Impact of clustering of substitutional impurities on quasiparticle lifetimes and localization

Jack G. Nedell [1,2,*] Michael Vogl [3,4] and Gregory A. Fiete [1,5]

[1]*Department of Physics, Northeastern University, Boston, Massachusetts 02115, USA*
[2]*Department of Physics, Cornell University, Ithaca, New York 14853, USA*
[3]*Department of Physics, King Fahd University of Petroleum and Minerals, 31261 Dhahran, Saudi Arabia*
[4]*Interdisciplinary Research Center (IRC) for Intelligent Secure Systems, KFUPM, Dhahran, Saudi Arabia*
[5]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

(Received 27 March 2023; revised 21 June 2023; accepted 7 September 2023; published 18 September 2023)

Motivated by the observation and prediction of clustering behavior for impurities substituted into the host lattice of a real material and the dramatic impact this can have on electronic properties, we develop a simple approach to describe such an effect via the electron self-energy. We employ a disorder-averaged T-matrix expansion taken to second order, which we modify to include a clustering probability parameter. This approach circumvents the need for specific cluster probability distributions, simplifying greatly the analysis of clustered impurities. To gain analytical insights, we study a nearest-neighbor square lattice tight-binding Hamiltonian with clustered impurity substitutions to investigate clustering of off-diagonal hopping impurities. We find that our T-matrix approach is in excellent agreement with exact numerical results from a tight-binding computation performed with the KWANT package. We observe a variety of interesting impurity clustering-induced effects in the self-energy such as the suppression of quasiparticle lifetimes at certain momenta and an increase in localization, as indicated by the inverse participation ratio. The KWANT results are reproduced in our modified T-matrix approach. In addition, our method allows for a full analytical treatment of clustering effects which can aid in physical insight.

DOI: 10.1103/PhysRevB.108.104206

I. INTRODUCTION

The interplay between disorder and the electronic properties of two-dimensional (2D) materials [1] is of great interest. Of particular importance are the questions of how to control and minimize undesirable disorder effects and how to harness disorder to achieve desirable physical properties (e.g., quantized responses such as what occurs in the quantum Hall effect or quantum anomalous Hall effect).

An interesting yet underappreciated aspect of disorder in materials is the spatial configuration of impurities, which most often is not taken into account since a disorder average is typically employed. In a given material, do impurities cluster, or are they well separated? In general, it is nontrivial to find the energetically most favorable configuration of impurities, whether they are adsorbed atoms on a material surface or chemically substituted within a host material lattice. Various density functional theory studies were performed to address this question for numerous host materials and impurity types, for instance, in studies that investigate impurities substituted into [2] or adsorbed by [3] the surface or the bulk [4] of materials. Often, one finds that impurities show a preference for clustering with clusters of certain preferred sizes. The details of this effect depend on the host material, the impurity type, temperature, and the concentration of impurities.

Likewise, similar clustering behavior has been observed or induced experimentally. For instance, adatoms on the surface of graphene have been observed to emerge in clusters [3,5], and a recent study demonstrated the formation of large nitrogen clusters substituted into the graphene lattice, with a reported increase in conductivity [6].

There have also been theoretical studies of substitutional impurities in graphene [7–11] and studies of the effect of impurity clusters [12,13] on electronic properties. Substitutional impurities in materials are known to harbor localized states [14], which for graphene [9] has been suggested to play an important role in the transport properties of nanoribbons as transport varies dramatically with specific disorder configuration, relative to the edges of the system [7]. Another interesting effect is a recently predicted metal-to-insulator transition in disordered graphene [15].

Finally, it is important to contextualize this investigation in relation to earlier works. Studies on the theory of random binary alloys, which received heightened attention in the 1970s [16,17], were done in the framework of the coherent potential approximation (CPA). This previous work expanded on details of scattering by large clusters of impurities, including impurities with off-diagonal disorder, which will be our focus as well. In addition, early theoretical work on carrier lifetimes in disordered semiconductors identified that the impact of isolated impurities on carrier lifetimes could not readily be extrapolated to impurities which were clustered together [18]. These past and present research efforts set the stage for the investigation we present here.

Using a T-matrix expansion, we compute the self-energy of a nearest-neighbor tight-binding model on a square

*Corresponding author: jn479@cornell.edu

lattice with off-diagonal (hopping) impurities. We compare the results for an ordinary T-matrix description of isolated impurities to our T-matrix model, which allows us to take into account clustering of neighboring impurities. We pit these results against a disorder-averaged self-energy obtained from exactly solved finite systems realizing the same disorder model. We find that clustering contributes nontrivially to the self-energy, especially the imaginary part. We find our T-matrix approach is in excellent agreement with exact numerical results. Yet it requires almost no additional technical know-how beyond the standard T-matrix approach and thus is of great practical utility.

We proceed to use this self-energy to understand how electronic properties of the system vary with the concentration of impurity clustering. In particular, we compute the density of states, quasiparticle lifetime, and inverse participation ratio, which gives us an indication of localization properties. This analysis complements the results of previous studies, demonstrating increased localization and a decrease in the zero-energy density of states for increased clustering. These results are summarized in the figures.

Our paper is organized as follows. In Sec. II we introduce the tight-binding model, including disorder, on the square lattice that we will use for our study of impurity clustering. In Sec. III we introduce the T-matrix approach and present analytical results for isolated and clustered impurities. In Sec. IV we present numerical results for disorder effects related to the clustering of impurities. Finally, in Sec. V we present the main conclusions of our work.

II. MODEL

As a platform for our investigation of impurity clustering effects we choose a tight-binding model on a square lattice. This model consists of only nearest-neighbor hoppings. This choice of model has the advantage that much analytic progress is possible and will therefore allow for deeper insights into our treatment of impurity clustering. However, our main results are more general and do not depend on this specific Hamiltonian or lattice.

The Hamiltonian of the clean lattice may be written as a sum over the lattice sites of the model $\mathbf{r_i}$ and the nearest-neighbor displacement vectors $\boldsymbol{\delta}$,

$$\hat{H}_0 = -t \sum_{\mathbf{r}_i,\boldsymbol{\delta}} \hat{c}^\dagger(\mathbf{r}_i)\hat{c}(\mathbf{r}_i + \boldsymbol{\delta}). \tag{1}$$

Fourier transforming to a crystal momentum basis, one obtains the dispersion relation $\xi_{\mathbf{k}} = -2t[\cos(k_x a) + \cos(k_y a)]$, where a is the lattice spacing. We take $t = 1$ and $a = 1$.

Next, we introduce to this model a set of lattice site "substitutions." If the original lattice is made up of atoms of type A and we replace select lattice sites with atoms of type B, hoppings between atoms are no longer only $t = t_{AA}$ (hopping among A atoms) and may also be $t_{AB} = t_{BA}$ (hopping between A and B atoms) or t_{BB} (hopping among B atoms). The full Hamiltonian we consider is $\hat{H} = \hat{H}_0 + \hat{H}_{\text{imp}}$, where $\hat{H}_{\text{imp}}$ is the impurity Hamiltonian. We write the impurity Hamiltonian in terms of the set of impurity coordinates $\{\mathbf{R}_i\}$ and the hopping corrections $t' = t - t_{AB}$ and $t'' = \frac{1}{2}(t + t_{BB} - 2t_{AB})$,

$$\hat{H}_{\text{imp}} = \sum_{\mathbf{R}_i,\boldsymbol{\delta}} \Delta t(\mathbf{R}_i, \boldsymbol{\delta})[\hat{c}^\dagger(\mathbf{R}_i)\hat{c}(\mathbf{R}_i + \boldsymbol{\delta}) + \text{H.c.}], \tag{2}$$

with the hopping term $\Delta t(\mathbf{R}_i, \boldsymbol{\delta}) = t' - t'' \sum_{\mathbf{R}_j} \delta(\mathbf{R}_i + \boldsymbol{\delta} - \mathbf{R}_j)$ replacing hoppings appropriately. In the standard treatment of impurities which implicitly assumes they do not neighbor each other on the lattice, we have simply $\Delta t(\mathbf{R}_i, \boldsymbol{\delta}) = t'$, i.e., a dependence of the neighboring lattice positions does not enter the hopping correction and only one sum over the impurity coordinates is needed in the impurity Hamiltonian. Note that even in a "simpler" case of $t_{AB} = t_{BB}$, this additional sum over lattice sites and the associated terms in the T-matrix expansion are still needed to correctly account for neighboring impurities.

III. SELF-ENERGY COMPUTATION

We study the effect of the impurities on the self-energy appearing in the single-particle Green's function. The imaginary part will reflect electron lifetime effects, and the real part will reflect energy renormalizations. The self-energy serves as a scheme to describe the disorder-averaged Green's function of a system in terms of the undressed Green's function of the "clean" system, $G_0(i\omega_n, \mathbf{k}) = \frac{1}{i\omega_n - \xi_{\mathbf{k}}}$ (where ω_n are the Matsubara frequencies), as follows:

$$\langle G(i\omega_n, \mathbf{k})\rangle_{\text{dis}} = \frac{1}{i\omega_n - \xi_{\mathbf{k}} - \Sigma(i\omega_n, \mathbf{k})} = \frac{1}{G_0^{-1} - \Sigma}. \tag{3}$$

Note that while many self-energy approximations depend only on (Matsubara) frequency, here, our approach yields a momentum-dependent term as well.

A. T-matrix approach

In this work we make use of the disorder-averaged T matrix. It can be written in terms of the impurity Hamiltonian and the undressed square lattice Green's function G_0 as an expansion in the hopping correction [19]:

$$\hat{T} = \langle \hat{H}_{\text{imp}} + \hat{H}_{\text{imp}}\hat{G}_0\hat{H}_{\text{imp}} + \cdots \rangle_{\text{dis}}. \tag{4}$$

Here, the disorder average $\langle\cdots\rangle_{\text{dis}}$ consists of an integration over the impurity coordinates, divided by the total area, i.e., an average over a uniform distribution. Averaging over a nonuniform probability distribution for the impurity locations would require more care, but we circumvent this in our discussion of impurity clusters. In what follows we will consider first the case of isolated impurities, which follows a standard treatment. Afterward, we consider the case of clustered impurities, which we are able to treat semiheuristically via the introduction of a clustering probability.

1. Isolated impurities

Operating under the assumption that impurities are isolated, i.e., impurities do not occupy neighboring sites, the self-energy can be taken as the number of impurities times the disorder-averaged T matrix for a single impurity, $\Sigma_{\mathbf{k}\mathbf{k}'} = N_{\text{imp}} T_{\mathbf{k}\mathbf{k}'}$. The infinite T-matrix series can be summed analytically, which we discuss in Appendix A. From this series, the

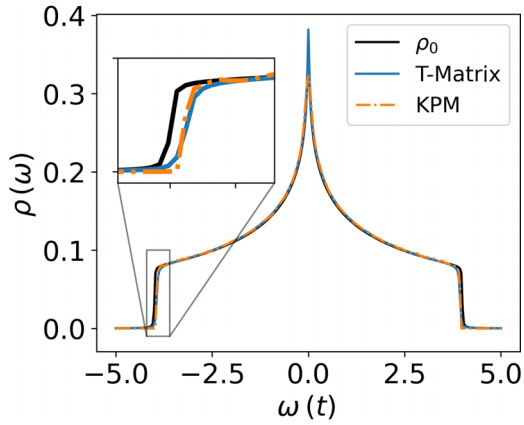

FIG. 1. Density of states for zero disorder ρ_0 and for disorder at a concentration of $\sigma = 0.1$. Hopping parameters used are $t = 1$ for the hoppings in the clean system and $t_{AB} = 0.9$ and $t_{BB} = 1.1$ for hoppings to impurity substitutions and between impurities. In the plot, we compare the kernel polynomial method (KPM) expansion on a finite model system of 401×401 lattice sites, which was computed in KWANT [21], with our T-matrix result (integrals were obtained with the QUADPY package [22]). Details of the model implementation for the finite system are discussed in Sec. III B.

T-matrix self-energy for isolated hopping impurities on the square lattice, in terms of the dimensionless hopping correction t' and the impurity concentration $\sigma = \frac{N_{\text{imp}}}{N}$, is given by

$$\Sigma_T(\omega, \mathbf{k}) = \frac{-t'\sigma}{d(\omega)}\left[a(\omega)\xi_{\mathbf{k}}^2 + b(\omega)\xi_{\mathbf{k}} + c(\omega)\right], \tag{5}$$

where we have defined the coefficients as

$$a(\omega) = t' g(\omega), \tag{6}$$

$$b(\omega) = -2[1 + t'(\omega g(\omega) - 1)], \tag{7}$$

$$c(\omega) = \omega t'[\omega g(\omega) + 1], \tag{8}$$

$$d(\omega) = t'^2[\omega g(\omega) - 1] - 2t'[\omega g(\omega) - 1] - 1. \tag{9}$$

The form of this T-matrix-based self-energy is similar to that obtained previously for an isolated impurity in graphene [9]. The term $g(\omega)$ appearing in the expressions above is given by $g(\omega) = \int_{\text{BZ}} \frac{d^2k}{(2\pi)^2} G_0(\omega, \mathbf{k})$, the integral of the undressed square lattice Green's function over the first Brillouin zone. The integral may be evaluated as an elliptic integral of the first kind, written in modulus form [20],

$$g(\omega) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \frac{d\mathbf{k}}{(2\pi)^2} \frac{1}{\omega + i\eta - \xi_{\mathbf{k}}} = \frac{2}{\pi\omega}\mathbf{K}\left(\frac{4t}{\omega}\right), \tag{10}$$

where the Matsubara frequency has been analytically continued to real frequencies as $\omega_n \to \omega + i\eta$ with $\eta = 0^+$. An interesting and measurable quantity that can be found from the T-matrix self-energy is the density of states (DOS), which we plot in Fig. 1.

From Fig. 1 we find that there is good agreement between the DOS as computed with the isolated-impurity T matrix and the random-impurity numerical result obtained in KWANT [21] with the kernel polynomial method (KPM), the details of which are given in Appendix B. It should be stressed that we have this good agreement despite the fact that our numeric computation allows for neighboring impurities with the different hopping parameter t_{BB}. This is in contrast to the T-matrix result, which assumed isolated impurities. Our result demonstrates the relatively wide range of applicability of the isolated-impurity assumption for random uniform distributions of impurities. This agreement is intuitive for small concentrations, as at low concentration in random uniform impurity distributions it is unlikely impurities will neighbor one another. We see good agreement here even up to the larger concentration of $\sigma = 0.1$ due to the small values of the corrections to the hoppings. We will later see that the most interesting clustering effect in this regime is not of the DOS.

One main difference from the computed DOS can be observed at $\omega = 0$, where the KPM result is suppressed compared with the isolated-impurity T-matrix result. Later analysis with our clustered-impurity method will show that while this type of suppression does occur due to clustering of impurities, the magnitude of the suppression is larger than would be expected from impurity clustering alone; a computation of the DOS performed using the KPM on an impurity distribution with zero clustering shows the same suppression, and as such, we can attribute this to a numerical quirk of the method rather than a sign of interesting physics.

2. *Clustered impurities*

To allow impurities to neighbor one another in our model, we must work with the full impurity Hamiltonian in Eq. (2). We take the first two terms in the T-matrix series expansion, which correspond to one- and twofold scattering from a single impurity. Terms corresponding to multiple-fold scattering from multiple distinct impurities or distinct clusters of neighboring impurities are discarded in accordance with the T-matrix approximation. Full summation of this T-matrix series, which at order n describes n-fold scattering by a single impurity or by up to n neighboring impurities in a cluster, is a much greater task in our case than in the case for isolated impurities. Thus, we restrict ourselves to the first two terms. We note that the reason for this is that the first nonvanishing imaginary contribution to the self-energy appears at quadratic order in the hopping corrections, and we therefore want to capture the impact of this contribution. For an analytic approach to systems where large impurity clusters and many-fold scattering are important, CPA methods [16,17] would be preferred.

To model the clustering of impurities, we employ a semi-heuristic approach. The T-matrix includes terms $\propto \sigma^2$, where σ is the probability of placing an impurity at a given lattice site. These terms, i.e., terms of higher than linear order in σ, correspond to scattering off neighboring impurities. In the random uniform distribution of impurities, the expected number of impurities occupying sites neighboring a known impurity (or any set of four sites, for that matter) is simply 4σ. In our approximation, we consider a clustering distribution of some kind, in which impurities are more likely to cluster than in the random uniform case, and make an appropriate replacement of the expected number of impurities neighboring a known impurity, $4\sigma \to p \times \langle N_{nn} \rangle$, where the statistical

parameter $p \in (0, 1)$ represents the proportion of impurities in the distribution which belong to clusters and $\langle N_{nn} \rangle \in (1, 4)$ is the expected number of impurity neighbors possessed by an impurity site known to belong to a cluster. For clusters of two impurities, this value is exactly $\langle N_{nn}^2 \rangle = 1$, as each site in a pair must neighbor a single other site, while for clusters of three it is $\langle N_{nn}^3 \rangle = \frac{4}{3}$ because, for all clusters of three neighboring lattice sites on the 2D square lattice, two of the sites neighbor one impurity while one site neighbors two. Meanwhile, in clusters of arbitrarily large size, we would have $\langle N_{nn}^\infty \rangle = 4$, with the assumption that clusters form compact domains with cluster area scaling faster than the cluster perimeter, which is energetically the valid assumption to make. As the numerics for this investigation use primarily small system sizes resulting in relatively small clusters, we simplify our equations by taking $\langle N_{nn} \rangle \approx 1$ in this paper. The general form with larger values of $\langle N_{nn} \rangle$ is not difficult to implement, but we choose instead to work only with the proportion of clustered impurities p in our numerical scheme. In applying this method to clustered impurities which follow some distribution, p is the general parameter representing the expected value of the number of impurity nearest neighbors of another impurity.

Let us now illustrate for our example how such a substitution is made by explicitly evaluating the first-order term in the T-matrix expansion, $\langle H_{\text{imp}} \rangle_{\text{dis}}$. This begins with evaluating the term in the random uniform distribution of impurities. Expressing this in terms of $\langle \mathbf{k}|H_{\text{imp}}|\mathbf{k}' \rangle$,

$$\begin{aligned} T^{(1)} &= \frac{1}{N^2} \sum_{\mathbf{R}_i, \mathbf{R}_j, \boldsymbol{\delta}} (e^{i\mathbf{k}'\cdot\boldsymbol{\delta}} + e^{-i\mathbf{k}\cdot\boldsymbol{\delta}}) \\ &\quad \times \int d^2R_i d^2R_j [t' - t''\delta(\mathbf{R}_i + \boldsymbol{\delta} - \mathbf{R}_j)] e^{i\mathbf{R}_i\cdot(\mathbf{k}'-\mathbf{k})} \\ &= \delta_{\mathbf{k},\mathbf{k}'} \frac{-2\xi_{\mathbf{k}}}{N^2} \left(N N_{\text{imp}} t' - N_{\text{imp}}^2 t'' \right) \\ &= -2\xi_{\mathbf{k}} \delta_{\mathbf{k},\mathbf{k}'} (\sigma t' - \sigma^2 t''), \end{aligned}$$

where the dependence on the number of impurities N_{imp} and the number of total sites in the system N may be reduced to their ratio, the impurity concentration σ.

The generalization to impurity clustering occurs with terms of order σ^2 and greater. In the term above, for example, σ^2 represents the probability that any one site and its neighbor in any one direction both host an impurity. In the random uniform distribution of impurities, these impurity occupancy probabilities are independently σ, leading to a proportionality with σ^2. We are interested, however, in clustered distributions, where impurity occupancy on a site and impurity occupancy of neighboring sites are not, *a priori*, independent. To model this, we keep the probability for impurities to occupy isolated sites as σ, but in the correction terms which correspond to the case of neighboring impurities, we assume that occupancies of these sites neighboring an existing impurity all occur independently with variable probability $\frac{p}{4}$, such that a random impurity belongs to a cluster with probability p. Using this simple scheme, the above term would be modified as $-2\xi_{\mathbf{k}}(\sigma t' - \sigma^2 t'') \to -2\xi_{\mathbf{k}}(\sigma t' - \frac{\sigma p}{4} t'')$.

Applying this modification to the impurity-averaged T matrix at second order, $T = \langle H_{\text{imp}} + H_{\text{imp}} G_0 H_{\text{imp}} \rangle_{\text{dis}}$, the self-energy may be written as follows:

$$\begin{aligned} \Sigma(\omega, \mathbf{k}) &= -\alpha\xi_{\mathbf{k}} \\ &\quad + \int_{\text{BZ}} \frac{d^2q}{(2\pi)^2} [\beta(\xi_{\mathbf{k}} + \xi_{\mathbf{q}})^2 - \gamma(\xi_{\mathbf{0}} + \xi_{\mathbf{k}+\mathbf{q}})] G_{\mathbf{q}}^0, \end{aligned} \tag{11}$$

with

$$\alpha = 2\sigma t' - \frac{\sigma p}{2} t'', \tag{12}$$

$$\beta = \sigma (t')^2 - \sigma p t' t'' + \frac{\sigma p^2}{4} (t'')^2, \tag{13}$$

$$\gamma = \sigma p (t'')^2. \tag{14}$$

As it turns out, the integral involved in Eq. (11) is again tractable in terms of elliptic integrals. We introduce the function $h(\omega)$, which is Eq. (3.148) in Ref. [23],

$$\begin{aligned} h(\omega) &= \int_{\text{BZ}} \frac{d^2q}{(2\pi)^2} \frac{\cos(q_x)}{\omega + i\eta - \xi_{\mathbf{q}}} \\ &= \frac{2}{\pi\omega} \left[\left(\frac{\omega}{t} + 2 \right) \mathbf{\Pi}\left(-\frac{4t}{\omega}, \frac{4t}{\omega} \right) - \left(\frac{\omega}{t} + 1 \right) \mathbf{K}\left(\frac{4t}{\omega} \right) \right], \end{aligned} \tag{15}$$

where $\mathbf{K}(k)$ and $\mathbf{\Pi}(n, k)$ are the complete elliptic integrals of the first and third kind, respectively, where an analytic continuation $\omega = \omega + i\eta$ is implicit. We have also kept the hopping parameter t symbolic in this expression to make the units clear, although, as mentioned, we set $t = 1$ for our numerics. In terms of $g(\omega)$ and the function $h(\omega)$, we can express the self-energy as

$$\begin{aligned} \Sigma(\omega, \mathbf{k}) &= -\alpha\xi_{\mathbf{k}} + \beta[(\omega + \xi_{\mathbf{k}})^2 g(\omega) - 2\xi_{\mathbf{k}} - \omega] \\ &\quad - \gamma[\xi_{\mathbf{0}} g(\omega) + \xi_{\mathbf{k}} h(\omega)]. \end{aligned} \tag{16}$$

The analytical result, Eq. (16), is the main result of this section. It is exact at the level of approximation at which we are working and contains the information about the clusters of the impurities from Eqs. (12)–(14).

B. Exact numerical treatment of clustered impurities

For a more complete investigation of the physics at hand and to corroborate the correctness of our treatment of impurity clustering, we also compute the self-energy of exact numerical realizations H of the Hamiltonian in Eq. (2). The Green's functions are computed as $G = [(\omega + i)\mathbb{1} - H]^{-1}$, where ω is the energy and η is a small numerical factor that was inserted for convergence, chosen to be $\eta = 0.01$.

In our computations we considered a lattice with $N = 61 \times 61$ sites. Inside the lattice we placed N_{imp} impurities, which corresponds to an impurity concentration $\sigma = N_{\text{imp}}/N$. Of the N_{imp} impurities a percentage p (called the clustering probability) was generated such that they are part of a cluster of impurities with random size. The remaining $(1 - p)N_{\text{imp}}$ impurities were placed randomly but isolated from other impurities or clusters. For more details on this algorithm, see Appendix C. This algorithm for generating impurity distributions was generated with the intent that it is nonspecific in the choice of cluster sizes in the same way that our analytic T-matrix approach is; that is, there is not a preference for a

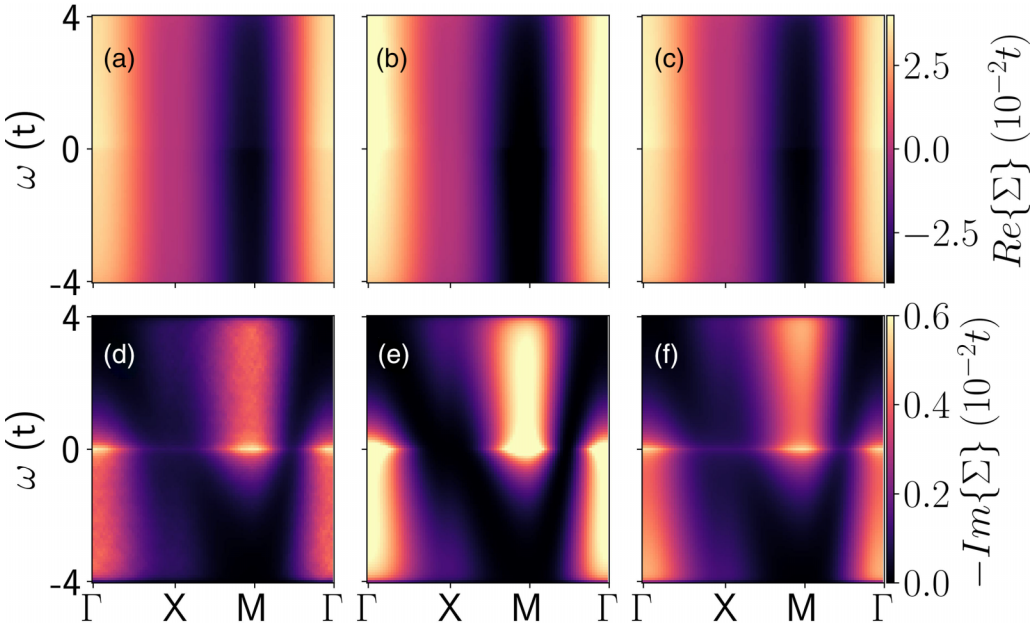

FIG. 2. The self-energy: real part (top row) and imaginary part (bottom row). (a) and (d) are the result of the numerical averaging of finite model Green's functions, while the other panels show the T-matrix results (b) and (e) neglecting and (c) and (f) including neighboring impurities. These plots are computed for parameters $(\sigma, t_{AB}, t_{BB}) = (0.05, 0.9, 1.1)$ at an impurity clustering of $p = 0.3$.

specific type of cluster, merely some fixed probability that each impurity neighbors another. This can also be expected to be true in various experimental scenarios, which is part of the reason that we consider this case.

To find the disorder-averaged self-energy we applied the following numerical procedure. For a desired concentration of σ and clustering probability p, a disorder-averaged Green's function was obtained by averaging $n_{\text{ave}} = 20$ disorder realizations. The averaged self-energy in position space was then obtained as $\Sigma = (G_0)^{-1} - \langle G \rangle_{\text{avg}}^{-1}$. We then transformed to an approximate crystal momentum basis with the discrete Fourier transform matrix,

$$U_{ij} = \frac{1}{\sqrt{N}} e^{i\mathbf{k}_i \cdot \mathbf{r}_j}, \tag{17}$$

where the self-energy was obtained as $\Sigma_{\mathbf{k}\mathbf{k}'} = U^\dagger \Sigma_{\mathbf{r}\mathbf{r}'} U$. The diagonal components, which dominate the numerically averaged self-energy, were then extracted for comparison with the T-matrix method. Off-diagonal elements correspond to scattering between crystal momentum eigenstates. While these elements may be relevant in a single disorder realization, they vanish analytically in the disorder-averaged framework and, likewise, are observed to vanish in our numerical disorder average. The motivation for this disorder average is self-averaging systems, as off-diagonal scattering is likewise negligible in these very large systems with homogeneous disorder.

IV. EFFECTS OF IMPURITY CLUSTERING

The exact numerical computation mentioned above provides us with an approximate self-energy for self-averaging systems hosting hopping impurities, characterized by the impurity density σ and clustering probability p. We may now investigate how clustering impacts the physics of the model. We perform this analysis with the same parameters, $t_{AB} = 0.9$ and $t_{BB} = 1.1$, as before. We do not discuss different values of (t_{AB}, t_{BB}) because for the regime where our approach is applicable the self-energy $\Sigma \propto (t'')^2$, which tells us that a change in parameters will mostly just increase the strength of an observed effect but not its nature.

First, we compare the self-energy for various approaches in Fig. 2. Note that for the purpose of visual clarity we have used $\eta = 0.05$ rather than $\eta = 0.01$ in the plot for the self-energy. In the perturbative regime, when t_{AB} and t_{BB} are both reasonably close to the original hopping t, we observe that the main effect of impurity clustering is imprinted on the imaginary part of the self-energy. In the real part it primarily leads to a renormalization of the bandwidth. It is dominated by the term at first order in the hopping corrections, which in the perturbative regime has a relatively small effect due to impurity clustering.

Next, in Fig. 3 we plot the spectral function and quasiparticle lifetime, computed from the self-energy as $A(\omega, \mathbf{k}) = -\frac{1}{\pi} \text{Im}\{G(\omega, \mathbf{k})\}$ and $\tau(\omega, \mathbf{k}) = \frac{-1}{2\,\text{Im}\{\Sigma\}}$. We find that the spectral function $A(\omega, \mathbf{k})$, much like the real part of the self-energy, changes very little due to impurity clustering. However, the quasiparticle lifetime is significantly different. Clustering of impurities here leads to significant suppression of quasiparticle lifetimes for most of the energy range.

We can visualize this effect even better if we examine the quasiparticle lifetime along the renormalized energy band of the model $\tau(\xi'_{\mathbf{k}}, \mathbf{k})$. In Fig. 4 we plot the quasiparticle lifetime for different values of the clustering probability p. Because the energy band of the model along which we want to compute the lifetime is renormalized by the disorder, we take this renormalization from the first order of the T matrix, $\xi'_{\mathbf{k}} = (1 + 2\sigma t' - \frac{\sigma p}{2} t'')\xi_{\mathbf{k}}$, which is the dominant real part

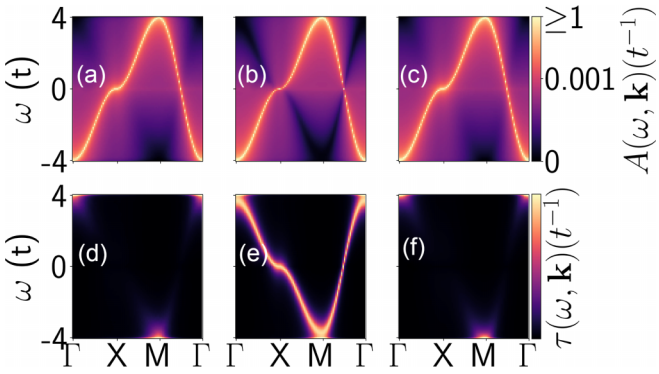

FIG. 3. Spectral function (top) and quasiparticle lifetime (bottom) for the three computation methods. (a) and (d) are the result of the numerical averaging of finite model Green's functions, while the other panels show the T-matrix results (b) and (e) neglecting and (c) and (f) including neighboring impurities. These plots are computed for parameters $(\sigma, t_{AB}, t_{BB}) = (0.05, 0.9, 1.1)$ at an impurity clustering of $p = 0.3$.

of the self-energy and thus the dominant contribution to the rescaling of the energy band. We observe that with increasing clustering probability p, there is a general flattening of the lifetime. Maxima, previously located between symmetry points M and Γ and at X (that is, specifically at the points where $\xi_{\mathbf{k}} = 0$), are significantly flattened, and we observe them splitting into two peaks. This effect does not depend on whether t_{BB} is larger or smaller than t, as the term responsible for these effects is $\propto (t'')^2$ and therefore depends on only the difference. An experimental result which may hold some relevance for this analysis is from Ref. [24], which studied optically excited states in disordered graphite and observed a similar phenomenon in which the lifetimes of states close to the Fermi energy were suppressed most rapidly. While our model differs significantly from this experiment, it is worth observing the phenomenological parallel.

We now return to an observation made previously during our discussion of the isolated-impurity approximation: at $\omega = 0$ the numerically computed DOS with random neighboring impurities was smaller than the value expected from the analytical result within the isolated-impurity approximation. To investigate this further and determine whether this may be a clustering-related effect, we numerically integrate the

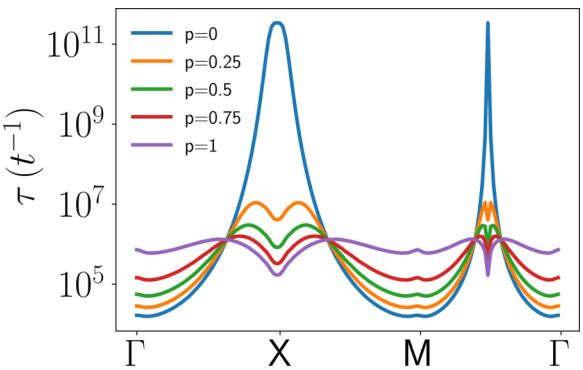

FIG. 4. Quasiparticle lifetime along renormalized band $\xi'_{\mathbf{k}}$, with varying cluster parameter p.

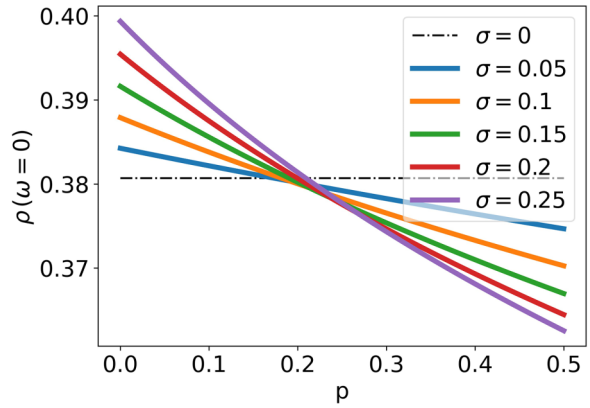

FIG. 5. Zero-energy density of states for different impurity concentrations as they vary with the clustering parameter.

Green's function as corrected with the impurity cluster self-energy at $\omega = 0$. The results are shown in Fig. 5. We observe that increasing clustering (larger clustering probability p) leads to a decrease in the density of states at $\omega = 0$. This decrease is again attributable to the term in the self-energy $\propto (t'')^2$ and will be observed with any values t_{AB} and t_{BB}. However, the magnitude of the decrease predicted to occur for larger clustering of impurities is smaller than the difference observed between the DOS curves in Fig. 1, suggesting that this particular discrepancy is related to the difference in the computation methods rather than representative of interesting physics.

Finally, we consider a long-standing feature of disorder models [14], localization. One may capture numerically how localized the states in a given physical system are with the so-called inverse participation ratio (IPR). To get an overall measure of the level of localization present in our system we sum the IPR for all states as

$$\mathrm{IPR} = \sum_{i,j} |\psi_j(\mathbf{x}_i)|^4. \tag{18}$$

To get a result valid for various impurity configurations we average the IPR over 10 realizations of the system. In Fig. 6 we plot the IPR for various values of impurity density σ and as a function of clustering probability p. We observe an

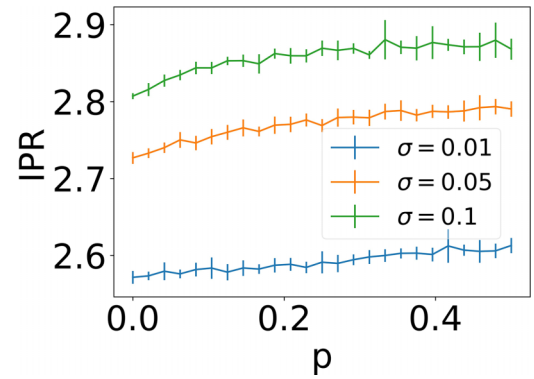

FIG. 6. Inverse participation ratio (IPR) varying with disorder concentration and clustering, computed numerically with an average of 10 systems. Larger IPR indicates more localized states. The error bars show the standard deviation of the numerical average.

increase in the IPR, indicating the degree of overall localization present in the states increases with an increase in clustering probability p.

V. CONCLUSION

Motivated by the known positional correlations of impurities which arise in the configuration of impurities in materials, we adopted multiple approaches to investigate the electronic properties of a square lattice with nearest-neighbor hopping disorder. In order to account for the clustering effect of impurities in a material, we used a semiheuristic approach that introduces a clustering probability p to determine the likelihood of having an impurity next to a given impurity. Working with the impurity-averaged T-matrix with this approach, we were able to obtain an important analytical result, Eq. (16), that takes into account the impurity distributions with statistical information contained in Eqs. (12)–(14).

We found good agreement between the numerical and analytic approaches taken for the perturbative regime of hopping disorder. A particularly interesting effect is the suppression of quasiparticle lifetime, with single peaks originally at $\omega = 0$ splitting into two subpeaks with an increase in impurity clustering, as shown in Fig. 4.

Additionally, computing the inverse participation ratio demonstrated an increase in localization for more clustering of impurities, as seen in Fig. 6. These combined numerical and analytical results, which are expressed through the electron self-energy, will appear in any quantity derived from the single-particle Green's function, including electrical and thermal transport, optical conductivity, and photoemission. (Note the transport and conductivity will depend on pairs of single-particle Green's functions, whereas photoemission depends only on a single Green's function.) Our work thus provides an important link between microscopic impurity distributions in a material and experimental observables [25]. This should help elucidate the more complex effects of impurity clustering which are not possible to describe with simple isolated-impurity approximations.

The data and code supporting the results of this paper are available from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support from the National Science Foundation through the Center for Dynamics and Control of Materials, an NSF MRSEC, under Cooperative Agreement No. DMR-1720595 and NSF Grant No. DMR-2114825. M.V. gratefully acknowledges the support provided by the Deanship of Research Oversight and Coordination (DROC) at King Fahd University of Petroleum and Minerals (KFUPM) for funding this work through Exploratory Research Grant No. ER221002. G.A.F. gratefully acknowledges support from the Alexander von Humboldt Foundation.

APPENDIX A: SINGLE-IMPURITY T MATRIX

For a single impurity at coordinate $\mathbf{R}$, the disorder component of the Hamiltonian can be written in the crystal momentum basis as

$$\langle \mathbf{k}|\hat{H}_{\text{imp}}|\mathbf{k}'\rangle = -\frac{t'}{N} e^{i\mathbf{R}\cdot(\mathbf{k}'-\mathbf{k})}(\xi_{\mathbf{k}} + \xi_{\mathbf{k}'}). \tag{A1}$$

The n-fold scattering terms in the T-matrix series, when written as integrals over $n-1$ free crystal momenta, form a recurrence. Following the evaluation of the disorder average, each term becomes

$$T_n = \sigma(-t')^n \int \left(\prod_{i=1}^{n-1} \frac{d^2k_i}{(2\pi)^2}\right) \times \cdots \times \left((\xi_{\mathbf{k_0}} + \xi_{\mathbf{k_1}}) \prod_{m=1}^{n-1} (\xi_{\mathbf{k_m}} + \xi_{\mathbf{k_{m+1}}}) G^0_{\mathbf{k_m}}\right). \tag{A2}$$

We find all such integrals may be evaluated in terms of the integral of the undressed square lattice Green's function over the first Brillouin zone,

$$g_{\square}(\omega) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \frac{d\mathbf{k}}{(2\pi)^2} \frac{1}{\omega + i\eta - \xi_{\mathbf{k}}} = \frac{2}{\pi\omega}\mathbf{K}\left(\frac{4t}{\omega}\right), \tag{A3}$$

which is evaluated as a complete elliptic integral of the first kind $\mathbf{K}$. The nth term in the T matrix, diagonal in the crystal momentum, is proportional to $\alpha_n\xi^2_{\mathbf{k}} + (\beta_n + \gamma_n)\xi_{\mathbf{k}} + \delta_n$, where the coefficients satisfy the following system of recurrence equations:

$$\alpha_n = (\omega g_{\square} - 1)\alpha_{n-1} + g_{\square}\gamma_{n-1}, \tag{A4a}$$

$$\beta_n = (\omega g_{\square} - 1)\beta_{n-1} + g_{\square}\delta_{n-1}, \tag{A4b}$$

$$\gamma_n = (\omega g_{\square} - 1)\gamma_{n-1} + \omega(\omega g_{\square} - 1)\alpha_{n-1}, \tag{A4c}$$

$$\delta_n = (\omega g_{\square} - 1)\delta_{n-1} + \omega(\omega g_{\square} - 1)\beta_{n-1}. \tag{A4d}$$

Using the easily computed values of these coefficients at $n = 0, 1$, exact expressions for the coefficients as a function of n can be obtained with the help of *Mathematica*'s RSOLVE. Finally, summing these coefficients to $n = \infty$ yields the form presented in the main text.

APPENDIX B: KERNEL POLYNOMIAL METHOD

One of the numerical tools we made use of in investigating disorder effects was the kernel polynomial method (KPM), as built into KWANT. Here, we briefly review this method, which allows efficient determination of the density of states. The algorithm in KWANT is based on Ref. [26]. For completeness, we summarize this method briefly.

We are interested in computing the spectral density $\rho(E)$ of the Hamiltonian operator $\hat{H}$, i.e., for the eigenenergies E_K,

$$\rho(E) = \frac{1}{N}\sum_{k=1}^{N} \delta(E - E_k). \tag{B1}$$

This is accomplished by way of an expansion in the Chebyshev polynomials T_n of $\hat{H}$, which may be implemented easily algorithmically because of the recursion relating the polynomials of successive orders. To perform this expansion, the Hamiltonian is first rescaled such that $E_k \in [-1, 1]\ \forall k$. This expansion, in terms of the rescaled energy argument and

Hamiltonian, takes the form

$$\rho(E) = \frac{1}{\pi\sqrt{1-E^2}}\left(\mu_0 + 2\sum_{n=1}^{\infty}\mu_n T_n(E)\right), \tag{B2}$$

where the coefficients μ_n are computed in terms of the rescaled Hamiltonian,

$$\mu_n = \frac{1}{N}\,\mathrm{Tr}[T_N(\hat{H})]. \tag{B3}$$

Without delving into the algorithmic details that make this method so efficient, we can still conclude that the KPM is of great use to us here—while matrix inversion and eigendecomposition of $n\times n$ matrices scale as around $O(n^3)$, KPM methods can be as good as $O(n)$ and at least $O(n^2)$.

APPENDIX C: CLUSTER GENERATION ALGORITHM

To assess the validity of the clustered-impurity T matrix, we generated distributions of impurities which had a specified proportion of impurities belonging to clusters. To generate such distributions, we implemented an algorithm which counted the number of said impurities as follows.

The core trick of the impurity-generating method was developing an efficient method for identifying which sites neighbor impurities and which of these are already occupied by impurities. For a square lattice of dimensions $N\times N$, we construct an impurity matrix $\mathcal{I}$, where $\mathcal{I}_{ij} = 1$ implies the site indexed (i, j) on the lattice hosts an impurity, while $\mathcal{I}_{ij} = 0$ implies the lattice site does not host an impurity. Consider then a 3×3 kernel matrix,

$$K = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{C1}$$

Now, if we perform a matrix convolution $K * \mathcal{I}$, padding $\mathcal{I}$ with rows and columns of zeros to yield an $N+2\times N+2$ matrix, we can construct a new matrix $\mathcal{N}$, which in a fashion analogous to $\mathcal{I}$ is defined such that $\mathcal{N}_{ij} = 1$ implies the site indexed (i, j) neighbors sites which host impurities and $\mathcal{N}_{ij} = 0$ implies the site indexed (i, j) does not neighbor sites which host impurities,

$$\mathcal{N}_{ij} = \begin{cases} 1 & \text{if } (K * \mathcal{I})_{ij} > 0, \\ 0 & \text{if } (K * \mathcal{I})_{ij} = 0. \end{cases} \tag{C2}$$

From there, a Hadamard (i.e., elementwise) product of $\mathcal{N}$ and $\mathcal{I}$ or $\neg\mathcal{I}$ can be taken to yield a final matrix with elements corresponding to impurities which neighbor other impurities or empty sites which neighbor impurities. In addition, logical operations can be added simply at any point in the process to alternatively return only isolated impurities or only sites which do not host impurities and do not neighbor impurities. It is additionally possible to use a larger kernel matrix or to perform these operations in higher dimensions, such that this process as described could enumerate all sites on a hypercubic lattice which have nth nearest-neighbor impurities.

In terms of actually generating the distributions, we applied the preceding method in a simple manner. Given the $N\times N$ sample and the desired impurity concentration and clustering proportion σ and p, we selectively generate σpN^2 clustered impurities and $\sigma(1-p)N^2$ isolated impurities. First, a random set of impurity elements are initially assigned, i.e., by checking whether a random number on [0,1] is greater than the desired proportion of initial impurities. Our algorithm used a random number of initial impurities, between 1 and $N_{\mathrm{imp}}/2$; the choice in initialization will affect the sizes of the impurity clusters. Then, using the above process, impurities are added to the system only to unoccupied sites which neighbor existing impurities, until the desired number of clustered impurities is obtained. Following this, the remaining isolated impurities are added to sites which do not neighbor impurities, one at a time.

Regardless of the choice in algorithm for placing the impurities themselves, the kernel convolution method is an efficient way to enumerate sites which are adjacent to existing impurities.

[1] X. Zou and B. I. Yakobson, An open canvas—2D materials with defects, disorder, and functionality, Acc. Chem. Res. **48**, 73 (2015).

[2] J. Kullgren, M. J. Wolf, P. D. Mitev, K. Hermansson, and W. J. Briels, DFT-based Monte Carlo simulations of impurity clustering at CeO_2 (111), J. Phys. Chem. C **121**, 15127 (2017).

[3] D. F. Förster, T. O. Wehling, S. Schumacher, A. Rosch, and T. Michely, Phase coexistence of clusters and islands: Europium on graphene, New J. Phys. **14**, 023022 (2012).

[4] J. Furthmüller, F. Hachenberg, A. Schleife, D. Rogers, F. Hosseini Teherani, and F. Bechstedt, Clustering of N impurities in ZnO, Appl. Phys. Lett. **100**, 022107 (2012).

[5] Y. Wang, S. Xiao, X. Cai, W. Bao, J. Reutt-Robey, and M. S. Fuhrer, Electronic transport properties of Ir-decorated graphene, Sci. Rep. **5**, 15764 (2015).

[6] L. Lin, J. Li, Q. Yuan, Q. Li, J. Zhang, L. Sun, D. Rui, Z. Chen, K. Jia, M. Wang, Y. Zhang, M. H. Rummeli, N. Kang, H. Q. Xu, F. Ding, H. Peng, and Z. Liu, Nitrogen cluster doping for high-mobility/conductivity graphene films with millimeter-sized domains, Sci. Adv. **5**, eaaw8337 (2019).

[7] G. R. Berdiyorov, H. Bahlouli, and F. M. Peeters, Effect of substitutional impurities on the electronic transport properties of graphene, Phys. E (Amsterdam, Neth.) **84**, 22 (2016).

[8] N. M. R. Peres, F. Guinea, and A. H. Castro Neto, Electronic properties of disordered two-dimensional carbon, Phys. Rev. B **73**, 125411 (2006).

[9] N. M. R. Peres, F. D. Klironomos, S.-W. Tsai, J. R. Santos, J. M. B. Lopes dos Santos, and A. H. Castro Neto, Electron waves in chemically substituted graphene, Europhys. Lett. **80**, 67007 (2007).

[10] V. M. Pereira, J. M. B. Lopes dos Santos, and A. H. Castro Neto, Modeling disorder in graphene, Phys. Rev. B **77**, 115109 (2008).

[11] M. Javan, R. Jorjani, and A. R. Soltani, Theoretical study of nitrogen, boron, and co-doped (B, N) armchair graphene nanoribbons, J. Mol. Model. **26**, 64 (2020).

[12] N. Sule, S. C. Hagness, and I. Knezevic, Clustered impurities and carrier transport in supported graphene, Phys. Rev. B **89**, 165402 (2014).
[13] M. I. Katsnelson, F. Guinea, and A. K. Geim, Scattering of electrons in graphene by clusters of impurities, Phys. Rev. B **79**, 195426 (2009).
[14] P. W. Anderson, Localized magnetic states in metals, Phys. Rev. **124**, 41 (1961).
[15] E. G. Kostadinova, C. D. Liaw, A. S. Hering, A. Cameron, F. Guyton, L. S. Matthews, and T. W. Hyde, Spectral approach to transport in a two-dimensional honeycomb lattice with substitutional disorder, Phys. Rev. B **99**, 024115 (2019).
[16] R. Mills and P. Ratanavararaksa, Analytic approximation for substitutional alloys, Phys. Rev. B **18**, 5291 (1978).
[17] H. W. Diehl and P. L. Leath, Pseudofermion approach to binary disordered systems. I. Effective-medium theories and the augmented-space method, Phys. Rev. B **19**, 587 (1979).
[18] W. P. Dumke, Optical transitions involving impurities in semiconductors, Phys. Rev. **132**, 1998 (1963).
[19] A. Altland and B. Simons, *Condensed Matter Field Theory*, 2nd ed. (Cambridge University Press, Cambridge, 2010).
[20] E. Kogan and G. Gumbs, Green's functions and DOS for some 2D lattices, Graphene **10**, 1 (2021).
[21] C. W. Groth, M. Wimmer, A. R. Akhmerov, and X. Waintal, Kwant: A software package for quantum transport, New J. Phys. **16**, 063065 (2014).
[22] N. Schlömer, N. Papior, D. Arnold, J. Blechta, and R. Zetter, nschloe/quadpy: None, https://zenodo.org/record/5541216.
[23] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., edited by A. Jeffrey and D. Zwillinger (Elsevier/Academic Press, Amsterdam, 2007).
[24] G. Moos, C. Gahl, R. Fasel, M. Wolf, and T. Hertel, Anisotropy of Quasiparticle Lifetimes and the Role of Disorder in Graphite from Ultrafast Time-Resolved Photoemission Spectroscopy, Phys. Rev. Lett. **87**, 267402 (2001).
[25] J. G. Nedell, J. Spector, A. Abbout, M. Vogl, and G. A. Fiete, Deep learning of deformation-dependent conductance in thin films: Nanobubbles in graphene, Phys. Rev. B **105**, 075425 (2022).
[26] A. Weiße, G. Wellein, A. Alvermann, and H. Fehske, The kernel polynomial method, Rev. Mod. Phys. **78**, 275 (2006).