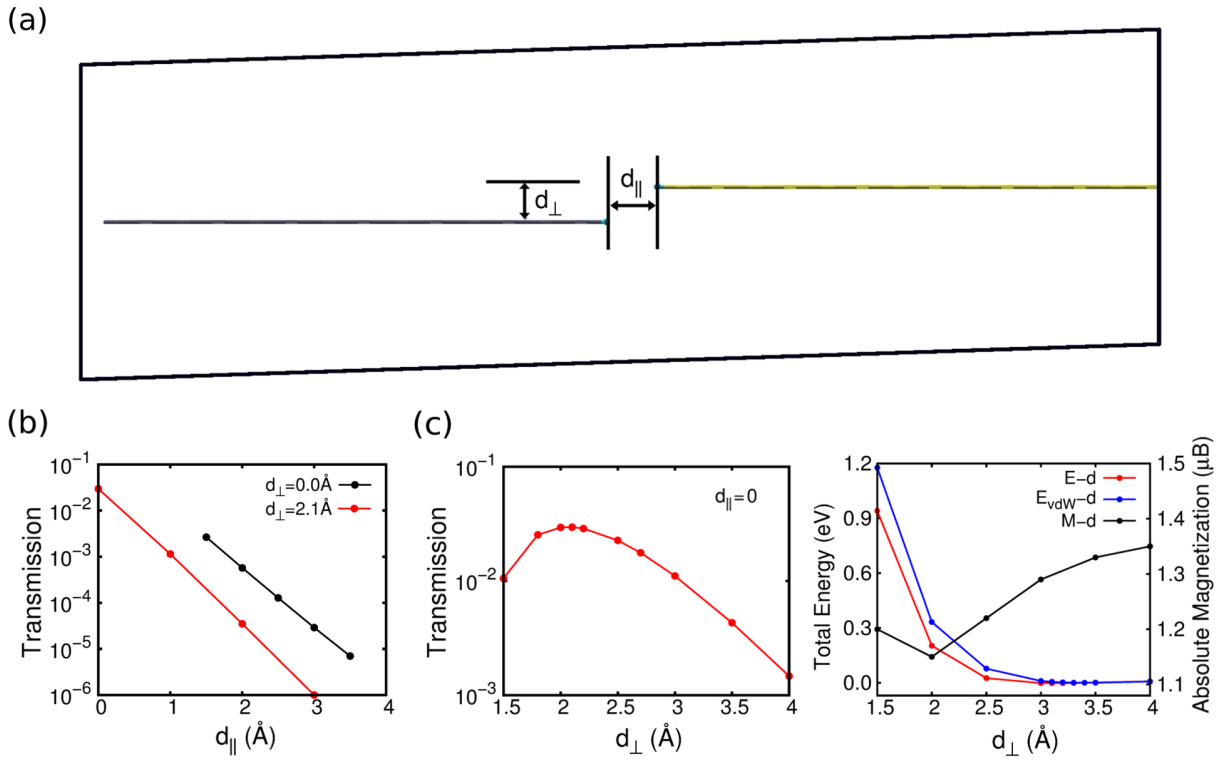

FIG. 2. (Color online) **(a)** Schematic side view of two layer graphene with horizontal distance $d_{\parallel}$ (parallel with graphene plane) and interlayer distance $d_{\perp}$ (perpendicular to graphene plane). The supercell is tilted so that the left of the bottom layer graphene can match the right of the top graphene layer with a periodic boundary condition. **(b)** Spin-up transmission as a function of $d_{\parallel}$ with $d_{\perp}$ = 0.0 Å and 2.1 Å. **(c)** Transmission for spin-up channel in the left panel and total energy with and without vdW correction (E_{vdW} and E), absolute magnetization (M) in the right panel as a function of $d_{\perp}$ with $d_{\parallel}$ = 0. The total energy is shifted with respect to the energy at $d_{\perp}$ = 3.4 Å such that the total energy at $d_{\perp}$ = 3.4 Å is zero for both E and E_{vdW}. y axis in both **b** and **c** for transmission is in logarithm scale.

edge "leaking" into the minority spin channel of the other side.

To understand the mechanism of this spin leakage, we examine the degree of the orbital hybridization between the edge atoms of the two graphene layers. In FIG. 3, we plot the projected density of states (PDOS) of edge C1 and C2 atoms at several interlayer distances. When the interlayer distance is large (e.g. $d_{\perp}$ = 4.0 Å, FIG. 3a), C1 has a peak below the Fermi level for spin-up and a peak above the Fermi level for spin-down and C2 is the opposite, which gives them the same magnitude of magnetization but with opposite signs. When the distance becomes smaller (below around 3.0 Å), we observe a small peak above the Fermi level in the spin-up PDOS of C1 (see FIG. 3b), and correspondingly a small peak above the Fermi level in the spin-down PDOS of C2 (not shown here). This clearly indicates hybridization between the orbitals of C1 and C2. The degree of hybridization between the orbitals of the two graphene edges can be estimated from the size of the small peak in the spin-up PDOS of C1. The small peak reaches its maximum when the interlayer distance is 2.1 Å, indicating that the orbital hybridization is strongest at this distance, leading to the largest transmission coefficient. The size of the small peak is almost the same for $d_{\perp}$ = 1.5 Å and 3.0 Å, consistent with Fig. 2c which shows that the transmission coefficient at these two interlayer distances is almost equal.

The degree of orbital hybridization as a function of $d_{\perp}$ can be visualized directly from the spin-up integrated local density of states (LDOS), as shown in FIG. 4(a-d), which is calculated by integrating the spin-up DOS from 0.05 eV below the incident energy to 0.05 eV above the incident energy (in compliance with our smearing parameter). These plots show that the orbitals around incident energy that carry current are mainly the carbon p_z orbitals. When $d_{\perp}$=4.0 Å (FIG. 4d), there is negligible LDOS on the top graphene layer (contains C1) indicating that there is almost no hybridization between the carbon orbitals from different graphene sheets. This explains the exponentially decaying part of the transmission curve at large interlayer distance. With decreasing interlayer distance, one can observe some spin-up LDOS on the top graphene layer due to the hybridization between the spin-up orbitals of two carbon atoms at the edges of two sheets. Because of the hybridization, some electrons with the spin opposite to the magnetization direction appear on the edge, as if they are "leaked" from the other edge whose magnetization is in the opposite direction, reducing the total absolute magnetization of the junction (see FIG. 2c). The p_z orbital in C1, which has the largest spin leakage compared to other carbon atoms on the top graphene layer, tilts with different angles at different interlayer distances, as shown in FIG. 4e. At an angle when the orbital points directly at the C2 atom on the other edge, the hybridization,

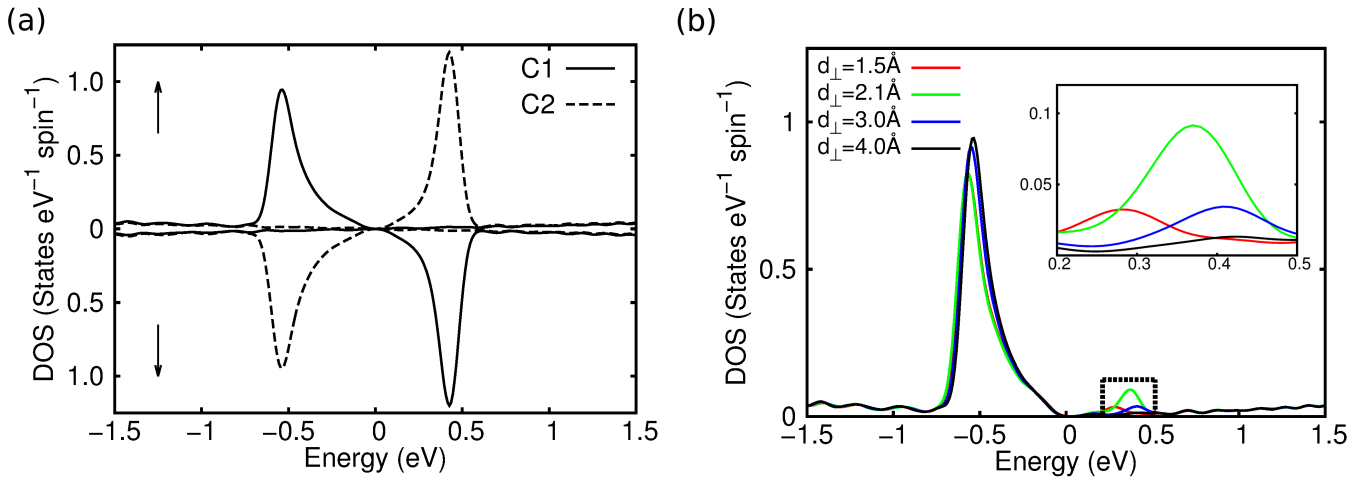

FIG. 3. (Color online) **(a)** Projected density of states (PDOS) for the p orbitals of the edge carbon atoms, C1 and C2, for $d_\perp = 4.0$ Å. At this distance, there is little evidence of coupling between the two graphene sheets. **(b)** Spin-up PDOS for the p orbital of C1 at different interlayer distances. Here we use a smearing of $\sigma = 0.005$ Ry, and the Fermi energy is at 0 eV.

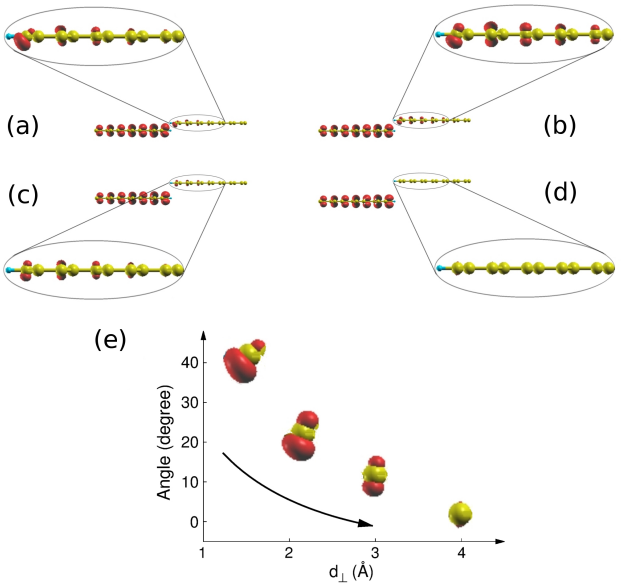

FIG. 4. (Color online) Isosurfaces (at 0.0002 states/Å^3, in red color) of integrated spin-up local density of states from 0.05 eV below the incident energy to 0.05 eV above the incident energy for two graphene layers with interlayer distance **(a)** 1.5 Å, **(b)** 2.1 Å, **(c)** 3.0 Å, **(d)** 4.0 Å. **(e)** The tilting angle of the p_z orbital in C1 of the top sheet as a function of the interlayer distance. Yellow balls are carbon atoms and blue balls are hydrogen atoms.

spin-leakage and degree of the LDOS delocalization in the top layer graphene are the largest. This happens at the interlayer distance of 2.1 Å (see FIG. 4b).

B. Tunneling between two overlapping graphene layers

In the case of two graphene layers overlapping each other, transmission depends sensitively on how the two layers are stacked together, and even whether there is a rotation angle between the two layers[12]. There are two typical types of stacking. One, in which every atom on the second layer lies over an atom of the first, is called the AA stacking. The other, in which half of the atoms in the second layer lie directly over the center of a hexagon in the lower sheet and the other half over an atom, is called the AB stacking. We first examine the AA stacking pattern. We fix the interlayer distance of graphene layers to that of bilayer graphene, which is about 3.4 Å[55]. By varying the overlapping area, expressed in the unit of graphene primitive unit cell area (S_0), we can calculate the transmission coefficient as a function of the overlapping area, which is plotted in FIG. 5a for AA stacking. The transmission first increases superlinearly with the overlapping area, as it varies from two to eight graphene primitive cells. Then the increase slows for larger overlapping areas until the transmission coefficient appears to converge when the overlapping area exceeds ten graphene primitive cells. For large overlapping areas, electrode-limited conduction is reached, evident from the transmission per conduction channel close to unity, a situation for which Eq. (13) is designed to uncover. When the overlapping area is small, it presents a constriction to the current in the junction. The transmission per conduction channel of electrode is expected to be much smaller than unity in this case, and it is confirmed by our calculation. Small transmission due to the constriction at the junction will be referred to as junction-limited transport.

For AB stacked junctions, the transmission as a function of overlapping area is shown in FIG. 5b. The transmission for AB stacking is about an order of magnitude smaller than that for AA stacking, as shown in FIG. 5b. Consequently the transmission per conduction channel is much smaller

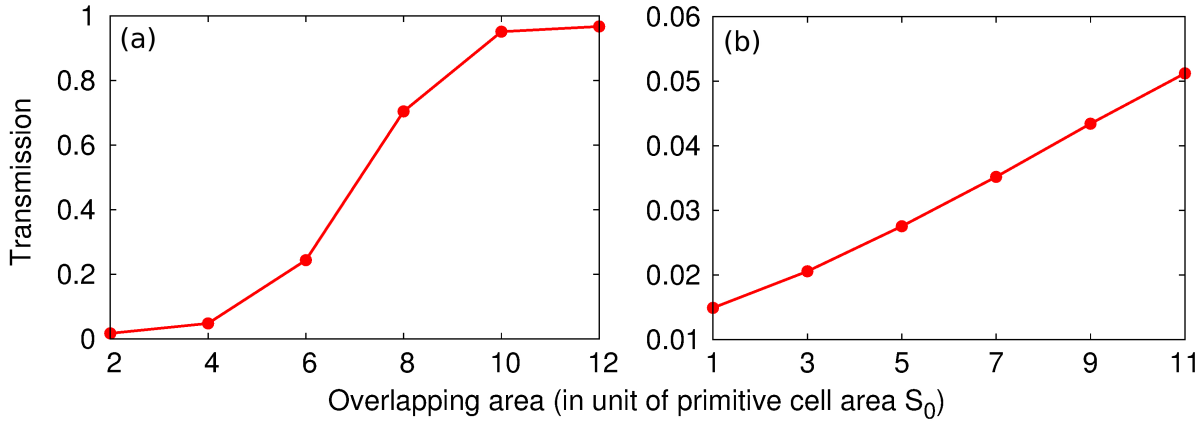

FIG. 5. (Color online) Transmission as a function of overlapping area for **(a)** AA stacking and **(b)** AB stacking.

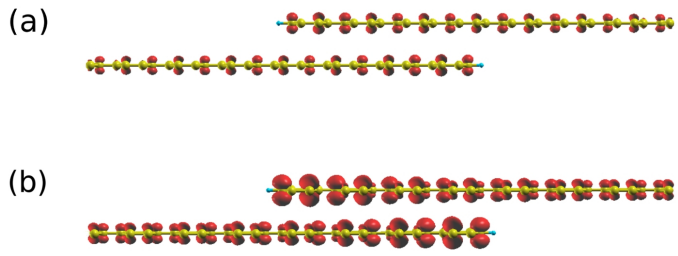

FIG. 6. (Color online) Isosurfaces (at 0.001 states/Å^3, in red color) of integrated total local density of states (summed over both spins) from 0.05 eV below the incident energy to 0.05 eV above the incident energy for **(a)** AB stacking with overlapping area $7S_0$, **(b)** AA stacking with overlapping area $8S_0$. S_0 is the area of graphene primitive unit cell. Yellow balls are carbon atoms and blue balls are hydrogen atoms.

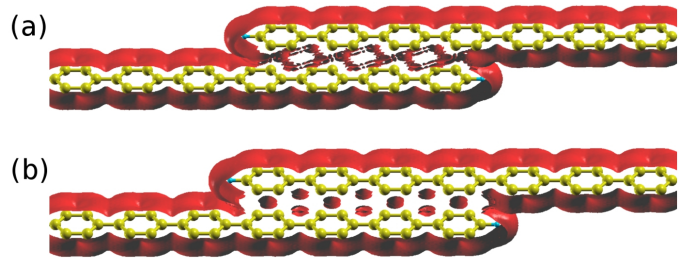

FIG. 7. (Color online) Isosurface of the effective potential that electrons experience at the incident energy for **(a)** AB stacking with overlapping area $7S_0$, **(b)** AA stacking with overlapping area $8S_0$. S_0 is the area of graphene primitive unit cell. The isosurface of the effective potential is in red. Yellow balls are carbon atoms and blue balls are hydrogen atoms.

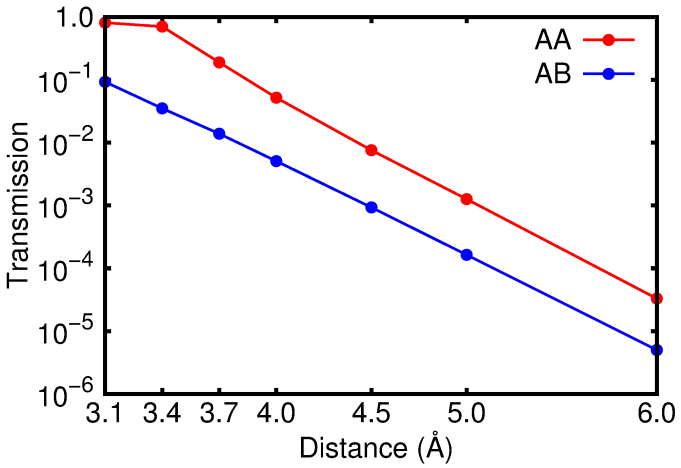

FIG. 8. (Color online) Transmission as a function of interlayer distance for both AA stacking with overlapping area $8S_0$ (red) and AB stacking with overlapping area $7S_0$ (blue). S_0 is the area of graphene primitive unit cell. y axis is in logarithm scale.

than unity, placing AB stacked junctions within the junction-limited regime. The linear dependence of the transmission on the overlapping area depicted in FIG. 5b indicates a simple scaling of the tunneling current with area. This may also be a consequence of weak coupling between the two layers, as evident from the much smaller the total integrated LDOS of the carbon p_z orbitals at the incident energy that carry the current for AB stacking than for AA stacking, as shown in FIG. 6.

To compare and relate the transport properties between AB and AA stacked graphene junctions, we examine the effective potential that electrons experience within the overlapping area between two graphene layers, plotted for the region outside the core radius, where the pseudopotential is equal to the all-electron potential. FIG. 7 shows the isosurface of the effective potential at the incident energy (red color). The region of interest is between the two layers. The volume is divided by the red surface into regions with higher potential than the incident energy, which are enclosed by the red surface and appear as solid red volumes, and regions with lower potential, which are the white colored regions between the two sheets (those outside both layers also have higher potentials). In the case of AA stacking, regions with higher potentials between two layers are contained within isolated pockets, and regions

with lower potentials form connected paths through the entire interlayer volume. For AB stacking, the white color low potential regions between the two layers are blocked off by the red color high potential regions. This difference in the topography of the potential can cause large difference in the transmission of the electron wave function at the incident energy, leading to both differences in the conductance as well as wave function hybridization.

We also plot in FIG. 8 the transmission coefficient as a function of the interlayer distance between the two layers for both AA and AB stacking. When the interlayer distance is larger than 4.5 Å, both AA and AB stacking reach the vacuum tunneling regime giving us the same decaying rate. But the AB stacking reach the vacuum tunneling regime at a shorter interlayer distance than AA stacking, confirming that the interaction between the two layers is much smaller for AB stacking.

IV. CONCLUSION

By implementing the multiple scattering theory within the plane-wave basis, we have improved over previous plane-wave based transport calculation in terms of both speed and parallel efficiency. We apply this method to study a graphene/graphene side contact junction system where the contact is formed by stacking two graphene layers through van der Waals interaction. The transmission through such a junction is closely related to spin leakage between the two graphene edges, a consequence of orbital hybridization between the carbon atoms across the layers which leads to the delocalization of the DOS. When the overlapping area is large, stacking pattern becomes an important factor in deciding transport properties across the layers. The transmission coefficients for AB stacking is one order of magnitude smaller than that for AA stacking, primarily due to the larger volume of the blocking potential within the overlapping region for AB stacking.

Acknowledgements

This work was supported by the US Department of Energy (DOE), Office of Basic Energy Sciences (BES), under Contract No. DE-FG02-02ER45995. A portion of this research was conducted at the Center for Nanophase Materials Sciences, which is sponsored at Oak Ridge National Laboratory by the Division of Scientific User Facilities. The computation was done using the utilities of the National Energy Research Scientific Computing Center (NERSC).

Appendix A: Layer multiple scattering theory for nonlocal pseudopotentials

The general solution of Eq. (1) can be expressed as

$$\psi(\mathbf{r}) = \sum_n a_n \psi_n(\mathbf{r}) + \sum_{\alpha lm} C_{\alpha lm} \psi_{\alpha lm}(\mathbf{r}), \tag{A1}$$

where $\psi_n(\mathbf{r})$ and $\psi_{\alpha lm}(\mathbf{r})$ are the solutions of the following homogeneous and inhomogeneous equations, respectively,

$$-\frac{\hbar^2}{2m}\nabla^2\psi_n(\mathbf{r}) + V_{loc}(\mathbf{r})\psi_n(\mathbf{r}) = E\psi_n(\mathbf{r}), \tag{A2}$$

$$-\frac{\hbar^2}{2m}\nabla^2\psi_{\alpha lm}(\mathbf{r}) + V_{loc}(\mathbf{r})\psi_{\alpha lm}(\mathbf{r}) + \sum_{\mathbf{R}_\perp} e^{i\mathbf{k}_\perp\cdot\mathbf{R}_\perp} W^\alpha_{lm}(\mathbf{r} - \tau^\alpha - \mathbf{R}_\perp) = E\psi_{\alpha lm}(\mathbf{r}). \tag{A3}$$

We divide the system along the z-direction into N slices. If the slices are sufficiently thin then within each slice V_{loc} can be approximated as independent of z. Applying Fourier transformation for the wavefunctions in the p^{th} slice, we have,

$$\psi^p_n(\mathbf{r}) = \phi^p_n(\mathbf{r}_\perp) e^{\pm ik^p_n(z-z_p)} = \sum_{\mathbf{G}_\perp} \phi^p_n(\mathbf{G}_\perp) e^{i(\mathbf{k}_\perp+\mathbf{G}_\perp)\cdot\mathbf{r}_\perp} e^{\pm ik^p_n(z-z_p)}, \tag{A4}$$

$$\psi^p_{\alpha lm}(\mathbf{r}) = \sum_j \phi^p_j(\mathbf{r}_\perp) f^p_{j,\alpha lm}(z) = \sum_{\mathbf{G}_\perp}\sum_j \phi^p_j(\mathbf{G}_\perp) e^{i(\mathbf{k}_\perp+\mathbf{G}_\perp)\cdot\mathbf{r}_\perp} f^p_{j,\alpha lm}(z), \tag{A5}$$

where $\phi^p_j(\mathbf{G}_\perp)$ obeys the eigenvalue Eq. (5); $f^p_{j,\alpha lm}(z) = \sum_{\mathbf{G}_\perp}[\phi^p_j(\mathbf{G}_\perp)]^* e^{-i(\mathbf{k}_\perp+\mathbf{G}_\perp)\cdot\tau^\alpha_\perp} \int_{z_{p-1}}^{z_p} dz' g^p_j(z-z') W^\alpha_{lm}(\mathbf{k}_\perp + \mathbf{G}_\perp, z' - \tau^\alpha_z)$ and $W^\alpha_{lm}(\mathbf{k}_\perp + \mathbf{G}_\perp, z) = \frac{1}{\Omega_{2D}} \int d^2\mathbf{r}_\perp W^\alpha_{lm}(\mathbf{r}) e^{-i(\mathbf{k}_\perp+\mathbf{G}_\perp)\cdot\mathbf{r}_\perp}$. Here Ω_{2D} is the cross-sectional area of the two-dimensional supercell, $g^p_j(z) = \exp(ik^p_j \cdot z)/2ik^p_j$ when $z > 0$ and $g^p_j(z) = \exp(-ik^p_j \cdot z)/2ik^p_j$ when $z < 0$. Calculations of $V_p(\mathbf{G}_\perp)$, $\phi^p_j(\mathbf{G})$ in Eq. (5), and $f^p_{j,\alpha lm}(z)$ are similar to the Choi-Ihm method[22].

The total wavefunction in the p^{th} slice in Eq. (7) is obtained from,

$$\psi^p_{\alpha lm}(\mathbf{r}) = \sum_n \phi^p_n(\mathbf{r}_\perp) f^p_{n,\alpha lm}(z), \tag{A6}$$

$$C_{\alpha lm,k_i} = \sum_{uv} D_{lm,uv} \sum_{p=1}^{N} \int_{z_{p-1}}^{z_p} dz \int d^2\mathbf{r}_\perp [W^\alpha_{uv}(\mathbf{r} - \tau^\alpha)]^* \psi^p_{k_i}(\mathbf{r}). \tag{A7}$$

Matching the boundary conditions for the wavefunction and its derivative between any two adjacent slices, we obtain the recurrence relation for the expansion coefficients A^p_{n,k_i} and B^p_{n,k_i},

$$A^p_{nk_i} = \frac{1}{2k^p_n}\left\{\sum_j A^{p+1}_{jk_i}\exp(-ik^{p+1}_j\Delta z)\left[(k^p_n + k^{p+1}_j)\sum_s (M^{p+1}_{js})^* M^p_{ns}\right]\right.$$
$$\left. + \sum_j B^{p+1}_{jk_i}\exp(ik^{p+1}_j\Delta z)\left[(k^p_n - k^{p+1}_j)\sum_s (M^{p+1}_{js})^* M^p_{ns}\right]\right\} + \sum_{\alpha lm} C_{\alpha lm,k_i} H^a_{n,\alpha lm}(p,p+1), \tag{A8}$$

$$B^p_{nk_i} = \frac{1}{2k^p_n}\left\{\sum_j A^{p+1}_{jk_i}\exp(-ik^{p+1}_j\Delta z)\left[(k^p_n - k^{p+1}_j)\sum_s (M^{p+1}_{js})^* M^p_{ns}\right]\right.$$
$$\left. + \sum_j B^{p+1}_{jk_i}\exp(ik^{p+1}_j\Delta z)\left[(k^p_n + k^{p+1}_j)\sum_s (M^{p+1}_{js})^* M^p_{ns}\right]\right\} + \sum_{\alpha lm} C_{\alpha lm,k_i} H^b_{n,\alpha lm}(p,p+1), \tag{A9}$$

where $H^a_{n,\alpha lm}(p,p+1)$, $H^b_{n,\alpha lm}(p,p+1)$ and M^p_{ns} are,

$$H^a_{n,\alpha lm}(p,p+1) \equiv \frac{1}{2k^p_n}\sum_j (k^p_n - k^{p+1}_j) f^{p+1}_{j,\alpha lm}(z_p)\sum_s (M^{p+1}_{js})^* M^p_{ns} - f^p_{n,\alpha lm}(z_p), \tag{A10}$$

$$H^b_{n,\alpha lm}(p,p+1) \equiv \frac{1}{2k^p_n}\sum_j (k^p_n + k^{p+1}_j) f^{p+1}_{j,\alpha lm}(z_p)\sum_s (M^{p+1}_{js})^* M^p_{ns}, \tag{A11}$$

$$M^p_{ns} = \frac{1}{\Omega_{2D}}\int d\mathbf{r}_\perp [\phi^p_n(\mathbf{r}_\perp)]^* \cdot \exp[i(\mathbf{k}_\perp + \mathbf{G}_{\perp,s})\cdot\mathbf{r}_\perp], \tag{A12}$$

and $\mathbf{G}_{\perp,s}$ denotes the s^{th} reciprocal lattice vector in $\{\mathbf{G}_\perp\}$ and Δz the thickness of a single slice along the z-direction.

The above recurrence relations can be rearranged to form an inhomogeneous equation,

$$\begin{bmatrix} A^p \\ B^p \end{bmatrix} = \begin{bmatrix} I_{11}(p,p+1) & I_{12}(p,p+1) \\ I_{21}(p,p+1) & I_{22}(p,p+1) \end{bmatrix}\begin{bmatrix} A^{p+1} \\ B^{p+1} \end{bmatrix} + \begin{bmatrix} H^a(p,p+1)C \\ H^b(p,p+1)C \end{bmatrix}, \tag{A13}$$

where A^p, B^p and C are matrices with elements $\{A^p_{nk_i}\}$, $\{B^p_{nk_i}\}$ and $\{C_{\alpha lm,k_i}\}$, respectively; $I(p,p+1)$ is the transfer matrix between adjacent slices p and $p+1$ in the following form,

$$I_{nj}(p,p+1) = \frac{1}{2k^p_n}\sum_s (M^{p+1}_{js})^* M^p_{ns}\begin{bmatrix} \exp(-ik^{p+1}_j\Delta z)(k^p_n + k^{p+1}_j) & \exp(ik^{p+1}_j\Delta z)(k^p_n - k^{p+1}_j) \\ \exp(-ik^{p+1}_j\Delta z)(k^p_n - k^{p+1}_j) & \exp(ik^{p+1}_j\Delta z)(k^p_n + k^{p+1}_j) \end{bmatrix}. \tag{A14}$$

The wave function is separated into forward and backward waves with the help of the scattering matrix, whose elements are calculated from the transfer matrix,

$$\begin{aligned} S_{11}(p,p+1) &= [I_{11}(p,p+1)]^{-1}, \\ S_{12}(p,p+1) &= -[I_{11}(p,p+1)]^{-1} I_{12}(p,p+1), \\ S_{21}(p,p+1) &= I_{21}(p,p+1) S_{11}(p,p+1), \\ S_{22}(p,p+1) &= I_{22}(p,p+1) + I_{21}(p,p+1) S_{12}(p,p+1), \\ h^a(p,p+1) &= -[I_{11}(p,p+1)]^{-1} H^a(p,p+1), \\ h^b(p,p+1) &= H^b(p,p+1) + I_{21}(p,p+1) h^a(p,p+1). \end{aligned} \tag{A15}$$

To stitch the slices together, we can derive the general recurrence relations for $1 \le m < n < k \le N$, as shown in Eq. (9). Each iteration of the recurrence relations results in the doubling of the number of slices represented by the the scattering matrix. This is called the "doubling technique". At the end of the iteration, we obtain the scattering matrix that represent the entire region [see Eq. (10)]. To further improve speed, we use a set of recurrence relations based on a mixture of transfer matrices I_{11}, I_{12} and

scattering matrices S_{21}, S_{22} which is faster but numerically less stable,

$$\begin{aligned}
I_{11}(m,k) &= [I_{11}(m,n) + I_{12}(m,n)S_{21}(n,k)]I_{11}(n,k),\\
I_{12}(m,k) &= [I_{11}(m,n) + I_{12}(m,n)S_{21}(n,k)]I_{12}(n,k) + I_{12}(m,n)S_{22}(n,k),\\
S_{21}(m,k) &= S_{21}(m,n) + S_{22}(m,n)S_{21}(n,k)[I_{11}(m,n) + I_{12}(m,n)S_{21}(n,k)]^{-1},\\
S_{22}(m,k) &= S_{22}(m,n)S_{22}(n,k) - S_{22}(m,n)S_{21}(n,k)[I_{11}(m,n) + I_{12}(m,n)S_{21}(n,k)]^{-1}I_{12}(m,n)S_{22}(n,k),\\
H^a(m,k) &= [I_{11}(m,n) + I_{12}(m,n)S_{21}(n,k)]H^a(n,k) + I_{12}(m,n)h^b(n,k) + H^a(m,n),\\
h^b(m,k) &= h^b(m,n) + S_{22}(m,n)h^b(n,k) - S_{22}(m,n)S_{21}(n,k)[I_{11}(m,n) + I_{12}(m,n)S_{21}(n,k)]^{-1}\times\\
&\quad [I_{12}(m,n)h^b(n,k) + H^a(m,n)].
\end{aligned} \tag{A16}$$

We iterate these equations until the condition number of I_{11} exceeds a predetermined criterion. Then we switch to Eq. (9) after computing S_{11} and S_{12} from I_{11} and I_{12} from Eq. (A15).

Appendix B: Complex band structure calculation

If the potential is periodic along the z-direction, the Bloch conditions can be applied as the boundary conditions,

$$\psi_{k_i}(\mathbf{r}_\perp + \mathbf{a}_{3\perp}, z = d) = e^{i\mathbf{k}_\perp\cdot\mathbf{a}_{3\perp} + ikd}\psi_{k_i}(\mathbf{r}_\perp, z = 0), \tag{B1}$$

$$\frac{\partial\psi_{k_i}(\mathbf{r}_\perp + \mathbf{a}_{3\perp}, z = d)}{\partial z} = e^{i\mathbf{k}_\perp\cdot\mathbf{a}_{3\perp} + ikd}\frac{\partial\psi_{k_i}(\mathbf{r}_\perp, z = 0)}{\partial z}, \tag{B2}$$

where $\mathbf{a}_{3\perp}$ is the component of $\mathbf{a}_3$ in the xy plane and d is the z-component of $\mathbf{a}_3$. In a transport problem, the wave functions (of either electrode) do not extend to infinity in all directions – they match to boundary conditions at the interfaces between the electrodes and the scattering region. Thus the requirement that the wave vector k is real is no longer necessary. Solutions with complex k (evanescent waves) are now allowed, changing the above boundary conditions to the generalized Bloch conditions[34]. States with real k's are the propagating (Bloch) states and those with complex k's are the evanescent states.

A complication of imposing the boundary conditions, Eqs. (B1) and (B2), on a nonlocal pseudopotential, is that one must account for the nonlocal spheres that cross one boundary plane and are thus folded to the boundary plane on the other side of the supercell by the Bloch boundary condition. This requires Eq.(A7) to be rewritten in the following form when the nonlocal spheres (characterized by W^α_{lm}) cross the left boundary of the unit cell at $z = 0$,

$$C_{\alpha lm,k_i} = \sum_{uv} D_{lm,uv}\left[\int_0^d dz\int d^2\mathbf{r}_\perp[W^\alpha_{uv}(\mathbf{r} - \tau^\alpha)]^*\psi_{k_i}(\mathbf{r}) + e^{-ikd}\int_0^d dz\int d^2\mathbf{r}_\perp[W^\alpha_{uv}(\mathbf{r} - \tau^\alpha - d\hat{z})]^*\psi_{k_i}(\mathbf{r})\right], \tag{B3}$$

and for those crossing the right boundary of the unit cell at $z = d$,

$$C_{\alpha lm,k_i} = \sum_{uv} D_{lm,uv}\left[\int_0^d dz\int d^2\mathbf{r}_\perp[W^\alpha_{uv}(\mathbf{r} - \tau^\alpha)]^*\psi_{k_i}(\mathbf{r}) + e^{ikd}\int_0^d dz\int d^2\mathbf{r}_\perp[W^\alpha_{uv}(\mathbf{r} - \tau^\alpha + d\hat{z})]^*\psi_{k_i}(\mathbf{r})\right]. \tag{B4}$$

The total number of those spheres is $N_{crosl} + N_{crosr}$ with $N_{crosl}(N_{crosr})$ being the total number of nonlocal spheres crossing left (right) boundary of the unit cell.

The unknowns in Eqs. (10), (A7), (B1) and (B2) are A^1, A^N, B^1, B^N, $C_{(\alpha lm)'}$ (nonlocal spheres completely fitting the electrode region) and $C_{(\alpha lm)}$ (nonlocal spheres crossing the boundaries of the electrode region). Some of these unknowns, A^N, B^1, and $C_{(\alpha lm)'}$, can be eliminated by expressing them in terms of A^1, B^N and $C_{(\alpha lm)}$. The latter are collected as a single vector X,

$$X = \begin{bmatrix} A^1_{nk_i}\\ B^N_{nk_i}\\ C_{(\alpha lm),k_i}\end{bmatrix}. \tag{B5}$$

Then the remaining equations from Eqs. (10), (A7), (B1) and (B2) are combined into a generalized eigenvalue problem, Eq. (11). The dimension of X is $(2N_{2D} + N_{crosl} + N_{crosr})\times$(number of incident waves). P and Q in Eq. (11) are $(2N_{2D} + N_{crosl} + N_{crosr}) \times (2N_{2D} + N_{crosl} + N_{crosr})$ matrices.

Next, we describe how to fold the complex bands of each electrode from the bigger first Brillouin zone of the two-dimensional primitive cell, into the smaller Brillouin zone corresponding to the larger two-dimensional supercell of the junction system. Both the two-dimensional primitive cell and its corresponding supercell is in the transverse plane formed by the first two lattice vectors. The third lattice vector $\mathbf{a}_3$ is the same for both the primitive cell and the supercell. Let $\mathbf{b_1}$, $\mathbf{b_2}$ be the reciprocal lattice vectors of

the electrode primitive cell in the plane perpendicular to $\mathbf{a}_3$ and $\mathbf{B_1}$, $\mathbf{B_2}$ the vectors for the supercell in the same plane. We know that the supercell's first two lattice vectors in real space can be expressed as a linear combination of the first two lattice vectors of the primitive cell with coefficients being integers, respectively. The inner product of the lattice vector in real space and its corresponding reciprocal lattice vector is a constant, 2π. Thus we can write $\mathbf{b_1} = c_{11}\mathbf{B_1} + c_{12}\mathbf{B_2}$ and $\mathbf{b_2} = c_{21}\mathbf{B_1} + c_{22}\mathbf{B_2}$, where c_{11}, c_{12}, c_{21} and c_{22} are integers. We start from the electrode complex band structure described within the larger Brillouin zone of the primitive cell. For each $\mathbf{k}_\perp$ in the first brillouin zone, the Bloch condition along the third lattice vector $\mathbf{a}_3$ is,

$$\psi_{\mathbf{k}}(\mathbf{r} + \mathbf{a}_3) = e^{i\mathbf{k}\cdot\mathbf{a}_3}\psi_{\mathbf{k}}(\mathbf{r}), \tag{B6}$$

where $\mathbf{k} = (\mathbf{k}_\perp, k_z)$. The complex band structure calculation can give us a series of eigenvalues k_z and eigenstates ψ_{k_z} [see Eq. (11)]. The Fourier transform of the electrode potential contains only vectors of the reciprocal lattice of the primitive cell, $m\mathbf{b_1} + n\mathbf{b_2}$, where m and n are integers. Thus each $\mathbf{k}_\perp$ in the first brillouin zone of the reciprocal space of the primitive cell couples only to $\mathbf{k}_\perp + m\mathbf{b_{1\perp}} + n\mathbf{b_{2\perp}}$. Then the eigenstate ψ_{k_z} is superposition of plane waves containing only the wave vector $\mathbf{k}_\perp$ and wave vectors differing from $\mathbf{k}_\perp$ by $m\mathbf{b_{1\perp}} + n\mathbf{b_{2\perp}}$,

$$\psi_{k_z} = \sum_m \sum_n a_{mnk}|\mathbf{k}_\perp + m\mathbf{b}_{1\perp} + n\mathbf{b}_{2\perp}\rangle, \tag{B7}$$

where a_{mnk} are the expansion coefficients; $|\mathbf{k}_\perp + m\mathbf{b}_{1\perp} + n\mathbf{b}_{2\perp}\rangle$ is the plane wave basis.

To fold $\mathbf{k} = (\mathbf{k}_\perp, k_z)$ of the primitive cell to $\mathbf{K} = (\mathbf{K}_\perp, K_z)$ of the supercell, we have the following relationship,

$$\mathbf{k} = \mathbf{K} + m^k\mathbf{B_1} + n^k\mathbf{B_2}, \tag{B8}$$

where m^k and n^k are integers. Then the Bloch condition along the third lattice vector $\mathbf{a}_3$ (supercell and the primitive cell have the same $\mathbf{a}_3$) for the supercell is,

$$\psi_{\mathbf{K}}(\mathbf{r} + \mathbf{a}_3) = \psi_{\mathbf{K}+m^k\mathbf{B_1}+n^k\mathbf{B_2}}(\mathbf{r} + \mathbf{a}_3) = \psi_{\mathbf{k}}(\mathbf{r} + \mathbf{a}_3) = e^{i\mathbf{k}\cdot\mathbf{a}_3}\psi_{\mathbf{k}}(\mathbf{r}) = e^{i(\mathbf{K}+m^k\mathbf{B_1}+n^k\mathbf{B_2})\cdot\mathbf{a}_3}\psi_{\mathbf{K}+m^k\mathbf{B_1}+n^k\mathbf{B_2}}(\mathbf{r}) = e^{i\mathbf{K}\cdot\mathbf{a}_3}\psi_{\mathbf{K}}(\mathbf{r}). \tag{B9}$$

Therefore, the complex band at $\mathbf{k}_\perp$ in the primitive cell will be folded to the complex band at $\mathbf{K}_\perp$ in the supercell. The eigenvalue $K_z = k_z - m^k\mathbf{B_{1z}} - n^k\mathbf{B_{2z}}$ and the corresponding eigenstate can be expanded in the new plane wave basis $|\mathbf{K}_\perp + M\mathbf{B_{1\perp}} + N\mathbf{B_{2\perp}}\rangle$, M and N are integers, as:

$$\psi_{k_z} = \sum_m \sum_n a_{mnk}|\mathbf{k}_\perp + m\mathbf{b_{1\perp}} + n\mathbf{b_{2\perp}}\rangle = \sum_m \sum_n a_{mnk}|\mathbf{K}_\perp + (m^k + mc_{11} + nc_{21})\mathbf{B_{1\perp}} + (n^k + mc_{12} + nc_{22})\mathbf{B_{2\perp}}\rangle. \tag{B10}$$

Appendix C: Reflection and transmission calculation

For transmission coefficient calculations, we need to match the boundary conditions at $z_0 = 0$ and $z_N = d$ for the wavefunction and its derivative, which can be viewed as adding two more slices. We denote the two added slices as the 0^{th} and the $(N+1)^{th}$ slices with wave coefficients A^0 and B^0, A^{N+1} and B^{N+1}. Then we have

$$\begin{aligned} A^1_{jk_i} = \frac{1}{2}\Bigg\{ &\sum_{k'_i\in R} A^0_{k'_ik_i} \sum_s (M^1_{js})^\dagger \left[\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0) - \frac{i}{k^1_j}\exp(ik^1_j\Delta z)\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0)}{\partial z}\right] \\ &+ \sum_{k'_i\in L} B^0_{k'_ik_i} \sum_s (M^1_{js})^\dagger \left[\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0) - \frac{i}{k^1_j}\exp(ik^1_j\Delta z)\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0)}{\partial z}\right]\Bigg\} \end{aligned} \tag{C1}$$

$$\begin{aligned} B^1_{jk_i} = \frac{1}{2}\Bigg\{ &\sum_{k'_i\in R} A^0_{k'_ik_i} \sum_s (M^1_{js})^\dagger \left[\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0) + \frac{i}{k^1_j}\exp(-ik^1_j\Delta z)\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0)}{\partial z}\right] \\ &+ \sum_{k'_i\in L} B^0_{k'_ik_i} \sum_s (M^1_{js})^\dagger \left[\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0) + \frac{i}{k^1_j}\exp(-ik^1_j\Delta z)\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0)}{\partial z}\right]\Bigg\} \\ &+ \sum_{\alpha lm} C_{\alpha lm,k_i}\left[-f^1_{j,\alpha lm}(0)\exp(-ik^1_j\Delta z)\right], \end{aligned} \tag{C2}$$

$$A^N_{jk_i} = \frac{1}{2}\left\{\sum_{k'_i\in R} A^{N+1}_{k'_i k_i}\sum_s (M^N_{js})\left[\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d) - \frac{i}{k^N_j}\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d)}{\partial z}\right]\right.$$
$$\left. + \sum_{k'_i\in L} B^{N+1}_{k'_i k_i}\sum_s (M^N_{js})\left[\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d) - \frac{i}{k^N_j}\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d)}{\partial z}\right]\right\}$$
$$+ \sum_{\alpha lm} C_{\alpha lm,k_i}\left[-f^N_{j,\alpha lm}(d)\right], \tag{C3}$$

$$B^N_{jk_i} = \frac{1}{2}\left\{\sum_{k'_i\in R} A^{N+1}_{k'_i k_i}\sum_s (M^N_{js})\left[\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d) + \frac{i}{k^N_j}\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d)}{\partial z}\right]\right.$$
$$\left. + \sum_{k'_i\in L} B^{N+1}_{k'_i k_i}\sum_s (M^N_{js})\left[\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d) + \frac{i}{k^N_j}\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d)}{\partial z}\right]\right\}, \tag{C4}$$

where L and R represent the sets of forward and backward waves. In the two new slices, we expand the wavefunction with the generalized Bloch basis (including both propagating waves and evanescent waves). Thus for each incident wave k_i, the dimensions of A^0, B^0 and A^{N+1}, A^{N+1} are $N_{2D} + N_{crosl}$ and $N_{2D} + N_{crosr}$, respectively. Recall that the dimensions of A^1, B^1 and A^N, A^N are all N_{2D}. Thus we rearrange the above equations into the following form,

$$\begin{bmatrix} A^1 \\ B^1 \end{bmatrix} = \begin{bmatrix} I_{11}(1,0) & I_{12}(1,0) \\ I_{21}(1,0) & I_{22}(1,0) \end{bmatrix}\begin{bmatrix} A^0 \\ B^0 \end{bmatrix} + \begin{bmatrix} H^a(1,0)C \\ H^b(1,0)C \end{bmatrix}, \tag{C5}$$

$$\begin{bmatrix} A^N \\ B^N \end{bmatrix} = \begin{bmatrix} I_{11}(N,N+1) & I_{12}(N,N+1) \\ I_{21}(N,N+1) & I_{22}(N,N+1) \end{bmatrix}\begin{bmatrix} A^{N+1} \\ B^{N+1} \end{bmatrix} + \begin{bmatrix} H^a(N,N+1)C \\ H^b(N,N+1)C \end{bmatrix}, \tag{C6}$$

where

$$H^a_{j,\alpha lm}(1,0) = 0, \tag{C7}$$
$$H^b_{j,\alpha lm}(1,0) = -f^1_{j,\alpha lm}(0)\exp(-ik^1_j\Delta z), \tag{C8}$$
$$H^a_{j,\alpha lm}(N,N+1) = -f^N_{j,\alpha lm}(d), \tag{C9}$$
$$H^b_{j,\alpha lm}(N,N+1) = 0, \tag{C10}$$

$$I_{jk'_i}(1,0) = \frac{1}{2}\sum_s (M^1_{js})^\dagger\begin{bmatrix} \psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0) - \frac{i}{k^1_j}\exp(ik^1_j\Delta z)\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0)}{\partial z} \\ \psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0) + \frac{i}{k^1_j}exp(-ik^1_j\Delta z)\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=0)}{\partial z} \end{bmatrix}, \tag{C11}$$

$$I_{jk'_i}(N,N+1) = \frac{1}{2}\sum_s (M^N_{js})\begin{bmatrix} \psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d) - \frac{i}{k^N_j}\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d)}{\partial z} \\ \psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d) + \frac{i}{k^N_j}\frac{\partial\psi_{k'_i}(\mathbf{G}_{\perp,s}, z=d)}{\partial z} \end{bmatrix}. \tag{C12}$$

In addition, Eq. (10) can be rewritten as

$$\begin{bmatrix} I & -S_{12}(1,N) \\ 0 & -S_{22}(1,N) \end{bmatrix}\begin{bmatrix} A^N \\ B^N \end{bmatrix} = \begin{bmatrix} S_{11}(1,N) & 0 \\ S_{21}(1,N) & -I \end{bmatrix}\begin{bmatrix} A^1 \\ B^1 \end{bmatrix} + \begin{bmatrix} h^a(1,N)C \\ h^b(1,N)C \end{bmatrix}. \tag{C13}$$

From Eqns. (C5), (C6) and (C13), we can construct the relationship between A^0, B^0 and A^{N+1}, A^{N+1},

$$\begin{bmatrix} I & -S_{12}(1,N) \\ 0 & -S_{22}(1,N) \end{bmatrix}\begin{bmatrix} I_{11}(N,N+1) & I_{12}(N,N+1) \\ I_{21}(N,N+1) & I_{22}(N,N+1) \end{bmatrix}\begin{bmatrix} A^{N+1} \\ B^{N+1} \end{bmatrix} = \begin{bmatrix} S_{11}(1,N) & 0 \\ S_{21}(1,N) & -I \end{bmatrix}\times$$
$$\begin{bmatrix} I_{11}(1,0) & I_{12}(1,0) \\ I_{21}(1,0) & I_{22}(1,0) \end{bmatrix}\begin{bmatrix} A^0 \\ B^0 \end{bmatrix} + \begin{bmatrix} H^a(1,0) + h^a(1,N) - H^a(N,N+1) \\ H^b(1,0) + h^b(1,N) - H^b(N,N+1) \end{bmatrix}C. \tag{C14}$$

Because the left electrode and the scattering region share those atoms whose nonlocal spheres lie across the $z = 0$ plane, we have N_{crosl} additional equations in the form,

$$C_{(\alpha lm),k_i} = \sum_{k_i' \in R} C^{LE}_{(\alpha lm),k_i'} A^0_{k_i',k_i} + \sum_{k_i' \in L} C^{LE}_{(\alpha lm),k_i'} B^0_{k_i',k_i}. \tag{C15}$$

Similarly, for those spheres lying across the $z = d$ plane, we have N_{crosr} additional equations

$$C_{(\alpha lm),k_i} = \sum_{k_i' \in R} C^{RE}_{(\alpha lm),k_i'} A^{N+1}_{k_i',k_i} + \sum_{k_i' \in L} C^{RE}_{(\alpha lm),k_i'} B^{N+1}_{k_i',k_i}. \tag{C16}$$

Here LE (RE) means left (right) electrode.

Therefore for each boundary conditon, altogether we have $2N_{2D} + N_{orb} + N_{crosl} + N_{crosr}$ equations from Eqns. (C14) ($2N_{2D}$), (C15) (N_{crosl}), (C16) (N_{crosr}), (A7) (N_{orb}). The number of unknowns for $\{A^{N+1}\}$, $\{B^0\}$ and $\{C\}$ are $\{N_{2D} + N_{crosl}\} + \{N_{2D} + N_{crosr}\} + \{N_{orb}\}$. Note that $\{A^0\}$ and $\{B^{N+1}\}$ provide the boundary conditions that specify the incident wave, e.g. $A^0 = I$, $B^{N+1} = 0$ for waves incident from the left electrode for which the corresponding transmission and reflection matrices are $\{A^{N+1}\}$ and $\{B^0\}$, respectively. $A^0 = 0$, $B^{N+1} = I$ represents waves incident from the right electrode for which the corresponding transmission and reflection matrices are $\{B^0\}$ and $\{A^{N+1}\}$, respectively. As explained in the main text, rearrange Eqs.(C14), (C15), (C16) and (A7), we can obtain a set of linear equations listed in Eq. (12), where M is a $(2N_{2D}+N_{crosl}+N_{crosr}+N_{orb})\times(2N_{2D}+N_{crosl}+N_{crosr}+N_{orb})$ matrix, D is $(2N_{2D} + N_{crosl} + N_{crosr} + N_{orb})\times$(number of incident waves), X is also $(2N_{2D} + N_{crosl} + N_{crosr} + N_{orb})\times$(number of incident waves) and has the following structure,

$$X = \begin{bmatrix} A^{N+1} \\ B^0 \\ C \end{bmatrix}. \tag{C17}$$

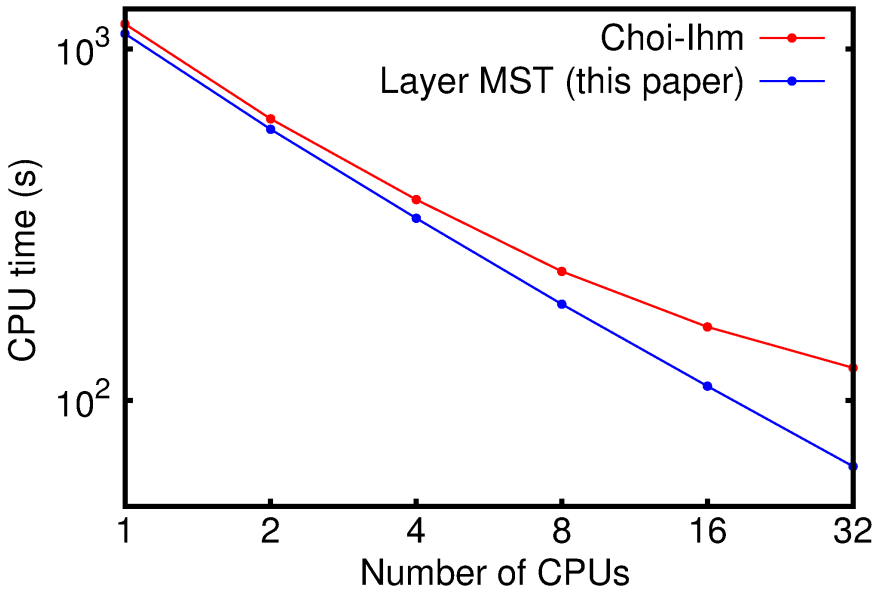

FIG. 9. (Color online) Calculated CPU time as a function of the number of CPUs. Both x and y axes are in logarithm scale.

Appendix D: Parallel efficiency

In order to compare the parallel efficiency of the layer multiple scattering theory method and the Choi-Ihm method, we perform a benchmark calculation using the system depicted in FIG . 2a with $d_\perp = 0$ and $d_\parallel = 3.5$ Å. The incident energy is 0.1 eV above the Dirac point of the electrode graphene. Setting $k_\perp = (1/3, 0)$, we calculate the corresponding transmission coefficient using both methods and record the CPU time per node as a function of the number of CPUs. Fig. 9 plots the calculated CPU time as a function of the number of the CPUs. The new method has a better parallel efficiency than the Choi-Ihm method. At 32 CPUs the two methods differ by about a factor of two.

* Email: cheng@qtp.ufl.edu; Tel: 352-392-6256

[1] A. K. Geim and K. S. Novoselov, Nat Mater. **6**, 183 (2007).
[2] A. K. Geim, Science **324**, 1530 (2009).
[3] P. Y. Huang, C. S. Ruiz-Vargas, A. M. van der Zande, W. S. Whitney, M. P. Levendorf, J. W. Kevek, S. Garg, J. S. Alden, C. J. Hustedt, Y. Zhu, J. Park, P. L. McEuen, and D. A. Muller, Nature. **469**, 389 (2011).
[4] K. Kim, Z. Lee, W. Regan, C. Kisielowski, M. F. Crommie, and A. Zettl, ACS Nano **5**, 2142 (2011).
[5] X. Li, C. W. Magnuson, A. Venugopal, J. An, J. W. Suk, B. Han, M. Borysiak, W. Cai, A. Velamakanni, Y. Zhu, L. Fu, E. M. Vogel, E. Voelkl, L. Colombo, and R. S. Ruoff, Nano Letters **10**, 4328 (2010).
[6] Y. Liu and B. I. Yakobson, Nano Letters **10**, 2178 (2010).
[7] O. V. Yazyev and S. G. Louie, Nat Mater. **9**, 806 (2010).
[8] A. Bagri, S.-P. Kim, R. S. Ruoff, and V. B. Shenoy, Nano Letters **11**, 3917 (2011).
[9] L. A. Jauregui, H. Cao, W. Wu, Q. Yu, and Y. P. Chen, Solid State Communications **151**, 1100 (2011).
[10] A. W. Tsen, L. Brown, M. P. Levendorf, F. Ghahari, P. Y. Huang, R. W. Havener, C. S. Ruiz-Vargas, D. A. Muller, P. Kim, and J. Park, Science **336**, 1143 (2012).
[11] O. V. Yazyev and S. G. Louie, Phys. Rev. B **81**, 195420 (2010).
[12] A. W. Robertson, A. Bachmatiuk, Y. A. Wu, F. Schffel, B. Rellinghaus, B. Bchner, M. H. Rmmeli, and J. H. Warner, ACS Nano **5**, 6610 (2011).
[13] L. Britnell, R. Gorbachev, R. Jalil, B. Belle, F. Schedin, A. Mishchenko, T. Georgiou, M. Katsnelson, L. Eaves, S. Morozov, *et al.*, Science **335**, 947 (2012).
[14] S. Seo, M. Min, S. M. Lee, and H. Lee, Nature communications **4**, 1920 (2013).
[15] R. Landauer, IBM Journal of Research and Development **1**, 223 (1957).
[16] M. Buttiker, IBM Journal of Research and Development **32**, 317 (1988).
[17] V. Mujica, M. Kemp, and M. A. Ratner, The Journal of Chemical Physics **101**, 6849 (1994).
[18] J. Taylor, H. Guo, and J. Wang, Phys. Rev. B **63**, 245407 (2001).
[19] M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, Phys. Rev. B **65**, 165401 (2002).
[20] Y. Xue, S. Datta, and M. A. Ratner, Chemical Physics **281**, 151 (2002).
[21] C. Zhang, X.-G. Zhang, P. S. Krstić, H.-p. Cheng, W. H. Butler, and J. M. MacLaren, Phys. Rev. B **69**, 134406 (2004).
[22] H. Joon Choi and J. Ihm, Phys. Rev. B **59**, 2267 (1999).
[23] H. J. Choi, J. Ihm, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **84**, 2917 (2000).
[24] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, Journal of Physics: Condensed Matter **21**, 395502 (2009).
[25] A. Smogunov, A. Dal Corso, and E. Tosatti, Phys. Rev. B **70**, 045417 (2004).
[26] M. A. Kuroda, J. Tersoff, D. M. Newns, and G. J. Martyna, Nano Letters **11**, 3629 (2011).
[27] O. V. Yazyev and A. Pasquarello, Phys. Rev. B **80**, 035408 (2009).
[28] J. P. Velev, C.-G. Duan, K. D. Belashchenko, S. S. Jaswal, and E. Y. Tsymbal, Phys. Rev. Lett. **98**, 137201 (2007).
[29] G. Sclauzero, A. Dal Corso, and A. Smogunov, Phys. Rev. B **85**, 165412 (2012).
[30] J. M. MacLaren, X.-G. Zhang, W. H. Butler, and X. Wang, Phys. Rev. B **59**, 5470 (1999).
[31] T.-H. Kim, X.-G. Zhang, D. M. Nicholson, B. M. Evans, N. S. Kulkarni, B. Radhakrishnan, E. A. Kenik, and A.-P. Li, Nano letters **10**, 3096 (2010).
[32] J. Zhang, Y. Wang, X.-G. Zhang, and X. F. Han, Phys. Rev. B **82**, 134449 (2010).
[33] Y. Wang, J. Zhang, X.-G. Zhang, H.-P. Cheng, and X. F. Han, Phys. Rev. B **82**, 054405 (2010).
[34] X.-G. Zhang, K. Varga, and S. T. Pantelides, Phys. Rev. B **76**, 035108 (2007).
[35] M. K. Srivastava, Y. Wang, X.-G. Zhang, D. M. C. Nicholson, and H.-P. Cheng, Phys. Rev. B **86**, 075134 (2012).
[36] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[37] A. Smogunov, A. Dal Corso, and E. Tosatti, Phys. Rev. B **70**, 045417 (2004).
[38] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[39] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light* (Oxford, Pergamon Press, 1964).
[40] D. Y. K. Ko and J. R. Sambles, J. Opt. Soc. Am. A **5**, 1863 (1988).
[41] N. P. K. Cotter, T. W. Preist, and J. R. Sambles, J. Opt. Soc. Am. A **12**, 1097 (1995).
[42] J. A. Wheeler, Phys. Rev. **52**, 1107 (1937).
[43] E. Merzbacher, *Quantum Mechanics, Chap.6, Sec. 6* (Wiley, New York, 1970).
[44] J. MacLaren, S. Crampin, D. Vvedensky, R. Albers, and J. Pendry, Computer Physics Communications **60**, 365 (1990).
[45] X.-G. Zhang, K. Varga, and S. T. Pantelides, Phys. Rev. B **76**, 035108 (2007).
[46] A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, Phys. Rev. B **41**, 1227 (1990).
[47] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[48] Y.-W. Son, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **97**, 216803 (2006).
[49] Y.-P. Wang, J. N. Fry, and H.-P. Cheng, Phys. Rev. B **88**, 125428 (2013).
[50] W. Butler, X.-G. Zhang, and J. MacLaren, Journal of Superconductivity **13**, 221 (2000).
[51] P. Lichtenberger, O. Morandi, and F. Schürrer, Phys. Rev. B **84**, 045406 (2011).
[52] C. R. Garibotti and G. Spiga, Journal of Physics A: Mathematical and General **27**, 2709 (1994).
[53] D. S. Fisher and P. A. Lee, Phys. Rev. B **23**, 6851 (1981).
[54] C. Cao, Y. Wang, H.-P. Cheng, and J.-Z. Jiang, Applied Physics Letters **99**, 073110 (2011).
[55] S. B. Trickey, F. Müller-Plathe, G. H. F. Diercksen, and J. C. Boettger, Phys. Rev. B **45**, 4460 (1992).