Accurate modeling of defects in graphene transport calculations

Lukas Linhart, Joachim Burgdörfer, and Florian Libisch
Institute for Theoretical Physics, Vienna University of Technology, 1040 Vienna, Austria, EU

(Received 10 October 2017; published 22 January 2018)

We present an approach for embedding defect structures modeled by density functional theory into large-scale tight-binding simulations. We extract local tight-binding parameters for the vicinity of the defect site using Wannier functions. In the transition region between the bulk lattice and the defect the tight-binding parameters are continuously adjusted to approach the bulk limit far away from the defect. This embedding approach allows for an accurate high-level treatment of the defect orbitals using as many as ten nearest neighbors while keeping a small number of nearest neighbors in the bulk to render the overall computational cost reasonable. As an example of our approach, we consider an extended graphene lattice decorated with Stone-Wales defects, flower defects, double vacancies, or silicon substitutes. We predict distinct scattering patterns mirroring the defect symmetries and magnitude that should be experimentally accessible.

DOI: 10.1103/PhysRevB.97.035430

I. INTRODUCTION

The first experimental realization of graphene by Novoselov *et al.* [1] in 2004 opened a new field of physics, the field of two-dimensional (2D) materials. Graphene consists of a single layer of carbon atoms arranged in a hexagonal honeycomb lattice featuring spectacular properties, with the manifestation of an anomalous quantum Hall effect at room temperature, extremely high electron mobility, and extraordinary mechanical strength being among the most prominent ones [2–4]. Possible new devices based on graphene cover a broad range from spintronics [5], valleytronics [6], and highly accurate sensors [7] to terahertz clocked field-effect transistors and stacked 2D material photovoltaics [8,9].

While new chemical-vapor-deposition growing techniques produce exceptionally clean flakes [10], any real device almost inevitably features imperfections due to defects and the finite size of the lattice. Changes in properties of graphene due to imperfections are crucial for the assessment of the performance and functionality of future devices, and therefore the understanding of such changes is key to engineering graphene-based devices. At the same time, it may become possible to tailor device properties by systematically placing defects. A variety of defect structures in graphene have already been observed and also theoretically predicted, including topological defects such as Stone-Wales and flower defects [11–17].

Beyond graphene, other emerging two-dimensional materials such as hexagonal boron nitride and transition-metal dichalcogenides as well as silicon-based new materials can be strongly tuned by local defects [18,19]. Modeling large-scale devices fabricated with such materials thus requires a flexible yet robust approach for accurate modeling of local defects. Full density-functional-theory (DFT)-based implementations such as the TranSIESTA code [20,21] and Wannierization approaches [22] for treating defects require us to treat the entire structure on the same level of sophistication. For narrow graphene nanoribbons the influence of a single localized defect on transport in graphene has been theoretically investigated for a variety of defects (e.g., Stone-Wales [23], boron impurities [24]) using these Wannierization methods. However, for larger devices with linear dimensions above 100 nm a full treatment of the entire geometry on the DFT level is hardly feasible. An approach to overcome this limitation consists of splitting a large geometry into building blocks of structures treatable by DFT [25]. This accurate approach has been successfully applied to defects and edge effects [26–28] yet includes most of the complexity of the Kohn-Sham Hamiltonian in the entire scattering geometry. Consequently, large-scale calculations remain numerically prohibitively expensive. Alternatively, tight-binding (TB) calculations of large graphene flakes provide detailed insights into transport properties [29,30]. Closely related density-functional-theory tight-binding (DFTB) [31] methods rely on extensive Slater-Koster parametrizations to describe the interactions between orbitals of neighboring atoms. More complicated defects or different hybridizations pose a considerable challenge for finding a suitably general parametrization. The goal of the present work is to combine the full DFT treatment of the defect with the TB treatment of the extended pristine 2D material by introducing a systematic algorithm for a locally varying parametrization. To achieve this goal we embed a local DFT treatment of a supercell containing around 50 atoms in the vicinity of the defect in a large-scale tight-binding calculation for the defect-free parts of the structure. In such a way, we are able to treat devices containing several millions of atoms. For one application we model an experiment where two scanning tunneling microscope (STM) tips approach a free-hanging graphene sheet to measure the conductance currently carried out by Eder *et al.* [32,33]. We benchmark our embedding technique for a variety of defects and discuss possible caveats and extensions.

II. EMBEDDING TECHNIQUE

We consider a tight-binding Hamiltonian of the form

$$H = \sum_{\mathrm{i}} \epsilon_i c_i^\dagger c_i + \sum_{\langle \mathrm{i,j} \rangle} \gamma_{i,j} c_i^\dagger c_j, \tag{1}$$

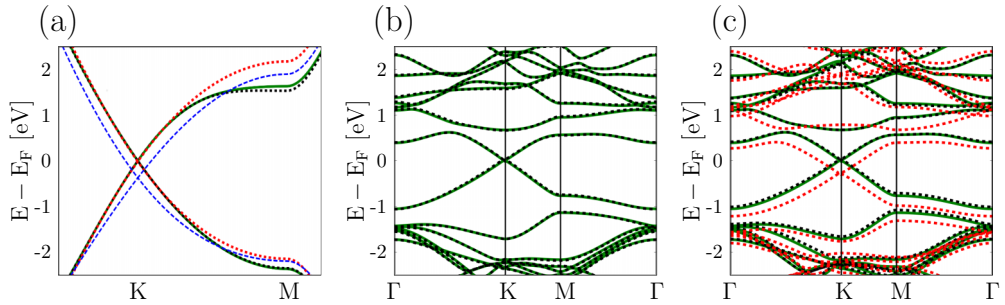

FIG. 1. (a) Band structures of bulk graphene around the K point. The band structures calculated with DFT (solid green line), with $\mathcal{S}^3_{\text{fit}}$ (dashed red line), with $\mathcal{S}^3_{\text{W}}$ (dashed blue line), and with $\mathcal{S}^{10}_{\text{W}}$ (dashed black line) are shown. (b) and (c) The band structure of a supercell consisting of 50 atoms with a Stone-Wales defect. Band structures shown are calculated from DFT [solid green line in (b) and (c)] with $\mathcal{S}^{10}_{\text{W}}$ [dashed black line in (b)], with $\mathcal{S}^{8}_{\text{W}}$ [dashed black line in (c)], and with $\mathcal{S}^{3}_{\text{W}}$ [dashed red line in (c)] NN description.

with c_{i} ($c_{\text{i}}^{\dagger}$) being the annihilation (creation) operators, ε_{i} being the diagonal on-site matrix elements, and $\gamma_{\text{i,j}}$ being the coupling elements between sites i and j. For a periodic, defect-free lattice we group the couplings $\gamma_{\text{i,j}}$ according to the distance between sites i and j and include only couplings to the n closest such neighbors. For example, first-nearest-neighbor tight binding ($n = 1$) neglects all interactions except those with direct neighbors. Application of Eq. (1) requires a judicious choice for the set of input parameters $\mathcal{S}^n = \{\epsilon_{\text{i}}, \gamma_{\text{i-j}}\}$ for a given n. Obviously, the larger the chosen n is, the more faithfully properties such as the band structure from *ab initio* theory or experiment are reproduced. However, a larger n also entails a much larger number of nonzero matrix elements in the tight-binding Hamiltonian and thus an increased numerical effort.

For the case of graphene investigated here, only π band orbitals contribute to transport properties [34]. We therefore restrict the tight-binding Hamiltonian to an orthogonal set of π states describing this band. A very common and straightforward approach is to calculate the underlying band structure employing an *ab initio* approach such as DFT [35,36]. The tight-binding parameters can then be obtained either by fitting the resulting TB bands to the DFT band structure or by first transforming the DFT description into a strongly localized basis from which the parameter set $\mathcal{S}^n$ can be directly deduced. Taking into account the lattice symmetries of bulk graphene only n γ_{i} need to be fitted in an nth-nearest-neighbor (NN) description; consequently, a fitting approach works well. We use the Levenberg-Marquardt algorithm [37,38] to fit a number of parameter sets $\mathcal{S}^n_{\text{fit}}$ ($n = 3, \ldots, 10$) to the bulk band structure.

With a higher number of NNs taken into consideration the accuracy of the tight-binding description increases. When focusing on the part of the band structure containing the Dirac cone near the K point (corresponding to the energy interval ± 1.2 eV around the Fermi energy), third-nearest-neighbor tight binding ($n = 3$) is sufficient for an accurate description [see Fig. 1(a)] [34], defining the quasiparticle energy range for which our bulk tight-binding model is valid.

A defect breaks the lattice symmetry and results in local deformations. Consequently, a much larger number of tight-binding parameters need to be fitted. As an example, consider a 6×6 supercell of graphene containing one double vacancy [see Fig. 2(a)]. Even when accounting for all residual preserved symmetries a third-NN description requires roughly 260 independent parameters. Finding a unique optimal parametrization by fitting to a DFT band structure therefore becomes unfeasible. Instead, an algorithm that allows the direct calculation of the coupling parameters from DFT is required. Within a periodic Kohn-Sham calculation, the orbitals are delocalized Bloch wave functions, with no obvious connection to the position of atoms. Obtaining tight-binding parameters thus requires a basis transformation to a strongly localized set of orbitals $\{\phi_{\text{i}}^{\text{loc}}\}$. We use WANNIER90 as a computational tool [39] to transform the plane-wave DFT basis into a strongly localized Wannier basis [40,41]. The matrix elements of the Hamiltonian $H^{\text{W}}_{\text{i,j}}$ described in the Wannier basis (W),

$$H^{\text{W}}_{\text{i,j}} = \langle \psi_{\text{i}}^{\text{W}} | \mathcal{H}^{\text{DFT}} | \psi_{\text{j}}^{\text{W}} \rangle = \begin{cases} \gamma^{\text{W}}_{\text{i,j}} & i \neq j, \\ \epsilon^{\text{W}}_{\text{i}} & i = j, \end{cases} \tag{2}$$

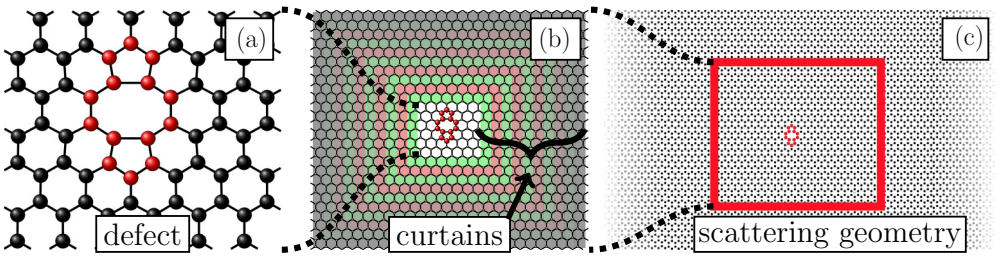

FIG. 2. (a) Example of a DFT-calculated structure consisting of 70 atoms with a double vacancy at the atoms indicated in red. (b) DFT-calculated defect geometry surrounded by a number of layers (highlighted by green and red) of primitive unit cells (curtains). The bracket indicates the spacing between the defect and the bulk where curtains are attached. (c) Schematic illustration of a scattering geometry within which the embedded defect is finally placed.

TABLE I. Tight-binding parameter set values obtained from fitting (except $\mathcal{S}_{\mathrm{W}}^{10}$, which is taken from the Wannier calculation and used as an initial guess). $\mathcal{S}_{\mathrm{fit}}^{\mathrm{i}}$ is the parameter set that accounts for up to i NN parameters. The set presented here is optimized in the energy range ± 1.2 eV around the Fermi level. The last row gives the averaged mean-square variations of the parameter values from one parameter set $\mathcal{S}_{\mathrm{fit}}^{i}$ to its neighboring set $\mathcal{S}_{\mathrm{fit}}^{i-1}$.

	$\mathcal{S}_{\mathrm{W}}^{10}$ (eV)	$\mathcal{S}_{\mathrm{fit}}^{9}$ (eV)	$\mathcal{S}_{\mathrm{fit}}^{8}$ (eV)	$\mathcal{S}_{\mathrm{fit}}^{7}$ (eV)	$\mathcal{S}_{\mathrm{fit}}^{6}$ (eV)	$\mathcal{S}_{\mathrm{fit}}^{5}$ (eV)	$\mathcal{S}_{\mathrm{fit}}^{4}$ (eV)	$\mathcal{S}_{\mathrm{fit}}^{3}$ (eV)
ϵ_0	0.29697	0.29499	0.29464	0.30003	0.30323	0.29570	0.08835	0.07893
γ_1	−2.91164	−2.92326	−2.95482	−3.03528	−3.11753	−3.14029	−3.09637	−3.28542
γ_2	0.22317	0.22675	0.22625	0.23194	0.23803	0.23385	0.03508	0.03186
γ_3	−0.28900	−0.29228	−0.28632	−0.29412	−0.30412	−0.31542	−0.29678	−0.37544
γ_4	0.02480	0.02503	0.02445	0.02512	0.02598	0.02730	0.03617	
γ_5	0.05493	0.05380	0.05358	0.05532	0.05752	0.06778		
γ_6	−0.02232	−0.02161	−0.02163	−0.02235	−0.02321			
γ_7	−0.01295	−0.01257	−0.01260	−0.01299				
γ_8	−0.02181	−0.02115	−0.02111					
γ_9	0.00716	0.00695						
γ_{10}	0.00400							
$\|\|\mathcal{S}_{\mathrm{fit}}^{i} - \mathcal{S}_{\mathrm{fit}}^{i-1}\|\|_2 / i$	0.00134	0.00365	0.01049	0.01202	0.00615	0.05982	0.05204	

are then just the tight-binding parameters $\gamma_{\mathrm{i,j}}^{\mathrm{W}}$ ($\epsilon_{\mathrm{i}}^{\mathrm{W}}$) in Wannier representation. We denote the corresponding set of the n' NN parameters by $\mathcal{S}_{\mathrm{W}}^{n'}$. *A priori*, there is no one-to-one relation between the band-structure-fitted set $\mathcal{S}_{\mathrm{fit}}^{n}$ and the Wannier representation $\mathcal{S}_{\mathrm{W}}^{n'}$. In particular the orders of the included NN couplings, to represent a given segment of the pristine bulk band structure in the absence of defects with a preserved level of accuracy, may not coincide ($n \neq n'$). Indeed, a fitted set of third-NN parameters $\mathcal{S}_{\mathrm{fit}}^{3}$ resembles the band structure of pristine graphene far better than the corresponding third-NN Wannier-based parameters $\mathcal{S}_{\mathrm{W}}^{3}$ [see Fig. 1(a)]. The reason is simple: the effective parameters in $\mathcal{S}_{\mathrm{fit}}^{3}$ are chosen such that the missing higher-neighbor interactions are included on a mean-field level, whereas such interactions are simply neglected in $\mathcal{S}_{\mathrm{W}}^{3}$. For a variety of defect supercells in graphene we find that a tenth-NN description resembles the defect band structure exceptionally well [see Fig. 1(b)] while lower-order NN descriptions based on Wannier functions lead, in general, to significant deviations [see Fig. 1(c)].

While an accurate description of moderate-sized structures on the level of tenth NN is readily possible, this straightforward approach has important drawbacks: the large number of NN interactions substantially reduces the sparsity of the Hamiltonian, leading to a much increased computational effort. Moreover, schemes that exploit the locality of the Hamiltonian, e.g., the modular recursive Green's function method [42,43], lose much of their efficiency.

An optimized strategy is thus to combine two ingredients: an accurate tenth-NN description $\mathcal{S}_{\mathrm{W}}^{10}$ of the local defect structure and a lower-order $\mathcal{S}_{\mathrm{fit}}^{3}$ description for the unperturbed regions of bulk graphene. We interface the two approaches by introducing a smooth (quasicontinuous) transition region (see Fig. 2). The key insight is that for pristine graphene an accurate fitted parametrization $\mathcal{S}_{\mathrm{fit}}^{n}$ can be achieved for variable n, $3 \leqslant n \leqslant 10$. For the maximum value employed, we have verified that to a high degree of accuracy $\mathcal{S}_{\mathrm{fit}}^{10} \approx \mathcal{S}_{\mathrm{W}}^{10}$ [see Fig. 1(a)]. Now successively dropping the most distant interaction and refitting the remaining parameters to the DFT band structure (see Table I for numerical values) ensures that values between adjacent parameter sets $\mathcal{S}_{\mathrm{fit}}^{n}$ and $\mathcal{S}_{\mathrm{fit}}^{n+1}$ are compatible. We can now smoothly change the number of NN interactions in the transition region: starting from a tenth-NN ($\mathcal{S}_{\mathrm{W}}^{10}$) description for the defect region, sequential layers (curtains) wrapped around the structure contain a decreasing number of interactions ($n = 9,8,7,\ldots,3$), until the outermost layer is described using the $\mathcal{S}_{\mathrm{fit}}^{3}$ bulk parametrization [see Fig. 2(b)]. Interactions $\gamma_{\mathrm{i,j}}$ between orbitals i and j where i and j both lie inside the defect supercell are described by $\mathcal{S}_{\mathrm{W}}^{10}$, while interactions between a defect supercell orbital and a curtain orbital in curtain n use $\mathcal{S}_{\mathrm{fit}}^{n}$. Interactions between orbital sites lying in different curtains described by different parameter sets $\mathcal{S}_{\mathrm{fit}}^{n}$ and $\mathcal{S}_{\mathrm{fit}}^{m}$ use parameters from the parameter set $\mathcal{S}_{\mathrm{fit}}^{\min(n,m)}$.

The presented embedding method implicitly requires a sufficiently smooth transition between adjacent parameter sets $\mathcal{S}_{\mathrm{fit}}^{i}$ and $\mathcal{S}_{\mathrm{fit}}^{i+1}$. The smoothness can be probed by the effective suppression of reflection at the transition region. For the test cases considered here we can compare the results of the present embedding algorithm with a full $\mathcal{S}_{\mathrm{W}}^{10}$ parametrization for the entire scattering geometry (discussed below). With the fitted tight-binding parameters in Table I and the sizes of our defect supercells the transmission for the employed parameter sets indeed agrees well with the $\mathcal{S}_{\mathrm{W}}^{10}$ benchmark. The relative parameter shifts between $\mathcal{S}_{\mathrm{fit}}^{\mathrm{i}}$ and $\mathcal{S}_{\mathrm{fit}}^{\mathrm{i+1}}$ are, on average, on the order of 10^{-2} (see Table I, last row). We find remarkably good agreement in all cases.

III. BENCHMARK CALCULATIONS

As a first test and application of our embedding approach we consider ballistic transport through a graphene zigzag nanoribbon [Fig. 3(a)]. We perform DFT calculations using VASP [44–47] in the local-density approximation (LDA) together with a Monkhorst-Pack grid for k-point sampling of $25 \times 25 \times 1$. The Wannierization is performed using p_z orbitals at each atomic lattice site and s orbitals in the center of each C-C bond as initial projections for the disentanglement procedure needed in the WANNIER90 code [39]. Having obtained tight-binding parameters, we employ the Landauer-Büttiker formalism [48] for transport calculations.

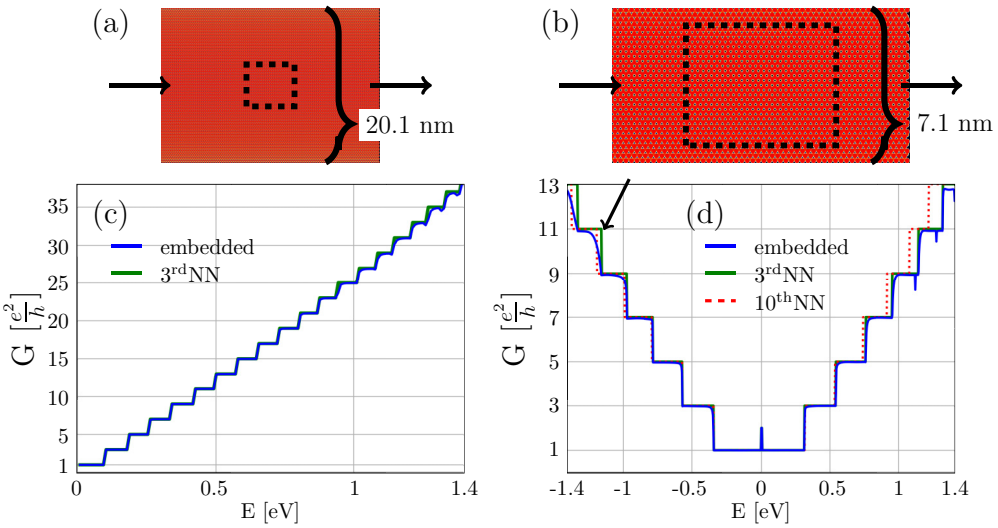

FIG. 3. Transmission benchmark calculations. Graphene ribbons with a size of (a) 20.1 and (b) 7.1 nm with the position and size of a defect-free embedded structure equivalent to that in Fig. 2. The arrows indicate the direction of transport. (c) Conductance of the ribbon shown in (a) calculated using a defect-free supercell described by $\mathcal{S}_{\mathrm{W}}^{10}$ embedded in a $\mathcal{S}_{\mathrm{fit}}^{3}$ bulk (solid blue line) and without embedding using $\mathcal{S}_{\mathrm{fit}}^{3}$ (solid green line). (d) Same as (c) for the narrower ribbon shown in (b). The dashed red line corresponds to a calculation using $\mathcal{S}_{\mathrm{W}}^{10}$ for the entire geometry. The black arrow indicates an example of the expanded mode opening region.

We first consider a uniform, defect-free zig-zag ribbon that should feature perfect transmission. We thus expect the transmission as a function of the energy above the Dirac point to be a staircase function with steps of height $\frac{2e^2}{\mathrm{h}}$ (not including physical spin) whenever a new transverse mode opens in the ribbon. Indeed, using $\mathcal{S}_{\mathrm{fit}}^{3}$ for the entire geometry results in a perfect staircase function [see Fig. 3(a)]. For a comparison we use our embedding approach where a small 5×5 supercell parameterized with the set $\mathcal{S}_{\mathrm{W}}^{10}$ is embedded into the $\mathcal{S}_{\mathrm{fit}}^{3}$ bulk using the curtain sequence outlined above. Together with the curtains the embedded structure has a width of 5.2 nm. Placing this embedded structure in a graphene ribbon of width 20.1 nm, we find almost no deviation from the perfect transmission curves up to energies ± 1.2 eV around the Fermi energy for which the bandstructure is faithfully reproduced [see Fig. 3(c), solid green and blue lines]. As we expect the influence of embedding to be largest for narrow ribbons, we consider a second benchmark ribbon which is only slightly wider than the defect structure [7.1 nm; see Fig. 3(b)]. Even for this extreme case where most of the ribbon consists of the embedded structure, the transmission curves of the ribbon with the embedded structure (solid green line) and without the embedded structure (solid blue line) agree very well for energies up to ± 1.2 eV. A comparison between a full $\mathcal{S}_{\mathrm{fit}}^{3}$ [solid blue line in Fig. 3(d)] and $\mathcal{S}_{\mathrm{W}}^{10}$ [dashed red line in Fig. 3(d)] treatment of the entire geometry shows deviations for energies $|E| > 1.2$ eV due to the different energies at which the lead modes open. For the embedded case the same effect leads to backscattering at the boundaries between different curtains as indicated, e.g., with a black arrow in Fig. 3. While, e.g., $\mathcal{S}_{\mathrm{fit}}^{3}$ predicts the opening of a new mode in the lead at -1.15 eV, the $\mathcal{S}_{\mathrm{W}}^{10}$ description used in the center of the embedded region allows for a new mode only at -1.19 eV, leading to backscattering in the energy window between the two thresholds. These problems emerge because $\mathcal{S}_{\mathrm{fit}}^{3}$ does not provide enough degrees of freedom to accurately reproduce the band structure over a broad energy region. To systematically improve this result one could use optimized parameter sets $\tilde{\mathcal{S}}_{\mathrm{fit}}^{n}$ for different energy regions of interest, correcting the small differences in mode opening energies. However, since the deviations are small compared to any defect influence and occur only at higher energies, we avoid such adjustments. Finally, in the embedded case the mode opening energies for different curtains may be slightly misaligned, causing scattering into modes blocked by inner curtains. Such an effect leads to sharp dips in the conductance immediately before a mode opening [49] as observed at 1.12 and at 1.27 eV in Fig. 3. With such limitations in mind, the present embedding method is found to be quite efficient and accurate.

As a second benchmark system we model the setup of the previously mentioned STM experiment [33]: a large, freestanding graphene membrane is contacted via an STM tip at site $\vec{r}_S$. The probability $|\psi_{\vec{r}_S}(\vec{r})|^2$ of finding an electron at site $\vec{r}$ is evaluated via the Green's function as

$$\psi_{\vec{r}_S}(\vec{r}) = G(\vec{r},\vec{r}_S). \tag{3}$$

Technically, we evaluate G on a rectangular sheet of graphene (54×62 nm; see Fig. 4) consisting of roughly 10^5 atoms with open boundary conditions [50] on all four sides emulating a large-scale graphene flake. Structures of this size are far beyond any size accessible by DFT but can be treated by a TB calculation involving the present embedding scheme. The obtained scattering probability distribution is, in principle, directly experimentally accessible by a second STM tip acting as a probe.

We compare probability distributions for a pristine flake with parametrization $\mathcal{S}_{\mathrm{fit}}^{3}$ [i.e., without embedding; Fig. 4(b)] and in the presence of a defect-free embedded region of $\mathcal{S}_{\mathrm{W}}^{10}$ [Fig. 4(c)] and find almost no difference. The sixfold scattering symmetry resulting in a starlike pattern [Figs. 4(a), 4(b) and 4(f)] is due to trigonal warping caused by the deviation from

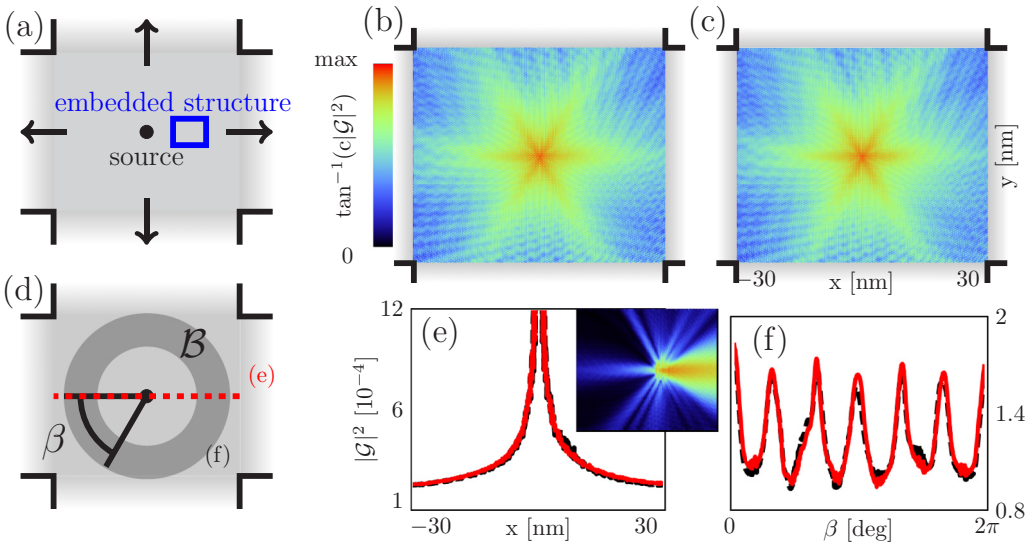

FIG. 4. Benchmark calculations. (a) Schematic of the geometry with an electron source in the center and a defect-free embedded structure placed as indicated by the rectangle. Open boundary directions are highlighted by black arrows; the black corners indicate the remaining hard-wall boundaries. Scattering probabilities (b) without embedding and (c) with embedding when an electron carrying 1 eV energy is added. (d) Schematic of the geometry used in plots (e) and (f). The dark gray region $\mathcal{B}$ indicates a ring with a radius between 12 and 18 nm from the source position r_S. (e) Plot of (b) and (c) along the cut marked by the dotted red line in (d). The inset shows the difference plot between (b) and (c), scaled by 10^2. (f) Angle-dependent plot of (b) and (c) with the angle β defined as in (d). The solid red line refers to (b), and the dashed black line refers to (c). The lines in (f) show the scattering probability from the source position r_S into region $\mathcal{B}$ at an angle $\beta \pm 0.05\pi$. A Gaussian smearing with $\sigma = 0.05$ is used.

the perfect conelike dispersion relation of graphene at higher energies [51].

The good agreement between embedded and embedding-free descriptions for both defect-free benchmark systems clearly showcases the validity of our embedding method: the slow variations in the tight-binding parameters caused by switching from a highly accurate, local $\mathcal{S}_W^{10}$ description to a bulk $\mathcal{S}_{fit}^3$ fit results in negligible artificial scattering for the range of energies for which the TB parametrization is valid. The embedding approach promises a fast and accurate defect description scheme to simulate transport in large flakes that combines the advantages of purely DFT based calculations with those of tight-binding schemes.

IV. DEFECTS

We now apply the embedding scheme to four different types of defects: a double vacancy, a silicon interstitial, a Stone-Wales defect, and a flowerlike closed-loop grain boundary (see insets in the top row of Fig. 5). For the DFT calculations we use supercells containing 32, 50, 70, and 128 atoms, respectively. The disentanglement and Wannierization of such large cells have to be done with great care and become more challenging with increasing size of the cells. Supercell sizes for the calculated defects were therefore kept reasonably small. These cell sizes may not fully account for very long range lattice deformations, i.e., dynamic relaxation of the lattice. To estimate the associated error, we perform benchmark calculations using larger cells (e.g., 126 instead of 70 atoms for the Stone-Wales defect), which show that the change in atomic positions are, on average, smaller than 1% of the lattice constant. The resulting error is thus likely quite small.

Again, VASP together with an LDA pseudopotential and a Monkhorst grid sampling is used with $5 \times 5 \times 1$ k points for the silicon and the Stone-Wales defects [Figs. 5(a) and 5(d)] and $3 \times 3 \times 1$ k points for the double vacancy and the flowerlike defect [Figs. 5(b) and 5(c)]. We place all four defects in both our benchmark geometries and evaluate the transmission (top row in Fig. 5) and the scattering probability distributions (middle row in Fig. 5). We compare to a full calculation using $\mathcal{S}_W^{10}$ in the entire transport structure (see Fig. 5, top row). Even for our relatively small benchmark systems, the runtime of the embedding calculation is, compared to the full-scale calculation, a factor of ∼16 faster, clearly highlighting the benefit of an efficient embedding algorithm to treat large-scale devices. Obviously, the exact runtime speedup strongly depends on the geometry of the structure, on the number and size of the embedded defects, and on the numerical implementation.

The various defects have strongly differing influence on transport properties. The very weak scattering at the double vacancy can be understood as a consequence of the conservation of the AB-sublattice symmetry that suppresses backscattering [49]. For the Stone-Wales defect we find some change in the transmission function around $E = 0.5$ eV [Fig. 5(c)]. The largest modifications are induced by the flowerlike defect [see Fig. 5(c), panel (i)]. It has thus by far the largest influence on transport and thus the best chance to be uniquely identified in the experiment. One obvious reason for this defect to show up so prominently is its size: the flower defect is substantially larger than the other defects discussed here, most importantly, much larger than the de Broglie wavelength associated with the electron at the K point (2.5 Å). Consequently, scattering at

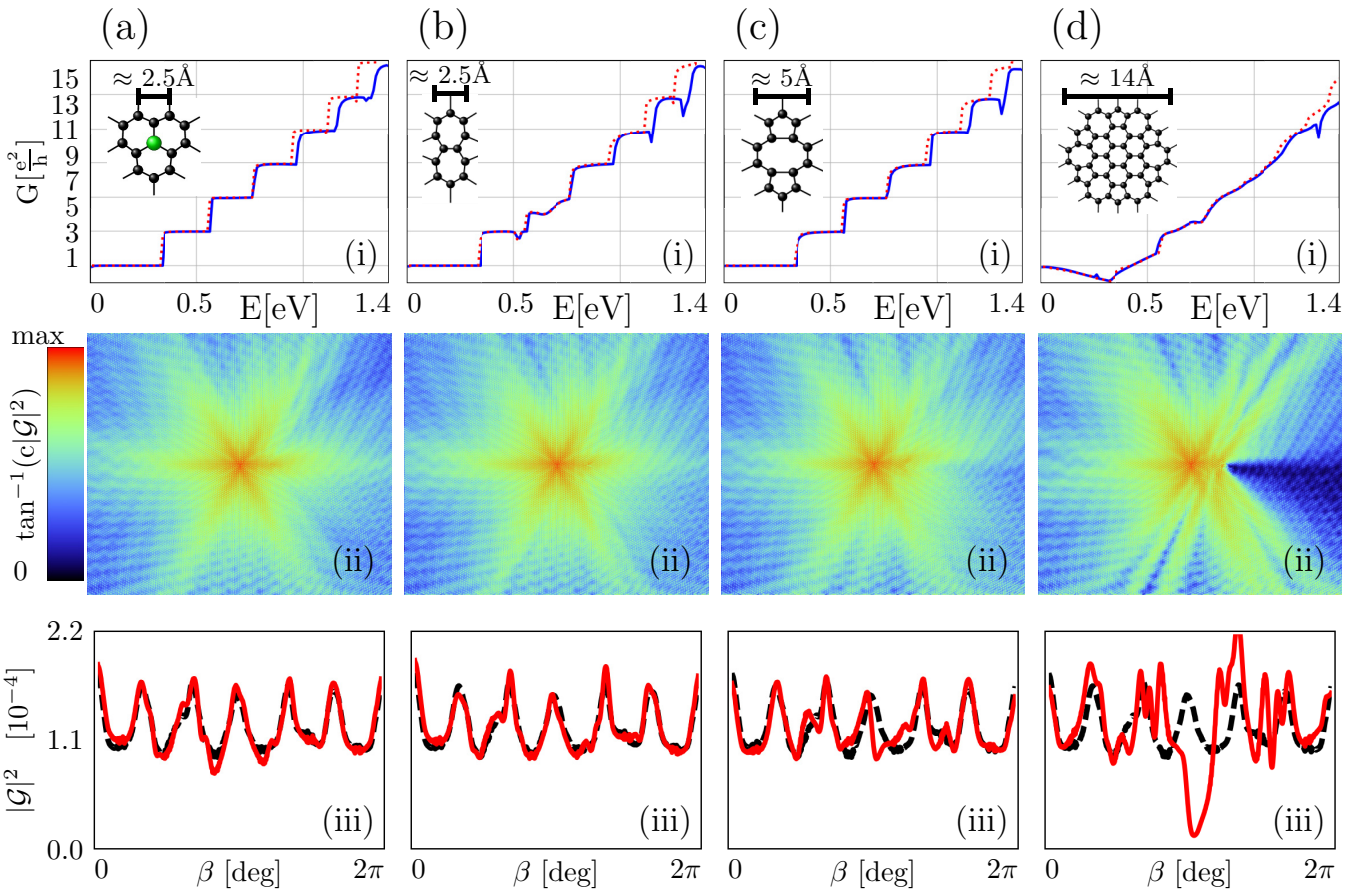

FIG. 5. Results for defects. Top: The conductance for four different embedded defect structures (solid blue line) using $\mathcal{S}^3$ in the bulk and without embedding (dashed red line) using $\mathcal{S}_W^{10}$ in the entire structure [geometry is as in Fig. 3(b)]. The insets show the defect types treated: (a) silicon interstitial, (b) Stone-Wales defect, (c) double vacancy, and (d) flowerlike grain boundaries. Middle: Scattering patterns for geometries as in Fig. 4 calculated at 1 eV electron energy with embedded defect structures as in Fig. 2 at the position indicated by the blue rectangle in Fig. 4(a). Bottom: Angle-resolved averaged scattering probabilities, with the angle and averaging area as defined in Fig. 4(d). The dashed black curves show defect-free scattering probabilities, and the red line shows the scattering probabilities in the presence of the corresponding defects.

this defect does not just lead to diffraction but casts a "shadow" leading to the pronounced minima in the scattering wave function behind it. For all investigated defect structures the transmission curves for the embedding and the full tenth-NN description match perfectly up to about 1.2 eV. Comparing the scattering patterns of different defects further corroborates these results: there are only minor changes compared to the defect-free calculation [Figs. 4(b) and 4(c)] for the double vacancy, the Stone-Wales defect, and the silicon interstitial (see Fig. 5, middle row). Only the flowerlike grain boundary defect causes significant backscattering, leading to a pronounced region of strongly suppressed transmission. A more quantitative comparison via angle-dependent scattering amplitudes [compare Fig. 4(d)] highlights the strong changes due to scattering at the flower defect (see Fig. 5, bottom row).

V. CONCLUSIONS

We have presented an embedding approach to incorporate *ab initio* tight-binding parameters from density functional theory into a large-scale tight-binding simulation. We combine a local tenth-nearest-neighbor description determined via projecting *ab initio* results onto Wannier functions with a lower-order tight-binding parametrization fitted to the pristine bulk band structure. Our approach allows the accurate treatment of large-scale devices containing local defects. We demonstrate our technique for four prototypical defects in graphene devices and find excellent agreement between our embedding approach and a computationally much more expensive full tenth-nearest-neighbor benchmark calculation. An extension of our method to extended one-dimensional defects such as grain boundaries and also to different materials seems straightforward and could lead to further generalizations. The calculated scattering patterns should become experimentally accessible by a double-tip STM setup. We therefore expect experimental tests in the near future [33].

ACKNOWLEDGMENTS

We acknowledge support from the FWF-SFB041 ViCom and the doctoral college TU-D funded by TU Wien. Numerical calculations were performed on the Vienna Scientific Cluster VSC2.

[1] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306**, 666 (2004).

[2] K. S. Novoselov, Z. Jiang, Y. Zhang, S. V. Morozov, H. L. Stormer, U. Zeitler, J. C. Maan, G. S. Boebinger, P. Kim, and A. K. Geim, Science **315**, 84173 (2007).

[3] K. I. Bolotin, K. J. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. L. Stormer, Solid State Commun. **146**, 351 (2008).

[4] C. Lee, X. Wei, J. W. Kysar, and J. Hone, Science **321**, 385 (2008).

[5] W. Han, APL Mater. **4**, 032401 (2016).

[6] J. R. Schaibley, H. Yu, G. Clark, P. Rivera, J. S. Ross, K. L. Seyler, W. Yao, and X. Xu, Nat. Rev. Mater. **1**, 16055 (2016).

[7] G. Lu, S. Park, K. Yu, R. S. Ruoff, L. E. Ocola, D. Rosenmann, and J. Chen, ACS Nano **5**, 1154 (2011).

[8] W. Han, R. K. Kawakami, M. Gmitra, and J. Fabian, Nat. Nanotechnol. **9**, 794 (2014).

[9] T. Mueller, F. Xia, and P. Avouris, Nat. Photonics **4**, 297 (2010).

[10] L. Banszerus, M. Schmitz, S. Engels, J. Dauber, M. Oellers, F. Haupt, K. Watanabe, T. Taniguchi, B. Beschoten, and C. Stampfer, Sci. Adv. **1**, e1500222 (2015).

[11] F. Banhart, J. Kotakoski, and A. V. Krasheninnikov, ACS Nano **5**, 26 (2011).

[12] J. T. Robinson, J. S. Burgess, C. E. Junkermeier, S. C. Badescu, T. L. Reinecke, F. K. Perkins, M. K. Zalalutdniov, J. W. Baldwin, J. C. Culbertson, P. E. Sheehan, and E. S. Snow, Nano Lett. **10**, 3001 (2010).

[13] V. M. Pereira, J. M. B. L. dos Santos, and A. H. C. Neto, Phys. Rev. B **77**, 115109 (2008).

[14] F. Libisch, S. Rotter, and J. Burgdörfer, Phys. Status Solidi B **248**, 2598 (2011).

[15] E. Cockayne, G. M. Rutter, N. P. Guisinger, J. N. Crain, P. N. First, and J. A. Stroscio, Phys. Rev B **83**, 195425 (2011).

[16] H. J. Park, V. Skákalová, J. Meyer, D. S. Lee, T. Iwasaki, C. Bumby, U. Kaiser, and S. Roth, Phys. Status Solidi B **247**, 2915 (2010).

[17] J. Kotakoski, C. Mangler, and J. C. Meyer, Nat. Commun. **5**, 3991 (2014).

[18] M. Chhowalla, Z. Liu, and H. Zhang, Chem. Soc. Rev. **44**, 2584 (2015).

[19] T. Markussen, R. Rurali, M. Brandbyge, and A.-P. Jauho, Phys. Rev. B **74**, 245313 (2006).

[20] M. Brandbyge, J.-L. Mozos, P. Ordejon, J. P. Taylor, and K. Stokbro, Phys. Rev. B **65**, 165401 (2002).

[21] K. Stokbro, J. Taylor, M. Brandbyge, and P. Ordejón, Ann. NY Acad. Sci. **1006**, 212 (2003).

[22] A. Calzolari, N. Marzari, I. Souza, and M. B. Nardelli, Phys. Rev. B **69**, 035108 (2004).

[23] H. Zeng, J.-P. Leburton, Y. Xu, and J. Wei, Nanoscale Res. Lett. **6**, 254 (2011).

[24] B. Biel, X. Blase, F. Triozon, and S. Roche, Phys. Rev. Lett. **102**, 096803 (2009).

[25] M. Shelley, N. Poilvert, A. Mostofi, and N. Marzari, Comput. Phys. Commun. **182**, 2174 (2011).

[26] G. Cantele, Y. Lee, D. Ninno, and N. Marzari, Nano Lett. **9**, 3425 (2009).

[27] Y.-S. Lee, M. B. Nardelli, and N. Marzari, Phys. Rev. Lett. **95**, 076804 (2005).

[28] A. R. Rocha, M. Rossi, A. Fazzio, and A. J. R. da Silva, Phys. Rev. Lett. **100**, 176803 (2008).

[29] I. Deretzis, G. Fiori, G. Iannaccone, and A. La Magna, Phys. Rev. B **81**, 085427 (2010).

[30] G. I. Márk, P. Vancsó, C. Hwang, P. Lambin, and L. P. Biró, Phys. Rev. B **85**, 125443 (2012).

[31] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, Phys. Rev. B **58**, 7260 (1998).

[32] M. Settnes, S. R. Power, D. H. Petersen, and A. P. Jauho, Phys. Rev. Lett. **112**, 096801 (2014).

[33] F. R. Eder, J. Kotakoski, K. Holzweber, C. Mangler, V. Skakalova, and J. C. Meyer, Nano Lett. **13**, 1934 (2013).

[34] S. Reich, J. Maultzsch, C. Thomsen, and P. Ordejon, Phys. Rev. B **66**, 035412 (2002).

[35] W. Hohenberg and P. Kohn, Phys. Rev. **136**, B864 (1964).

[36] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[37] K. Levenberg, Q. Appl. Math. **2**, 164 (1944).

[38] D. W. Marquardt, J. Soc. Ind. Appl. Math. **11**, 431 (1963).

[39] A. A. Mostofi, J. R. Yates, G. Pizzi, Y. S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, Comput. Phys. Comm. **185**, 2309 (2014).

[40] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[41] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2001).

[42] F. Libisch, S. Rotter, and J. Burgdörfer, New J. Phys. **14**, 123006 (2012).

[43] S. Rotter, J.-Z. Tang, L. Wirtz, J. Trost, and J. Burgdörfer, Phys. Rev. B **62**, 1950 (2000).

[44] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).

[45] G. Kresse and J. Hafner, Phys. Rev. B **49**, 14251 (1994).

[46] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).

[47] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[48] R. Landauer, IBM J. Res. Dev. **1**, 223 (1957).

[49] F. Libisch, A. Kliman, S. Rotter, and J. Burgdörfer, Phys. Status Solidi B **253**, 2366 (2016).

[50] F. Libisch, T. Hisch, R. Glattauer, L. Chizhova, and J. Burgdörfer, J. Phys. Condens. Mater. **29**, 114002 (2017).

[51] A. H. C. Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).